PN-ABI-614
77440

FINAL REPORT FOR THE UNITED STATES DEPARTMENT OF STATE AGENCY FOR INTERNATIONAL DEVELOPMENT

PROJECT No.: 7.393

GRANT No.: DPE-5542-G-SS-7060-00

PROJECT TITLE: "DEVELOPMENT OF MOLECULAR GENETIC MARKERS IN *Phaseolus* AS BREEDING TOOLS: DISEASE RESISTANCE"

PRINCIPAL INVESTIGATOR:
C. EDUARDO VALLEJOS
1255 FIFIELD HALL
DEPARTMENT OF VEGETABLE CROPS
UNIVERSITY OF FLORIDA
GAINESVILLE, FL 32611

JUN 2 2 1992

RESEARCH ACCOMPLISHMENTS

The specific objectives of the proposal were to:

1) Construct a saturation linkage map in *Phaseolus vulgaris*, the common bean, using isozymes and DNA restriction fragment length polymorphisms (RFLPs), and
2) To use these markers to map and tag, by linkage analysis, genes conferring resistance to common bacterial blight of beans (CBB) and to bean common mosaic virus (BCMV).

The work plan for this project was presented in detail at the International Bean Trials Workshop held at CIAT in 1987:

Vallejos CE & Chase CD (1988) Development of DNA markers for *Phaseolus vulgaris* genetics and breeding. Proceedings of the International Bean Trials Workshop. International Center for Tropical Agriculture (CIAT). Pp 397-409.

A genomic library of size selected *Pst*I fragments was constructed and later used as the source of probes for the mapping project. A random sample of clones from this library were used to assess the extent of polymorphism in the common bean. It was determined that the clones were very useful in detecting polymorphisms among a set of different bean genotypes that were provided by our collaborators at CIAT. These probes were capable of discriminating between the mesoamerican and andean gene pools. The description of this work was published in 1991:

Chase CD, Ortega VM & Vallejos CE (1991) DNA restriction fragment length polymorphisms correlate with isozyme diversity in *Phaseolus vulgaris* L. Theor. Appl. Genet. 81: 806-811.

The survey mentioned above indicated that maximum allelic differences could be detected between the mesoamerican breeding line 'XR-235-1-1' and the andean cultivar 'Calima.' For this reason, a backcross population between these two genotypes was used as the mapping population. These genotypes also differ dramatically in seed size and in their ability to respond to *Xanthomonas campestris* f. sp. *phaseoli*, the bacterial pathogen that causes CBB. 'XR-235-1-1' is a small seeded (ca. 250 mg) line developed for resistance to CBB. On the

other hand, 'Calima' is a large seeded (ca. 500 mg) cultivar susceptible to CBB. These genotypes were first characterized for polymorphisms at isozyme and seed storage protein loci, and later with RFLP markers.

The backcross population was variable for seed size and prompted us to weigh and measure the seeds. Isozyme analysis revealed the presence of a major gene that controls seeds size in beans. This finding is relevant in bean research because it has been determined that seed size and yield are inversely correlated. Thus, by detecting genes involved in the control of seed size it is possible to detected some of the components of yield. This work was also published in 1991:

Vallejos CE & Chase CD (1991) Linkage between isozyme markers and a locus affecting seed size in *Phaseolus vulgaris* L. Theor. Appl. Genet. 81: 413-419.

Further analysis of seed storage proteins, isozymes and a locus responsible for seed pigmentation revealed a large linkage group containing the phaseolin locus (*Pha*). This work was also published in 1991:

Vallejos CE & Chase CD (1991) Extended linkage map for the phaseolin linkage group of *Phaseolus vulgaris* L. Theor. Appl. Genet. 82: 353-357.

The bean *Pst*I genomic library was screened to identify clones that revealed polymorphisms between the parental genotypes with at least one of four restriction enzymes. Probes known to detect polymorphic loci were hybridized to genomic blots of the segregating progeny. Segregation data were collected for 245 markers. Linkage analysis of these markers was performed using the software program **Mapmaker** (Whitehead Institute for Biological Research/ MIT). The segregating markers were assorted into 11 linkage groups and encompass 960 cM of the bean genome. We have estimated with partial linkage data that the size of the bean genome is approximately 1200 cM in length. The construction of the linkage map has been described in detail in a publication that will appear in the July 1992 issue of the scientific journal "Genetics:"

Vallejos CE, Sakiyama NS & Chase CD (1992) A molecular marker-based linkage map of *Phaseolus vulgaris* L. Genetics 131: (in press).

A total of 216 clones of mapped bean sequences have been transferred to the Biotechnology Research Unit of CIAT in Cali, Colombia. CIAT will be the repository of the library and will also be in charge of distributing the clones to researchers around the world interested in bean genetics and breeding.

We are continuing to add markers to the bean map. A new set of markers called randomly amplified DNA polymorphisms (RAPD) are being incorporated into the map. I have recently submitted for publication a chapter describing the bean linkage map (See attachment); this chapter will appear in a book dedicated to molecular markers in plant breeding edited by Dr. I. Vasil and Dr. R. Phillips:

Vallejos CE (Submitted) *Phaseolus vulgaris*: the common bean. In: DNA-based markers in plants. Eds I. Vasil & R. Phillips. Kluwer Academic Publishers.

Although a chromosome segment associated with an isozyme was identified as carrier of a gene affecting seed size, a complete analysis of the genome with the entire set of markers is still pending. Upon completion of this work, a copy of the manuscript describing this analysis will be sent to AID as an addendum to this report.

Unfortunately, we were unable to accomplish the second objective of the proposal. the construction of the linkage map took longer than anticipated. Construction of the map was started with the first backcross progeny in 1987. Too many individual plants of this progeny died before additional DNA samples were collected. After performing linkage analysis for 110 markers with a very small progeny (n=36), it was decided to analyze a new and larger progeny. Furthermore, we found that the breeding line 'XR-235-1-1,' chosen for its resistance to CBB, had not been fully characterized for its response to inoculation with the pathogen. We have found, as part of a different project now, that susceptibility in leaves of 'Calima' is developmentally regulated, and that resistance in 'XR-235-1-1' is temperature sensitive. The mapping of genes involved in resistance to CBB and those involved in resistance to BCMV are being supported by two different grants at this time.

This work has been presented at several national and international meetings:

International Bean Trials Workshop
Title: "Development of DNA markers for *Phaseolus vulgaris* genetics and breeding"
Oral presentation by C. E. Vallejos
Authors: C. E. Vallejos and C. D. Chase
Place: CIAT, Cali, Colombia
Date: October 12-16, 1987

Bean Improvement Cooperative 1987 Meeting
Title: "Development of DNA markers in *Phaseolus vulgaris*
Oral presentation by C. E. Vallejos
Authors: C. E. Vallejos, C. D. Chase
Place: Denver, Colorado
Date: October 27-29, 1987

III Congresso Brasileiro de Fisiologia Vegetal. Simposio de Biotecnologia de Plantas
Title: "The Use of Molecular Markers for Bean Breeding and Genetics"
Oral presentation by C. E. Vallejos, invited speaker
Place: Viçosa, Minas Gerais, Brazil
Date: February 24 - March 1, 1991

Biological Nitrogen Fixation. Networking Workshop
Title: "Development of Molecular Genetic Markers in *Phaseolus* as Breeding Tools: Disease Resistance"
Oral presentation by C. E. Vallejos
Place: Banff, Canada
Date: September 1-6, 1991

The International Society for Plant Molecular Biology. Third International Congress
Title: "The *Phaseolus vulgaris* Linkage Map"
Poster presentation by: C. E. Vallejos
Authors: Sakiyama, N. S., C. D. Chase, C. E. Vallejos
Place: Tucson, Arizona

Date: October 1-6, 1991

Bean Improvement Cooperative 1991 Annual Meeting
Title: "Construction of the *Phaseolus vulgaris* Linkage Map
Oral presentation by: C. E. Vallejos
Authors: Sakiyama, N. S., C. D. Chase, C. E. Vallejos
Place: Lincoln, Nebraska
Date: November 4-6, 1991

TECHNOLOGY TRANSFER FROM UF TO CIAT

During the term of the grant several members of the Biotechnology Research Unit (BRU) and other units from CIAT have visited my lab at the University of Florida in Gainesville. Dr. William Roca, Director of BRU visited my lab in April of 1988. Dr. Daniel Debouk also visited my lab (6/10 to 12/88), he was the Curator of the Bean Germ Plasm Collection at CIAT at that time. Mr. Hernando Ramirez, a lab technician in BRU, visited my lab for two months (7/1 to 9/2, 1988). During his visit Mr Ramirez was trained on all techniques related to RFLP analysis, these included: plant and plasmid DNA isolation, restriction enzyme digestions, nucleic acid fractionation by agarose gel electrophoresis, Southern blotting, DNA labeling, hybridization and autoradiography. Both Dr. Joseph Thome, staff member of BRU, and Mr. Alejandro Calderon, a Research Associate at the BRU, have also visited my lab in June and July of 1989 to discuss some techniques and future applications and approaches to research with the aid of RFLP markers. Today, BRU has trained personnel and modern lab facilities fully equipped for RFLP analysis.

I have visited CIAT in October of 1987 and in September of 1989. During both visits I have discussed with many staff members about the technology itself and plant breeding approaches that can be taken using DNA markers. Dr. Ney Sakiyama, the post-doctoral associate from Brazil that conducted most of the lab work to construct the map, also visited CIAT in September of 1989. He stayed at CIAT for a longer period. During his visit, Dr. Sakiyama shared information with staff members in the BRU about the use of the polymerase chain reaction (PCR) to amplify insert DNA, and also to share some of the "lab tricks" he had learned at the University of Florida.

Throughout the duration of the grant we have supplied CIAT with mapped

clones. We have now completed the transfer of 216 bean genomic clones to the BRU of CIAT for internal use and distribution.

RESEARCH PERSONNEL

C. Eduardo Vallejos, Associate Professor
Christine D. Chase, Associate Professor
Ney S. Sakiyama, Post-doctoral Research Associate
Mr. Samuel Camp, Biological Scientist
Mr. Victor M. Ortega, Biological Scientist
Ms. Cassia Sakiyama, Laboratory Assistant

ADDENDA

Development of DNA markers for *Phaseolus vulgaris* genetics and breeding

C. Eduardo Vallejos*
Christine D. Chase*

Introduction

Phaseolus vulgaris beans are one of the main sources of proteins in lesser developed countries. Current average yields of 500 kg/ha in these countries are low compared to onfarm potential yields of 3000 kg/ha. To a great extent, the difference between actual and potential yields can be attributed to environmental and biotic stresses.

Significant increases in yield can be effected by exploiting CIAT's rich germplasm collection of almost 23,000 working accessions. A successful global bean-breeding program would include the identification of traits of economic importance, and the incorporation of these traits into commercial cultivars. These objectives are complicated by strong regional preferences for bean type according to shape, color, and size. Nevertheless, an efficient implementation of this program can be accomplished with the use of molecular markers, that is, with isozymes and DNA restriction fragments.

Unlike morphological markers, molecular markers possess a number of features that make them very useful in plant breeding and genetics (Beckman and Soller, 1983; Burr et al., 1983; Tanksley, 1983). Molecular markers are naturally occurring—variation does not have to be induced artificially. They segregate in a codominant fashion, and the exact genotype of a segregant can be determined. Their expression as isozymes or presence as DNA restriction fragments are not influenced by environment, nor are they subject to epistatic interactions. Segregation of an almost infinite number of these markers can therefore be studied in a single progeny, allowing study of the entire genome.

* Department of Vegetable Crops, University of Florida, Gainesville, FL, USA.

Isozyme markers have been used to tag genes for qualitative traits via linkage analysis. Such traits include nematode resistance (Rick and Fobes, 1974), nuclear male sterility (Tanksley et al., 1984), self-incompatibility (Tanksley and Loaiza-Figueroa, 1985) in tomato, bean yellow mosaic virus resistance (Weeden et al., 1984), and resistance to *Fusarium* race 1 (Hunt and Barnes, 1982) in peas. Isozymes have also been used to detect quantitative trait loci involved in morphological characteristics (Tanksley et al., 1982), cold tolerance (Vallejos and Tanksley, 1983; Zamir et al., 1982), and yield (Pollak et al., 1984). In addition, these markers can be used as selection criteria for the expedient recovery of recurrent parent alleles in a backcross program (Tanksley and Rick, 1980) or in effecting high rates of fixation in a pedigree breeding program (Mendlinger, 1983).

Unfortunately, the usefulness of isozymes is restricted by their limited number. They represent a biased sample of the genome. To be useful as genetic markers, isozymes must show detectable electrophoretic variation via activity-staining. Excluded are those enzymes with stainable activity that are present in quantities below the limits of detection. Also excluded are enzymes not amenable to activity-staining after electrophoresis, and all structural proteins. These limitations prevent the exclusive use of isozymes to probe entire genomes.

To overcome these limitations, DNA restriction fragments have been proposed as additional genetic markers. To cover the entire genome, linkage maps saturated with molecular markers have been constructed for tomato (Bernatzky and Tanksley, 1986; Helentjaris et al., 1986), maize (Helentjaris et al., 1986), and lettuce (Landry et al., 1987). DNA restriction fragment loci have been used to locate specific sequences in the genome (Vallejos et al., 1986), and map genes controlling for disease resistance (Landry et al., 1987).

The main objective of our program is the construction of a saturation linkage map of DNA markers for *Phaseolus vulgaris*. Three basic steps have been taken to accomplish this objective. First, segregation progenies were obtained between two distinct genotypes: Calima, a cultivar adapted to Colombia and East Africa and obtained from Dr. W. Roca, CIAT, and XR-235-1, a *Xanthomonas campestris* resistant breeding line derived from the interspecific cross *P. vulgaris* x *P. coccineus* (Freytag et al., 1982) —seeds were obtained from Dr. M. J. Bassett, University of Florida.

Second, genomic and cDNA libraries have been constructed and are being screened for clones that reveal DNA restriction fragment length polymorphisms between the parents. Screening is accomplished by digesting DNAs from both parental lines with a number of restriction

enzymes and probing the digests with individual clones from our libraries. This has allowed the identification of restriction enzyme-clone combinations that reveal polymorphism between the parents.

Finally, as appropriate enzyme-clone combinations are identified, DNA from the progeny are digested with selected restriction enzymes and probed with appropriate clones. Once probing with a number of clones has been completed, results will be used for the segregation and linkage analyses that will allow the construction of the map. What follows is a detailed description of the methodology used to accomplish our objective.

DNA Restriction Fragment Length Polymorphisms (RFLPs)

DNA isolation

Plant cells have three genomes: mitochondrial (mt), chloroplast (cp), and nuclear (n). The organelle genomes are small and simple. They range in size from 0.85×10^5 to 1.9×10^5 base pairs (bp) for chloroplasts (Whitfeld and Bottomley, 1983), and 0.14×10^6 bp for mitochondria (Pring and Lonsdale, 1985). In contrast, nuclear genomes are large and more complex. They vary in size from 2×10^8 to 8×10^{10} bp (Bennett and Smith, 1976). The size of the common bean genome has been estimated at 1.6×10^9 bp (Ayonoadu, 1974). It has been calculated that only 5%-10% of higher plant genomes is expressed—this fraction is present in low copy number and distributed throughout the genome (Flavell, 1980). Polymorphisms among DNA fragments present in low copy number are used in the construction of a linkage map.

The first step taken for the analysis of RFLPs is the isolation and purification of nuclear DNA. Young bean leaf samples are collected and immediately frozen in liquid nitrogen (-196 °C). Samples can be stored at -70 °C for an indefinite period of time. Extraction of DNA is started by grinding 10-40 g of sample with liquid nitrogen in a mortar. The tissue powder is added to 4 volumes of chilled (4 °C) extraction buffer (tris-HCl, pH 7.9, 0.1 M/Na_2EDTA 10 mM/sucrose 0.5 M/β-mercaptoethanol 1%)[1] added before use.

1. M = molar solution
EDTA = ethylenediaminetetraacetic acid.

The sample is further homogenized in a polytron homogenizer for 30-40 seconds at a medium setting. The homogenate is filtered through four layers of cheesecloth and one layer of miracloth. Triton X-100 is added to a final concentration of 0.5%. This detergent dissolves the organelles' membranes, releasing their DNA, while nuclear structures are not significantly disrupted.

The homogenate is centrifuged at 4000 x g_{avg} for 15 minutes. The supernatant is discarded and the nuclear pellet is resuspended in 1 ml of resuspension buffer (tris-HCl, pH 7.9, 0.1 M/Na_2EDTA 20 mM) for every 10 g of tissue. An equal volume of 2X lysis buffer (tris-HCl, pH 7.9, 0.1 M/NaCl 2 M/CTAB 2%)[2] is added to the resuspension and the mixture is incubated at 65 °C for 30 minutes with gentle inversion of the tubes every 5 minutes. At the end of the incubation period, 1/2-volume of chloroform (chloroform to octanol, 24:1) is added, and tubes are inverted gently about 20 times.

Phases are separated by centrifugation in a swinging bucket rotor at 4000 x g_{avg} for 20 minutes. The chloroform is in the lower phase, containing all lipid compounds, including chlorophyll. The interphase contains denatured proteins and starch. The upper aqueous phase, containing nucleic acids (DNA and RNA) and other water-soluble compounds, is transferred to another tube and the nucleic acids are precipitated by adding 2/3-volume of isopropanol (at -20 °C).

The precipitable is removed with a glass hook and immersed in ethanol (76%)/Na acetate (0.2 M) for at least 30 minutes. It can also be stored in this solution for longer periods of time. After a short dip (20 seconds) into ethanol (76%)/ammonium acetate (10 mM), the pellet is transferred to a new tube, air-dried, and dissolved in tris-EDTA buffer (tris-HCl, pH 8.0, 10 mM/EDTA 1 mM). The resuspended DNA occasionally contains some starch which can be removed by centrifugation.

DNA prepared in this way is suitable for restriction digestions. Small amounts of RNA and other impurities can be removed by centrifugation in a cesium chloride gradient, although this step is not necessary. DNA concentrations are determined by absorbance at 260 nm (A_{260} = 1 unit = 50 µg/ml). This procedure results in high molecular weight DNA with yields varying from 7 to 20 µg of fresh tissue. Higher yields are obtained with young leaves.

2. CTAB = cetyltrimethylammonium bromide.

Restriction enzymes

Restriction enzymes are the most important tools used to detect variation at the DNA level. These enzymes were discovered in bacterial cells and have the distinct property of recognizing precision specific DNA sequences (usually 4 to 6 bases) and hydrolizing both DNA strands at a specific point. Consider the recognition sequence for the enzyme EcoRI:

```
G G G C C T T G*A A T T C T A T A G G
C C C G G A A C T T A A*G A T A T C C
```

This enzyme recognizes the six base-pair sequence indicated in bold letters and hydrolizes both DNA strands at the points indicated by the asterisks. A single base substitution, deletion, or addition within this sequence would prevent recognition. Thus, a single base difference between two plant lines can result in a heritable restriction fragment.

Consider a hypothetical region of the genome of two inbred lines P1 and P2 (Figure 1). In this example, both parents possess EcoRI sites A and C, but a single base change has resulted in the loss of the recognition site B in P2. This changes the pattern of fragments from 250 and 1000 base pairs in P1 to a single fragment of 1250 base pairs in P2. Currently, there are over 50 different restriction enzymes commercially available. By employing several different enzymes and examining numerous regions of the genome, a larger number of DNA markers can be analyzed in a single progeny.

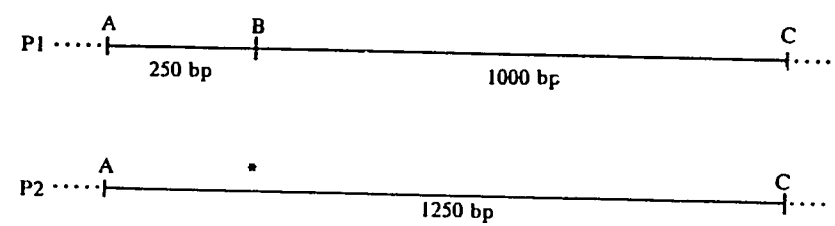


Figure 1. Hypothetical example in which a single base change (*) results in the loss of an EcoRI recognition site in P2 and a heritable RFLP between P1 and P2. (bp = base pair; RFLP = DNA restriction fragment length polymorphism.)

Electrophoresis

The different DNA fragments generated by digestion with a restriction enzyme can be physically separated according to size by horizontal agarose gel electrophoresis. For each gel lane, we normally digest 5 µg of DNA in a 50-µl volume under conditions specified by the manufacturer. DNA

samples are positioned in small wells at one end of the gel. When an electric field is applied to the gel, the negatively charged DNA fragments move toward the other end of the gel (the anode). The rate of migration of the DNA fragments will depend on the fragment size, the agarose concentration, and the applied current. The rate of migration is inversely proportional to the $\log_{10}$ of the molecular weight.

After electrophoretic separation, the DNA fragments are visualized by staining with ethidium bromide. This dye intercalates into the DNA and fluoresces when illuminated with ultraviolet light. Nuclear DNA restriction fragments appear to have a continuous size distribution which is seen as a smear in the gel (Figure 2). This size distribution pattern simply reflects the complexity and large size of the nuclear genome. In contrast, simple genomes such as those of chloroplasts or mitochondria, would yield discrete and easily identifiable bands.

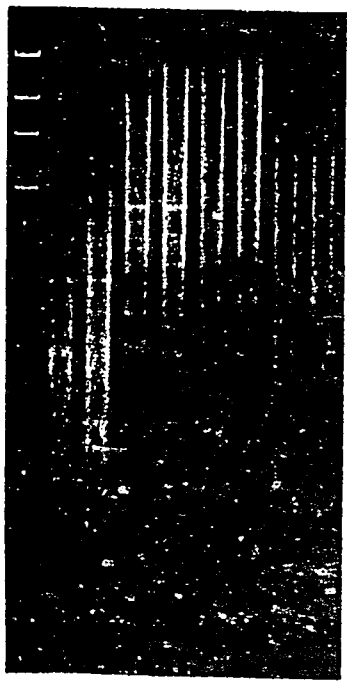


Figure 2. Agarose gel after staining DNA with ethidium bromide. Lane on the leaf is bacteriophage lambda DNA digested with HindIII. It provides molecular weight standards. Lanes on the right are DNAs of Calima and XR-235-1 digested with DraI, EcoRI, EcoRV, and HindIII.

Southern blots

The agarose gel is fragile, bulky, and does not lend itself to continuous manipulation. To overcome these problems, DNA fragments in the gel are blotted onto a solid support (nitrocellulose, or nylon membranes) by the Southern procedure (Southern, 1975). To facilitate the transfer of large fragments, the DNA in the gel is first depurated in HCl (0.25 N, 10 min). In order to allow future hybridization with complementary strands, the DNA is denatured in NaOH (0.4 N)/NaCl (0.6 *M*) for 30 minutes. This step ensures strand separation.

Alkaline transfers are normally made onto nylon membranes. The transfers are obtained by placing the gel over 3MM filter paper on a plate, submerging both ends of the filter in the transfer media. A sheet of nylon membrane, the size of the gel, is placed on top of the gel. Wet filter paper and dry paper towels are placed on top of the membrane. The transfer media moves from the tank, through the gel, toward the paper towels by capillarity, creating a mass flow. The single-stranded DNA fragments are trapped on the filter and conserve the same size distribution they had in the gel. Detection of particular DNA fragments is effected through hybridization with a labeled sequence of DNA, usually from a selected clone.

Detection Tools—Complementary DNA and Genomic Clones

Detection of a DNA restriction fragment immobilized on a blot is accomplished by hybridization with a radio-labeled cloned sequence of DNA contained in the restriction fragment. Two types or clones can be used for this purpose: complementary DNA (cDNA), or genomic.

DNA library

The normal flow of genetic information proceeds from DNA to RNA and then to proteins. Thus, the messenger RNAs (mRNAs), extracted from a given tissue, represent the fraction of the genome being expressed in that tissue at the time of sampling. Different procedures are used to construct cDNA libraries (Gubler and Hoffman, 1983). However, the general procedure starts with the isolation of polyadenylated mRNA. This RNA is then used as template to synthesize DNA with the enzyme "reverse transcriptase."

The product of this synthesis is a set of single-stranded cDNAs, each representing at least part of the coding sequences of an active gene in the nuclear genome. Double-stranded cDNAs are obtained by further enzymatic action. The set of double-stranded cDNA is inserted into a cloning vector at a specific restriction site. The most common vector used for this purpose is the plasmid. A plasmid is an extrachromosomal genetic element that can be found in bacterial cells. It is a circular and double-stranded DNA molecule. Inside the bacterial cells it uses the host's enzymes for replication.

Recombinant plasmids containing a cDNA are then introduced into bacterial cells, a process called transformation. The transformed bacterial cells are plated at a density at which each cell gives rise to a single colony. Each colony will contain a clone of a single gene. The collection of these colonies, each containing a recombinant plasmid, represents a "cDNA library" or a "cDNA clone bank." A colony can be amplified in large volumes of liquid media for the purpose of extracting significant quantities of a recombinant plasmid. This clone can be used to detect restriction fragments containing homologous sequences.

A cDNA library will contain copies of the genes actively expressed in the tissue from where the mRNA was originally extracted. Thus, depending on the tissue and its stage of development, a cDNA library could represent a biased sample of the genome. For the purpose of constructing a linkage map, a cDNA library should be constructed by using mRNA from all possible tissues at different stages of development.

Genomic library

A genomic library can be constructed by simply digesting nuclear DNA with an appropriate restriction enzyme, inserting the digested DNA into an appropriate vector, and transforming a bacterial host in much the same way as described for the construction of a cDNA library. However, such a library will contain many clones of repetitive DNA sequences. Such clones give complex hybridization patterns which are difficult to score in segregating progenies. To be useful in mapping, a genomic library must be enriched for low copy number sequences.

One strategy for enrichment is based upon the observation that those DNA sequences which are expressed contain fewer methylated bases than sequences which are inactive, and that certain restriction enzymes will not cut at their normal recognition sites if the sites contain methylated bases

(Doefler, 1983). We took advantage of these observations to produce a genomic library enriched for expressed sequences of *Phaseolus vulgaris*.

Genomic DNA was isolated from the snap bean cultivar Sprite by techniques described above. We digested 50 µg of DNA with the enzyme **Pst**I, which does not cut methylated recognition sequences (Helentjaris et al., 1985). Digested DNA was layered onto a 10-ml, 10%-30% sucrose gradient. Gradients were centrifuged in an SW41 rotor at 30,000 rpm for eight hours and fractionated by pumping from the bottom of the tube. Approximately 40 fractions of 250 µl were collected; 25 µl of each fraction was run on an agarose gel in parallel with molecular weight standards (bacteriophage lambda DNA digested with **Pst**I). Gradient fractions containing DNA fragments of 0.5-2.5 kb were identified and pooled. DNA was recovered by ethanol precipitation, ligated with 0.5 µg of **Pst**I digested plasmid vector (pTZ18R) (Pharmacia), and transformed into *E. coli* strain TB1.

Detection of DNA Restriction Fragments

Probe labeling

Cloned probes are normally labeled with a base containing the radioisotope of phosphorus, ^{32}P. One common procedure is "nick translation" (Rigby et al., 1977). In this procedure, 100 ng of plasmid DNA, or the insert alone, are incubated in a small volume (50 µl) with four bases (dATP, dGTP, dTTP, and ^{32}P-dCTP), DNase, DNA polymerase I, Mg^{2+}, and buffer. While the DNase removes bases from the intact plasmid (including the insert), the polymerase repairs the gaps by introducing bases from the medium, including the radio-labeled base ^{32}P-dCTP. At the end of the incubation period, the unincorporated bases are removed by gel filtration and we are left with radio-labeled DNA. Specific activities of 1×10^8 cpm per µg of DNA are obtained.

Hybridizations

Detection of DNA fragments is accomplished through hybridizations. For this purpose, the genomic Southern blot is enclosed in a plastic bag that contains buffer and salts. The radio-labeled DNA is first denatured by boiling and then introduced into the bag. The bag is sealed, and hybridization is allowed to proceed for 20-36 h at 65 °C. As mentioned

before, the DNA in the gels was also denatured before being blotted onto the nylon membrane. Thus, hybridization occurs between added single-stranded, radio-labeled DNA and homologous sequences bound to the membrane. At the end of the hybridization period, the blot is removed from the bag and is washed to remove the unhybridized probe. The membrane is wrapped in polyethylene film (Saran Wrap), and zones of hybridization are detected by autoradiography. X-ray (X-Omat, Kodak) film is exposed to the membranes at -70 °C for a period of 7 to 14 days (Figure 3). The hybridized probe can be washed off the membrane with a low salt solution at high temperature (100 °C) and the membrane can be used for another hybridization with a different probe.

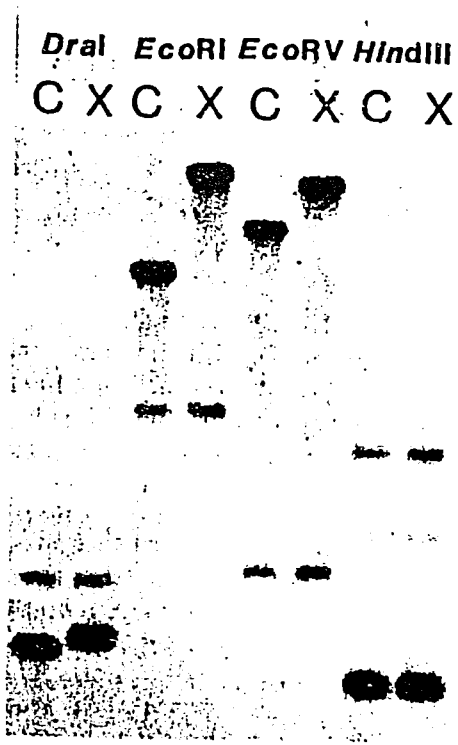


Figure 3. Autoradiogram of Southern blot hybridized with a genomic clone. Blot obtained from gel shown in Figure 2. (C = genomic clone; X = hybrid.)

Acknowledgments

We thank Ms. Cynthia Ragland and Mr. Samuel Camp for their technical assistance. We also thank Dr. M. J. Bassett for providing seeds of XR-235-1. This project has been supported in part by a grant from CIAT.

Resumen

Los autores obtuvieron progenies segregantes entre dos genotipos distintos de *Phaseolus vulgaris* y generaron clones genómicos y de cADN adecuados para la detección de los RFLP. Hasta el momento, se han hibridado, hasta digestos de restricción, 22 clones genómicos y de cADN de los cultivares Calima y XR-235-1. De éstos, 11 han presentado polimorfismo entre las dos líneas. Continúa el proceso de evaluación de este potencial.

El empleo de estas dos líneas en el esfuerzo de localización de genes hecho por los autores hará posible el trazado de mapas de los principales genes de resistencia a *X. campestris* pv. *phaseoli*; este gen está presente en XR-235-1.

Recientemente se inició el proceso de sondeo del ADN de una generación F_2 que había sido evaluada por su resistencia a las enfermedades. En el futuro, se emplearán marcadores de ADN para señalar otras características de importancia económica del frijol.

References

Ayonoadu, U. W. U. 1974. Nuclear DNA variation in *Phaseolus*. Chromosoma (Berl.) 48:41-49.

Beckmann, J. S. and Soller, M. 1983. Restriction fragment length polymorphisms in genetic improvement: methodologies, mapping and costs. Theor. Appl. Genet. 67:35-43.

Bennett, M. D. and Smith, J. B. 1976. Nuclear-DNA amounts in angiosperms. Philos. Trans. R. Soc. Lond. B Biol. Sci. 274(933):227-274.

Bernatzky, R. and Tanksley, S. D. 1986. Toward a saturated linkage map in tomato, based on isozymes and random cDNA sequences. Genetics 112(4):887-898.

Burr, B.; Evola, E.; Burr, F. A.; and Beckmann, J. S. 1983. The application of restriction fragment length polymorphisms to plant breeding. In: Setlow, J. K. and Hollaender, A. (eds.). Genetic engineering: principles and methods. Plenum Press, New York, NY, USA. Vol. 5, p. 45-59.

Doerfler, W. 1983. DNA methylation and gene activity. Annu. Rev. Biochem. 52:93-124.

Flavell, R. 1980. The molecular characterization and organization of plant chromosomal DNA sequences. Annu. Rev. Plant Physiol. 31:569-596.

Freytag, G. F.; Bassett, M. J.; and Zapata, M. 1982. Registration of XR-235-1-1 bean germplasm. Crop Sci. 22(2):1268-1269.

Gubler, U. and Hoffman, B. J. 1983. A simple and very efficient method for generating cDNA libraries. Gene (Amster.) 25:263-269.

Helentjaris, T.; King, G.; Slocum, M.; Siedenstrang, C.; and Wegman, S. 1985. Restriction fragment polymorphisms as probes for plant diversity and their development as tools for applied plant breeding. Plant. Mol. Biol. 5(2):109-118.

———; Slocum, M.; Wright, S.; Schaefer, A.; and Nienhuis, J. 1986. Construction of genetic linkage maps in maize and tomato, using restriction fragment length polymorphisms. Theor. Appl. Genet. 72:761-769.

Hunt, J. S. and Bernes, M. F. 1982. Molecular diversity and plant disease resistance: an electrophoretic comparison of near-isogenic lines of wilt-resistant or susceptible *Pisum sativum* L. cv. William Massey. Euphytica 31(2):341-348.

Landry, B. S.; Kesseli, R. V.; Farrara, B.; and Michelmore, R. W. 1987. A genetic map of lettuce (*Lactuca sativa* L.) with restriction fragment length polymorphism, isozyme, disease resistant and morphological markers. Genetics 116(2):331-337.

Mendlinger, S. 1983. Electrophoretic variability and the pedigree breeding method. In: Tanksley, S. D. and Orton, T. J. (eds.). Isozymes in plant genetics and breeding, 2 parts. Developments in plant genetics and breeding, 1A and 1B. Elsevier, Amsterdam, Netherlands. p. 377-380.

Pollak, L. J.; Gardner, C. O.; and Parkhurst, A. M. 1984. Relationships between enzyme marker loci and morphological traits in two mass selected maize populations. Crop. Sci. 24(6):1174-1179.

Pring, D. R. and Lonsdale, D. M. 1985. Molecular biology of higher plant mitochondria DNA. Int. Rev. Cytol. Rick, C. M. and Fobes, J. F. 1974. Association of an allozyme with nematode resistance. Rep. Tomato Genet. Coop. 25:25.

Rigby, P. W. J.; Dieckman, M.; Rhodes, C.; and Berg, P. 1977. Labeling deoxyribonucleic-acid to high specific activity in vitro by nick translation with DNA-polymerase I. J. Mol. Biol. 113(1):237-251.

Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel-electrophoresis. J. Mol. Biol. 98(3):503-517.

Tanksley, S. D. 1983. Molecular markers in plant breeding. Plant Mol. Biol. Rep. 1:3-8.

——— and Loaiza-Figueroa, F. 1985. Gametophytic self-incompatibility is controlled by a single major locus on chromosome 1 in *Lycopersicon peruvianum*. Proc. Natl. Acad. Sci. U.S.A. 82:5093-5096.

———; Medina-Filho, H.; and Rick, C. M. 1982. Use of naturally-occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. Heredy 49(1):11-25.

——— and Rick, C. M. 1980. Isozyme gene linkage map of the tomato: applications in genetics and breeding. Theor. Appl. Genet. 57:161-170.

———; ———; and Vallejos, C. E. 1984. Tight linkage between a nuclear male-sterile locus and an enzyme marker in tomato: potential application in plant breeding. Theor. Appl. Genet. 68(1-2):109-113.

Vallejos, C. E. and Tanksley, S. D. 1983. Segregation of low temperature tolerance and isozymes in an interspecific backcross of tomato. Theor. Appl. Genet. 66:241-247.

———; ———; and Bernatzky, R. 1986. Localization in the tomato genome of DNA restriction fragments containing sequences homologous to the rRNA (45s), the major chlorophyll a/b binding polypeptide and the ribulose bisphosphate carboxylase genes. Genetics 112(1):93-105.

Weeden, N. F.; Provvidenti, R.; and Marx, G. A. 1984. An isozyme marker for resistance to bean yellow mosaic-virus in *Pisum sativum*. J. Hered. 75(5):411-412.

Whitfeld, P. R. and Bottomley, W. 1983. Organization and structure of chloroplast genes. Annu. Rev. Plant. Physiol. 34:279-310.

Zamir, D.; Tanksley, S. D.; and Jones, R. A. 1982. Haploid selection for low temperature tolerance of tomato pollen. Genetics 101(1):129-137.

Theor Appl Genet (1991) 81:806-811

THEORETICAL AND APPLIED GENETICS

© Springer-Verlag 1991

DNA restriction fragment length polymorphisms correlate with isozyme diversity in *Phaseolus vulgaris* L.

C. D. Chase *, V. M. Ortega and C. E. Vallejos

Vegetable Crops Department, University of Florida, Gainesville, FL 32611, USA

Received July 16, 1990; Accepted November 3, 1990
Communicated by F. Salamini

Summary. Genetic variation in *Phaseolus vulgaris* L. (*P. vulgaris*) was investigated at the isozyme and DNA levels. We constructed a library of size-selected *Pst*I clones of *P. vulgaris* nuclear DNA. Clones from this library were used to examine 14 *P. vulgaris* accessions for restriction fragment length polymorphisms (RFLPs). DNAs from each accession were analyzed with three restriction enzymes and 18 single copy probes. The same accessions were also examined for variability at 16 isozyme loci. Accessions included four representatives of the T phaseolin group and five representatives each of the C and S phaseolin groups. One member of the S group (the breeding line XR-235-1-1) was derived from a cross between *P. vulgaris* and *P. coccineus*. Isozymes and RFLPs revealed very similar patterns of genetic variation. Little variation was observed among accessions with C and T phaseolin types or among those with the S phaseolin type. However, both isozyme and RFLP data grouped accessions with S phaseolin separately from those accessions with C or T phaseolin. The highest degree of polymorphism was observed between XR-235-1-1 and members of the C/T group. RFLP markers will supplement isozymes, increasing the number of polymorphic loci that can be analyzed in breeding, genetic, and evolutionary studies of *Phaseolus*.

Key words: Common bean – Molecular markers – Phaseolin – *Phaseolus coccineus* – Gene pools

Introduction

Restriction fragment length polymorphisms (RFLPs) provide a powerful tool for studies of plant genetics and evolution (Beckman and Soller 1986; Helentjaris and Burr 1989; Helentjaris et al. 1985). Application of this technology to the study of a species requires polymorphism at the DNA level, and different higher plant species exhibit different levels of DNA variability. While inbred lines of maize exhibit RFLPs with a high percentage of probes, RFLPs among lines of cultivated tomato are rare. RFLPs are readily detected, however, among *Lycopersicon* species which can be crossed with the cultivated tomato (Helentjaris et al. 1985). The level of DNA polymorphism in these species correlates with the level of polymorphism observed through isozyme analysis (Helentjaris et al. 1985).

In *P. vulgaris*, variability at the protein level has been well documented (Weeden 1984; Bassiri and Adams 1978b; Koenig and Gepts 1989). Isozyme analysis (Koenig and Gepts 1989) and the analysis of phaseolin seed storage proteins (Gepts and Bliss 1985; Gepts et al. 1986) demonstrate two genetically distinct groups of *P. vulgaris*. Accessions from Mesoamerica are primarily of the S phaseolin type (with some exhibiting the M phaseolin pattern). Accessions from the Andes are primarily of the T phaseolin type, with some accessions exhibiting C, H, A, J, or I types. *P. vulgaris* cultivars include members of both the S and the C/T phaseolin groups, suggesting multiple domestication events (Brown et al. 1982; Gepts et al. 1986).

P. vulgaris has been successfully hybridized with other *Phaseolus* species (Coyne 1964; Smartt 1970), and interspecific variation has been reported for isozymes (Bassiri and Adams 1978a) and seed proteins (Sullivan and Freytag 1986). Therefore, lines developed from interspecific crosses may also provide useful variants for DNA analysis.

Our interest in constructing a molecular marker-based linkage map of the *Phaseolus vulgaris* genome led

* To whom correspondence should be addressed

16

us to investigate the frequency of DNA polymorphisms within and between the genetically distinct groups. The assessment of genetic variation at the DNA level is an essential prerequisite to the design of evolutionary, breeding, and genetic studies employing DNA probes. Accessions with the S, C, and T phaseolin types were analyzed. The breeding line XR-235-1-1, which was developed from a *P. vulgaris* × *P. coccineus* cross (Freytag et al. 1982), was included to determine whether introgression of DNA from close relatives of *P. vulgaris* can contribute to variation at the DNA level. DNAs were examined with 18 single-copy genomic probes. Accessions were also analyzed for polymorphisms at 16 isozyme loci, to allow direct comparison of DNA and isozyme variability.

Materials and methods

Plant materials

With the exception of XR-235-1-1 and Sprite, seeds of all accessions used in this study were provided by J. Tohme and W. Roca, Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia. Seeds of XR-235-1-1 and Sprite were provided by M. Bassett, University of Florida, Gainesville/FL. The accessions analyzed for isozyme and DNA polymorphisms are listed in Table 1. The breeding line Ica Pijao (CIAT No. G05773) was included as a source of reference alleles in the isozyme studies.

Isozyme analysis

A combination of starch gel electrophoresis and enzyme activity staining was used to screen for polymorphisms of: aconitase (ACO) (E.C. 1.6.4.3); alcohol dehydrogenase (ADH) (E.C. 1.1.1.1); acid phosphatase (APS) (E.C. 3.1.3.2); Beta-N-acetylglucosaminidase (BNAG) (E.C. 3.2.1.30); diaphorase (DIA) (E.C. 1.6.4.3); glutamate oxaloacetate transaminase (GOT) (E.C. 2.6.1.1); malate dehydrogenase (MDH) (E.C. 1.1.1.37); 6-phosphogluconate dehydrogenase (6PGDH) (E.C. 1.1.1.44); phosphoglucoisomerase (PGI) (E.C. 5.3.1.9); phosphoglucomutase (PGM) (E.C. 2.2.5.1); and shikimate dehydrogenase (SKDH) (E.C. 1.1.1.25). These analyses were performed as described previously (Vallejos and Chase 1991).

Table 1. *Phaseolus vulgaris* accessions analyzed for isozyme and restriction fragment length polymorphisms

CIAT Number	Name	Phaseolin	Type
G15416	XR-235-1-1	S	Breeding line
G03645	Jamapa	S	Landrace
G04489	Culiapa 72	S	Landrace
G04493	Ica Bunsi	S	Breeding line
G04459	Nep 2	S	Mutant
G04435	Diacol Calima	C	Breeding line
G09011	Taylor Cramberry	C	Breeding line
G12172	Unnamed	C	Landrace
G12207	Canario	C	Landrace
G17668	Cran 28	C	Breeding line
G00051	Swedish Brown	T	Landrace
G01293	Algarrobo	T	Landrace
G03668	Sangretorro	T	Landrace
G14192	Horoz Fasulyesi	T	Landrace

Enzymes and reagents

Restriction enzymes and T4 DNA ligase were purchased from Bethesda Research Laboratories (BRL). All enzyme reactions were performed as recommended by the supplier in buffers provided by the supplier. [Alpha-^{32}P]dCTP (3,000 Ci/mmol) was purchased from Dupont/New England Nuclear Research Products.

Genomic library construction

Nuclear DNA was isolated by the procedure of Kislev and Rubenstein (1980) from 7-day-old etiolated seedlings of the cultivar Sprite. *Pst*I-digested DNA was layered onto a 10-ml, 10–30% (w/v) linear sucrose gradient. (Sucrose was dissolved in 10 m*M* TRIS, 1 m*M* EDTA, pH 8.) Gradients were centrifuged in an SW 41 rotor at 30,000 rpm for 8 h. Fractions of 250 µl were collected. Aliquots of each fraction were visualized following agarose gel electrophoresis in parallel with molecular weight standards. Gradient fractions containing DNA fragments of 0.5–2.5 kb were identified and combined. DNA was recovered by ethanol precipitation, rehydrated in sterile distilled water, and ligated with *Pst*I-digested plasmid vector (pTZ18R purchased from Pharmacia). The ligation reaction was used to transform the *Escherichia coli* strain TB1 by the procedure of Hanahan (1983). Recombinant colonies were identified by blue/white colony selection on LB-X-gal-ampicillin agar (Vieria and Messing 1982).

Probe preparation

Plasmid DNAs were prepared by the triton lysis method (Lonsdale et al. 1981). GeneClean (BIO 101) was used (as directed by the manufacturer) to recover the cloned *Pst*I inserts from agarose gels. The concentrations of insert DNAs were estimated by visualization; 10% of each insert preparation was visualized following agarose gel electrophoresis in parallel with DNAs of known concentrations. Random priming (Feinberg and Vogelstein 1984) DNA labeling kits (Boehringer, Mannheim) were used to label 100 ng of insert DNA in the presence of 50 µCi [alpha-^{32}P]dCTP.

Total nucleic acid preparations

Young trifoliolate leaves (one-third to one-half expanded) were harvested from a single plant of each accession. Leaf samples of 1 g were wrapped in foil, frozen in liquid nitrogen, and stored at −80°C. Total nucleic acids were purified from 1 g leaf samples by the method of Dellaporta et al. (1983).

Restriction, electrophoresis, and hybridization of nucleic acids

Samples containing 27 µg of nucleic acid (DNA and RNA) were digested with 20 units of restriction enzyme according to the supplier's instructions. Digested samples were loaded onto 0.8% agarose gels, which were run at 2 V/cm for 18.75 h. Molecular size standards (*Pst*I fragments of bacteriophage lambda-DNA) were included on each gel. Gels were stained for 1 h in ethidium bromide (0.5 µg/ml) and photographed over a UV (302 nm) transilluminator.

Following photography, DNAs were blotted to Hybond-N+ nylon membranes (Amersham Corp.) as described by Southern (1975). DNAs were fixed to membranes by incubation for 20 min on Whatman 3-mm chromatography paper soaked in 0.4 *M* NaOH. Each 20 × 20 cm blot was hybridized with 100 ng of heat-denatured, radiolabeled insert DNA prepared as de-

- 17

scribed above. Blots were prehybridized, hybridized, and washed as described by Church and Gilbert (1984). X-ray films (Kodak XAR-5) were exposed under Lightning Plus intensifier screens (Dupont) to radiolabeled blots for 5 days. Following autoradiography, probes were removed from blots by placing the blots in a boiling solution of 0.5% w/v SDS and allowing the solution to cool to room temperature.

Cluster analysis

Cluster analysis was accomplished through use of the BIOSYS-I computer program, version 1.6 (D. L. Swofford and R. B. Selander, University of Illinios at Urbana-Champaign). Isozyme and RFLP banding patterns were converted to allele frequencies. Pair-wise distances (Nei 1973) between accessions were computed based upon comparisons at the 18 DNA or 16 isozyme loci. Dendrograms based upon genetic distances were constructed by the unweighted paired group method (Sneath and Sokal 1973).

Results

The results of isozyme analysis are summarized in Table 2. Of the 16 isozyme loci examined, 9 revealed polymorphisms. Relatively little variation was observed among members of the C/T phaseolin group. Only two loci (*Aco-1* and *Mdh-1*) distinguished any members of this group, and only one member was distinguished in each case. The S group was slightly more variable, with four loci (*Adh-1*, *Bnag*, *Got-2*, and *Dia-2*) revealing differences. However, two of the four loci distinguished only XR-235-1-1 (accession G15416) from the other members, and XR-235-1-1 is unique in that it was derived from a *P. vulgaris* × *P. coccineus* cross. All members of the S phaseolin group differed from all members of the C and T phaseolin groups with respect to *Aco-2*, *Dia-1*, and *Skdh*. Maximum variation was observed between XR-235-1-1 and members of the C/T phaseolin group, with 6 of 16 isozyme loci revealing differences.

For the investigation of DNA polymorphism, we constructed a library of size-selected *Pst*I clones of *P. vulgaris* nuclear DNA. Others report this type of library to be enriched in single-copy sequences, because single-copy sequences include those which are actively expressed and therefore undermethylated. These sequences are preferentially digested by methylation-sensitive enzymes such as *Pst*I (Burr et al. 1988; Helentjaris et al. 1988). This strategy was highly effective for *P. vulgaris*, as 95% of cloned inserts shorter than 2.5 kb gave simple (one to three band) hybridization patterns expected for single-copy probes. Representative hybridization data are shown in Fig. 1.

DNAs prepared from the 14 different accessions were digested with each of three different restriction enzymes (*Eco*RI, *Eco*RV, and *Hin*dIII) and probed with 18 different probes, for a total of 54 different probe-enzyme combinations. Results of the *Eco*RI hybridizations are summarized in Table 3. Each of the three restriction enzymes resulted in useful polymorphism. Polymorphisms were revealed in *Eco*RI digests by five different probes, in *Eco*RV digests by five probes, and in *Hin*dIII digests by seven probes. With four exceptions, polymorphisms revealed with one probe-enzyme combination were not revealed by the same probe in combination with another enzyme. In three of the four exceptions, the polymorphic individual was XR-235-1-1.

Variation observed at the DNA level matched that detected by isozyme analysis (Table 2; Koenig and Gepts

Table 2. Zymotypes[a] of *P. vulgaris* accessions

Accession	Phaseolin	Enzyme locus								
		Aco-1	*Aco-2*	*Adh-1*	*Bnag*	*Dia-1*	*Dia-2*	*Got-2*	*Mdh-1*	*Skdh*
G15416	S	+	+	a_4	[illegible]	+	c_{30}	r_7	+	+
G03645	S	+	+	+	+	+	c_{30}	+	+	+
G04459	S	+	+	+	+	+	+	+	+	+
G04489	S	+	+	+	+	+	+	+	+	+
G04493	S	+	+	+	a_5	+	c_{30}	+	+	+
G04435	C	r_2	a_1	+	+	a_5	c_{30}	+	a_{10}	r_5
G09011	C	+	a_1	+	+	a_5	c_{30}	+	+	r_5
G12172	C	+	a_1	+	+	a_5	c_{30}	+	+	r_5
G12207	C	+	a_1	+	+	a_5	c_{30}	+	a_{10}	r_5
G17668	C	+	a_1	+	+	a_5	c_{30}	+	+	r_5
G00051	T	+	a_1	+	+	a_5	c_{30}	+	+	r_5
G01293	T	+	a_1	+	+	a_5	c_{30}	+	+	r_5
G03668	T	+	a_1	+	+	a_5	c_{30}	+	a_{10}	r_5
G14192	T	+	a_1	+	+	a_5	c_{30}	+	+	r_5

[a] Electrophoretic mobilities were determined relative to reference alleles of Ica Pijao (G05773); + indicates mobility equivalent to that of the reference allele; a_x indicates mobility advanced x mm relative to the reference allele; r_x indicates mobility retarded x mm relative to the reference allele; c_x indicates cathodal mobility x mm relative to the origin while mobility of the reference allele was anodal. The accessions showed no variability at the *Aps-1*, *Aps-2*, *Got-1*, *Mdh-2*, *6Pgdh*, *Pgi-1*, or *Pgm-1* loci

18

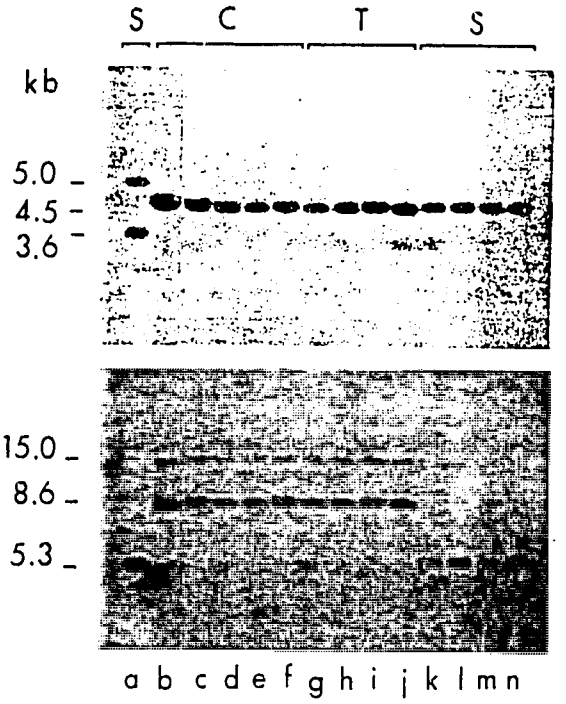


Fig. 1. Autoradiographs of *P. vulgaris* DNAs hybridized with genomic probes. The autoradiograph in the top panel showns *Hin*dIII-digested DNA hybridized with probe 2F5; this probe-enzyme combination distinguished XR-235-1-1 (accession G15416) from the other accessions. The autoradiograph in the *bottom panel* shows *Eco*RI-digested DNA hybridized with probe 2F7; this probe-enzyme combination distinguished all accessions with S phaseolin from those with C or T phaseolin. DNAs were prepared from accessions: G15416 (*a*), G04435 (*b*), G09011 (*c*), G12172 (*d*), G12207 (*e*), G17668 (*f*), G00051 (*g*) G01293 (*h*), G03668 (*i*), G14192 (*j*), G03645 (*k*), G04459 (*l*), G04489, (*m*), and G04493 (*n*). *S*, *C*, and *T* indicate the phaseolin types of the accessions

Table 3. Polymorphic *Eco*RI fragments[a] of *P. vulgaris* DNA detected by hybridization with genomic probes

Accession	Phaseolin	Probe					
		2C11	2E7	2F3	2F7	2F12	4A9
G15416	S	4.6	3.7	4.8	5.3	7.3	9.1
G03645	S	4.6/3.3	3.4	4.8	5.3	7.3	9.1
G04459	S	4.6/3.3	3.7	4.8	5.3	7.3	9.1
G04489	S	4.6/3.3	3.7	4.8	5.3	7.3	9.1/7.4
G04493	S	4.6/3.3	3.7	4.8	5.3	7.3	9.1
G04435	C	4.6/3.1	3.7	2.6	15/8.6	10.5	5.0
G09011	C	4.6/3.1	3.4	2.6	15/8.6	10.5	5.0
G12172	C	4.6/3.1	3.4	2.6	15/8.6	10.5	5.0
G12207	C	4.6/3.1	3.4	2.6	15/8.6	10.5	5.0
G17668	C	4.6/3.1	3.4	2.6	15/8.6	7.3	5.0
G00051	T	4.6/3.1	3.4	2.6	15/8.6	10.5	5.0
G01293	T	4.6/3.1	3.4	2.6	15/8.6	10.5	5.0
G03668	T	4.6/3.1	3.4	2.6	15/8.6	10.5	5.0
G14192	T	4.6/3.1	3.4	2.6	15/8.6	10.5	5.0

[a] Sizes of fragments are given in kilobase pairs (kb). No variation was observed among these accessions when *Eco*RI digest were probed with 2B12, 2C6, 2E9, 2E10, 2F5, 2F6, 2F8, 2F9, 2H4, 2H9, 4B10, or 4C2

1989) and by the analysis of seed storage proteins (Gepts et al. 1986; Gepts and Bliss 1985). Relatively little variation was revealed within each of the genetically distinct groups defined by phaseolin type. Of the 18 probes used, only 4 distinguished any of the C/T accessions from the others, and no more than 3 of the 8 C/T accessions were differentiated by any single probe. Members of the S group were more variable than those of the C/T group; 9 of the 18 probes distinguished one or more members of the S group. However, 4 of the 9 probes distinguished only XR-235-1-1 from the others. Variation was readily detected between the two groups; 6 of the 18 probes distinguished all members of the S group from all members of the C/T group. Maximum variation was observed between XR-235-1-1 and members of the C/T group, with 12 of the 18 probes revealing polymorphisms.

Agreement of the isozyme and RFLP data was apparent in dendrograms based upon genetic distances calculated from these data. Dendrograms based upon *Eco*RI, isozyme, and combined *Eco*RI and isozyme data are shown in Fig. 2a, b, and c, respectively. Features common to these dendrogramms were also present in dendrograms (not shown) generated from *Eco*RV or *Hin*dIII data. All data sets grouped the five accessions with S phaseolin separately from those accessions with C or T phaseolin, and showed XR-235-1-1 (G15416) to be the most divergent accession of the S phaseolin group. Other subgroups within the S or the C/T phaseolin groups varied from dendrogram to dendrogram, and there was no clear separation of accessions with the C phaseolin from those with the T phaseolin type.

Discussion

The limited genomic variation observed within each of the genetically divergent groups of *P. vulgaris* was comparable to that observed in cultivated tomato, where only 2 of 22 probes revealed polymorphisms between two lines (Helentjaris et al. 1985). No more than 1 of the 18 probes used in our study distinguished any pair of accessions with the C or T phaseolin types or any pair of accessions with the S phaseolin type (excluding XR-235-1-1 from the analysis). This is in contrast to the highly polymorphic nature of maize inbreds (Evola et al. 1986; Helentjaris et al. 1985). Helentjaris et al. (1985) noted that DNA variation in maize and tomato correlates with variation observed at isozyme loci. Our results (Tables 2–3, Fig. 2), taken with the results of others (Gepts et al. 1986; Koenig and Gepts 1989), demonstrate that DNA variation is well correlated with isozyme variation in *P. vulgaris*.

Helentjaris et al. (1985) suggested a number of factors that could account for differences in genomic variability; these include differences in mating systems, breeding systems, and domestication events. Studies of natural plant

19

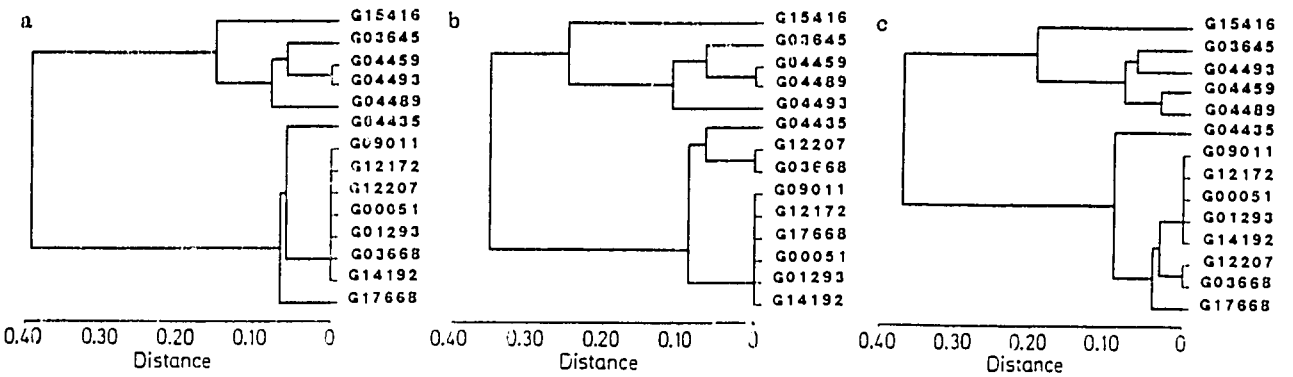


Fig. 2a–c. Dendrograms of *P. vulgaris* accessions based upon comparisons at isozyme and RFLP loci. Dendrogram a was based upon comparison of the accessions at DNA loci detected by probing EcoRI digests with the 18 probes described in Table 3. Dendrogram b was based upon comparisons at the 16 isozyme loci described in Table 2. The data used to generate dendrograms a and b were combined to generate dendrogram c

populations demonstrate high variability at isozyme loci in cross-pollinating populations and low levels of variability in self-pollinating populations (Hamrick et al. 1979; Rick et al. 1977, 1979). The two genetically distinct groups of *P. vulgaris* are believed to originate from two distinct wild gene pools – Andean and Mesoamerican (Gepts and Bliss 1985; Gepts et al. 1986; Koenig and Gepts 1989). *P. vulgaris* therefore provides an example of a naturally self-pollinating species that exhibits low levels of DNA and isozyme variation within gene pools.

Self-pollination, however, also results in increased genetic variability between gene pools (Loveless and Hamrick 1984). Variability between the two major genetic groups of *P. vulgaris* has been documented through analysis of seed storage proteins (Gepts and Bliss 1985; Gepts et al. 1986) and isozymes (Koenig and Gepts 1989). We have extended these analyses to the DNA level. RFLP data, considered independently from phaseolin or zymotype data, group accessions in a manner consistent with isozyme and seed storage protein groupings (Fig. 2).

Additional variability is present in accessions of wild *P. vulgaris*. Novel phaseolin types have been identified in wild populations of Mexico (Romero-Andreas and Bliss 1985), Colombia (Gepts and Bliss 1986), and Andean South America (Gepts et al. 1986). RFLPs may be useful in investigating the relationships among accessions with novel phaseolins and those with C, T, and S types. RFLPs may also be useful in investigating the genetic diversity of *P. vulgaris* accessions from the Colombia-Peru region, a minor center of genetic diversity and domestication for common bean (Gepts and Bliss 1986; Koenig and Gepts 1989).

XR-235-1-1 contributed a high degree of variability to the S phaseolin group in this study and was distinguished from members of the C/T group by more probes and isozymes than any other S type accession examined. This high degree of variation may result from the *P. coccineus* genome contributions to the genome of XR-235-1-1. High pollen abortion has been observed in F_2 plants from crosses between *P. vulgaris* and *P. coccineus*. While Cheng et al. (1981) reported evidence that chromosomes of *P. vulgaris* and *P. coccineus* differ by two inversions, others found the chromosomes of the two species to be very similar if not identical (Shii et al. 1982). This discrepancy could be due to the different plant accessions analyzed. Most of the DNA polymorphisms observed in this study were probably the result of single base changes, as polymorphisms revealed by a given probe-enzyme combination were generally not revealed by the same probe in combination with other restriction enzymes. We note that in three of the four exceptions to this generalization the polymorphic individual was XR-235-1-1. This may be indicative of minor structural differences (short insertions, deletions, or inversions) between the genomes of the two species.

Construction of a linkage map (based upon RFLP and isozyme markers) of the *P. vulgaris* genome is in progress. The breeding lines XR-235-1-1 (G15416) and Diacol Calima (G04435) were selected as appropriate mapping parents exhibiting maximum levels of polymorphism. Maximum variation is an obvious advantage in linkage studies. However, application of RFLPs to investigations of *P. vulgaris* will not necessarily require materials generated by interspecific crosses, as one-third of our probes revealed polymorphisms between two genetic groups of this species. RFLPs will therefore supplement isozymes, increasing the number of polymorphic loci that can be analyzed in breeding, genetic, and evolutionary studies of *Phaseolus*.

Acknowledgements. We thank Drs. J. Tohme, W. Roca, and M. Bassett for supplying seeds of the accessions used in this study, and for helpful discussions. This work was supported in part by USAID Grant No. DPE-5542-G-SS-7060-00, and is part of a collaborative program between the University of Florida and CIAT. Florida Agricultural Experiment Station Journal Series No. R-00853.

20

References

Bassiri A, Adams MW (1978a) An electrophoretic survey of seedling isozymes in several *Phaseolus* species. Euphytica 27:447–459

Bassiri A, Adams MW (1978b) Evaluation of common bean cultivar relationships by means of isozyme electrophoretic patterns. Euphytica 27:707–720

Beckman JS, Soller M (1986) Restriction fragment length polymorphisms and genetic improvement of agricultural species. Euphytica 35:111–124

Brown JWS, McFerson JR, Bliss FA, Hall TC (1982) Genetic divergence among commercial classes of *Phaseolus vulgaris* in relation to phaseolin pattern. Hort Science 17:752–754

Burr B, Burr FA, Thompson KH, Albertson MC, Stuber CW (1988) Gene mapping with recombinant inbreds in maize. Genetics 118:519–526

Cheng SS, Bassett MJ, Quesenberry KH (1981) Cytogenetic analysis of interspecific hybrids between common bean and scarlet runner bean. Crop Sci 21:75–79

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

Coyne DP (1964) Species hybridization in *Phaseolus*. J. Hered 55:5–6

Dellaporta SL, Wood J, Hicks J (1983) Maize DNA minipreps. Maize Genet Coop Newsl 57:26–27

Evola SV, Burr FA, Burr B (1986) The suitability of restriction fragment length polymorphisms as genetic markers in maize. Theor Appl Genet 71:765–771

Feinberg AP, Vogelstein B (1984) A technique for radiolabeling DNA restriction fragments to high specific activity. Anal Biochem 137:266–269

Freytag GF, Bassett MJ, Zapata M (1982) Registration of XR-235-1-1 bean germ plasm. Crop Sci 22:1268–1269

Gepts P, Bliss FA (1985) F_1 hybrid weakness in the common bean. J Hered 76:447–450

Gepts P, Bliss FA (1986) Phaseolin variability among wild and cultivated common beans (*Phaseolus vulgaris*) from Colombia. Econ Bot 40:469–478

Gepts P, Osborn TC, Rashka K, Bliss FA (1986) Phaseolin protein variability in wild forms and landraces of the common bean (*Phaseolus vulgaris*); evidence for multiple centers of domestication. Econ Bot 40:451–468

Hamrick JL, Linhart YB, Mitton JB (1979) Relationships between life history characteristics and electrophoretically detectable genetic variation in plants. Annu Rev Ecol Syst 10:173–200

Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol 166:557–580

Helentjaris T, Burr B (eds) (1989) Development and application of molecular markers to problems in plant genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor/NY, 165 pp

Helentjaris T, King G, Slocum M, Siedenstrang C, Wegman S (1985) Restriction fragment polymorphisms as probes for plant diversity and their development as tools for applied plant breeding. Plant Mol Biol 5:109–118

Helentjaris T, Slocum M, Wright S, Schaefer A, Nienhuis J (1986) Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. Theor Appl Genet 72:761–769

Helentjaris T, Weber D, Wright S (1988) Identification of the genomic locations of duplicate nucleotide sequence in maize by analysis of restriction fragment length polymorphisms. Genetics 188:353–363

Kislev N, Rubenstein I (1980) Utility of ethidium bromide in the extraction from whole plants of high-molecular-weight maize DNA. Plant Physiol 66:1140–1143

Koenig R, Gepts P (1989) Allozyme diversity in wild *Phaseolus vulgaris*: further evidence for two major centers of genetic diversity. Theor appl Genet 78:809–817

Lonsdale DM, Thompson RD, Hodge TP (1981) The integrated forms of the S1 and S2 DNA elements of maize male-sterile DNA are flanked by a large repeated sequence. Nucleic Acids Res 9:3657–3669

Loveless MD, Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. Annu Rev Ecol Syst 15:65–95

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323

Rick CM, Fobes JF, Holle M (1977) Genetic variation in *Lycopersicon pimpinellifolium*: evidence of evolutionary change in mating systems. Plant Syst Evol 127:139–170

Rick CM, Fobes JF, Tanksley SD (1979) Evolution of mating systems in *Lycopersicon hirsutum* as deduced from genetic variation in electrophoretic and morphological characters. Plant Syst Evol 132:279–298

Romero-Andreas J, Bliss FA (1985) Heritable variation in the phaseolin protein of nondomesticated common bean, *Phaseolus vulgaris* L. Theor Appl Genet 71:478–480

Shii CT, Rabakoarihanta A, Mok MC, Mok DWS (1982) Embryo development in reciprocal crosses of *Phaseolus vulgaris* L. and *P. coccineus* Lam. Theor Appl Genet 62:59–64

Smartt J (1970) Interspecific hybridization between cultivated American species of the genus *Phaseolus*. Euphytica 19:480–489

Sneath PHA, Sokal R (1973) Numerical taxonomy. Freeman, San Francisco/CA

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517

Sullivan JG, Freytag G (1986) Predicting interspecific compatibilities in beans (*Phaseolus*) by seed protein electrophoresis. Euphytica 35:201–209

Vallejos CE, Chase CD (1991) Linkage between isozyme markers and a locus affecting seed size in *Phaseolus vulgaris* L. Theor Appl Genet 81:413–419

Vieria J, Messing J (1982) The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene 19:259–268

Weeden NF (1984) Distinguishing among white-seeded bean cultivars by means of allozyme genotypes. Euphytica 33:199–208

- 21

Theor Appl Genet (1991) 81:413–419

THEORETICAL AND APPLIED GENETICS

© Springer-Verlag 1991

Linkage between isozyme markers and a locus affecting seed size in *Phaseolus vulgaris* L.*

C. E. Vallejos ** **and C. D. Chase**

University of Florida, Institute of Food and Agricultural Sciences, Department of Vegetable Crops, 1255 Fifield Hall, Gainesville, FL 32611, USA

Received May 29, 1990; Accepted August 22, 1990
Communicated by F. Salamini

Summary. Backcross and F_2 progenies were produced between two bean genotypes, 'XR-235' and 'Calima,' which differ in seed weight by a factor of two. The small-seeded 'XR-235' was used as the pistillate and recurrent parent. These genotypes showed polymorphisms at nine isozyme loci and at the phaseolin locus. Seed size parameters (weight, length, width, and thickness) were determined for each BC_1 and F_2 individual, i.e., for seeds harvested from 'XR-235' after pollination with F_1 and from the F_1 after selfing, respectively. A combination of starch gel electrophoresis and enzyme activity staining was used to determine the genotype of each BC_1 and F_2 individual at the segregating loci. SDS-PAGE and Coomassie blue staining were used to determine genotype at the phaseolin locus. Tests for independent assortment using two-way contingency and maximum likelihood tables revealed three linkage pairs: *Aco-1* – 20 cM – *Dia-1*; *Adh-1* – 2 cM – *Got-2*; and *Est-2* – 11 cM – *Pha*. Statistical comparisons were made between the means of genotype classes at each segregating locus for all seed size parameters. The results from two independently obtained BC_1s and the F_2 consistently indicated that the *Adh-1*–*Got-2* segment was linked to a locus that affected seed size and overcame maternal control over seed size. This locus has been designated *Ssz-1*. This gene exhibited additive gene action and accounted for 30–50% of the seed size difference between the parents.

Key words: Beans – Seed size – QTL – Isozymes – Linkage

Introduction

Seed size (weight) is considered a component of yield in dry bean breeding programs (Al-Mukhtar and Coyne 1981; Conti 1982, 1985). Negative correlations between seed size and yield have been reported for some crosses (Coyne 1968); however, such correlations can impose limitations in breeding programs of large-seeded types. Seed size in cultivated beans has been described as polygenic (Sax 1923; Yarnell 1965) and, from crosses between cultivated and wild types, it has been estimated that there are at least ten genes involved (Motto et al. 1978). The use of genetic markers to identify genes involved in seed size control may help assess more clearly their role. Sax (1923) pointed out that the genes controlling continuous variation could be identified through linkage with qualitatively inherited characters. In fact, he identified a gene affecting seed size in beans, which was linked to a gene that controls pigmentation, the *P* locus. More recently, quantitative trait loci have been identified and mapped using isozyme loci (Tanksley et al. 1982) and DNA restriction fragment length polymorphisms (RFLPs) (Paterson et al. 1988) in tomato.

We are constructing a linkage map in beans based on isozymes and RFLPs. The general strategy has been to produce backcross and F_2 progenies from a pair of contrasting genotypes. Some of this contrast encompasses a twofold size difference in seed weight. Traditionally, the genetic analysis of bean seed weight has been carried out with progeny tests, using the average seed weight of seeds produced by each individual of a BC_1 and/or F_2 generation. This procedure is dictated by the control of the maternal tissue genotype over the size of the seed it bears (Bassett 1982). However, when we harvested F_1 seeds, i.e., seeds harvested from 'XR-235' after pollination with 'Calima,' it was noticed that these seeds had a weight

* Florida Agricultural Experiment Station, Journal Series No. R00696
** To whom correspondence should be addressed

22

intermediate between the parents. Furthermore, seed size variation was observed in the BC_1 seed harvested from the small-seeded parent ('XR-235') after pollination with the F_1, and in the F_2 seed harvested from F_1 plants after selfing. This variation suggested that some genes from the large-seeded 'Calima' parent could exert a direct effect on the size of the embryo and overcome the maternal effect on seed size. The objective of this project was to determine whether any of the segregating isozyme loci could be used to detect genes that affect seed size in the mode described above.

Materials and methods

Plant material

Two *Phaseolus vulgaris* L. inbreds that have allelic differences at nine enzyme loci and a twofold difference in seed size were selected to produce backcross and F_2 progenies: 'XR-235-1', a breeding line obtained from Dr. M.J. Bassett (University of Florida, Department of Vegetable Crops), which carries resistance to common bacterial blight (Freytag et al. 1982) and has an average seed weight of 275 mg, and 'Calima', a cultivar obtained from Dr. W. Roca (CIAT), which is adapted to Colombia and East Africa with an average seed weight of 500 mg. 'XR-235' was used as the pistillate parent to obtain the F_1 progeny and as the recurrent parent for the backcross progenies.

Isozyme analysis

A combination of starch gel electrophoresis and enzyme activity staining was used to detect polymorphisms between parental lines for eight enzymes: alcohol dehydrogenase (ADH) (E.C. 1.1.1.1), aconitase (ACO) (E.C. 4.2.1.3), β-N-acetylglucosaminidase (BNAG) (E.C. 3.2.1.30), diaphorase (DIA) (E.C. 1.6.4.3), esterase (EST) (E.C. 3.2.1.2), glutamate oxaloacetate transaminase (GOT) (E.C. 2.6.1.1), malate dehydrogenase (MDH) (E.C. 1.1.1.37), and shikimate dehydrogenase (SKDH) (E.C. 1.1.1.25). Enzyme activity stains were those described by Tanksley (1984) for ACO, by Vallejos (1983) for ADH, EST, GOT, MDH, and SKDH, and by Weeden (1984 and 1986) for DIA and BNAG. ADH, ACO, and MDH were resolved with a histidine buffer system: gel buffer (5 m*M* histidine · HCl, adjusted to pH 7.0 with NaOH), electrode buffer (135 m*M* TRIS, 43 m*M* citric acid, final pH 7.0). Samples were loaded in the gel at 13.3 V/cm for 15 min, the wicks were then removed and the gel was run for 3 more hours. BNAG, EST, GOT, and SKDH were resolved using a TRIS-citrate system: gel buffer (15.2 m*M* TRIS, 3.9 m*M* citric acid, pH 7.8), electrode buffer (0.3 *M* H_3BO_3 adjusted to pH 7.9 with 4 *N* NaOH). Samples were loaded for 15 min at 13.3 V/cm, the wicks were then removed and the gel was run at 13.3 V/cm for 1 h, and at 20 V/cm until the borate front reached 10 cm from the origin. Alternatively, this system can be run at 2.6 mA/cm^2 on constant current at all stages. Finally, DIA was resolved in a TRIS-citrate lithium borate system: electrode buffer (190 m*M* borate adjusted to pH 8.1 with 2 *M* LiOH), gel buffer (9 vol. of 50 m*M* TRIS adjusted to pH 8.4 with citric acid and 1 vol. of electrode buffer). Samples were loaded at 15 V/cm for 20 min and, after the wicks were removed, the gel was run at the same conditions for 4 moure hours (Sprecher and Vallejos 1989).

Seed samples were used for ADH and MDH. Seed flour was obtained by scratching the cotyledons with a scalpel blade under the raphe after partially removing the seed coat; 20 mg of seed flour was incubated in a microcentrifuge tube with 60 µl of extraction buffer (0.1 *M* TRIS · HCl, pH 7.8) for 30 min on ice. Samples were collected from the supernatant with filter paper wicks after a 2-min spin. Samples for ACO, BNAG, GOT, and SKDH were obtained by grinding petioles with 1/3 vol. of buffer, and those for DIA and EST were obtained from root tissue in a similar manner.

Phaseolin type was determined by resolving seed extracts on SDS-PAGE. Briefly, seed flour was obtained as described above, and 5 mg of seed flour was incubated with 50 µl of extraction buffer [0.0625 *M* TRIS · HCl (pH 6.8), 0.5 *M* NaCl] for 30 min. The samples were centrifuged for 5 min at 16,000 g and 5 µl of the supernatant was mixed with 195 µl of SDS sample buffer (Laemmli 1970) and incubated at 37°C for 1 h. Protein extracts were separated by SDS PAGE in 12% gels.

Seed parameters and statistical analysis

Seed measurements in the different progenies were taken on the seed of the individuals themselves, i.e., seeds harvested from 'XR-235' after pollination with 'Calima' for the F_1, from 'XR-235' after pollination with the F_1 for BC_1, and from the F_1 after selfing for F_2. Individual seed weights were measured with an analytical balance (Mettler AE 100) to 0.1 mg of resolution. Length, width and thickness of individual seeds were measured with a Vernier caliper to a resolution of 0.1 mm. All statistical analyses were performed on a mainframe computer using the Statistical Analysis System (SAS Institute, Cary/NC).

Results

Quantitative trait-seed size

The descriptive parameters of seed size (weight, length, width, and thickness) and shape (ratios of the three dimensions) for the parents and their F_1 are listed on Table 1. The mean seed weight of 'Calima' was about twice that of the breeding line 'XR-235-1', whereas the seed weight of their hybrid was intermediate (Table 1). The three one-dimensional parameters of 'Calima' seeds also exceeded those of the breeding line, whereas the dimensions of the F_1 seeds were intermediate. Regardless of size, comparisons of the ratios of the three dimensions showed that 'Calima' seeds were more elongated than 'XR-235' seeds [$(L/W)_{Cal} > (L/W)_{XR}$, and $(L/T)_{Cal} > (L/T)_{XR}$] and that they were also proportionally thinner [$(W/T)_{Cal} < (W/T)_{XR}$] (Table 1). As expected from the previous measurements the F_1 had intermediate ratios.

Seeds of the parents and F_1 used for the measurements were obtained in the greenhouse at the same time. However, those of the segregating progenies were obtained at different times and under slightly different greenhouse conditions. For this reason, direct comparisons between parents and progeny, and heritability calculations cannot be made with the available data. The two BC_1 progenies (BC_1A and BC_1B) were obtained at different times. Seed weight and sizes were slightly higher for the BC_1A than for the BC_1B, although only the differences in width and W/T ratio were significantly different (0.01 level). This difference lead to their independent analysis. On the other hand, the BC_1A and F_2 progenies could

Table 1. Seed size and shape measurements of the parents ('XR-235-1' and 'Calima'), their F_1 ('XR-235-1' × 'Calima'), two BC_1 ('XR-235' × F_1) sets (A and B), and the F_2

		'XR-235'	F_1	'Calima'		BC_1A[b]	BC_1B	F_2
Weight (mg)	x[a]	276.7	405.0	505.5	x	415.3a	383.6	518.2
	s_x	7.6	11.8	9.9	s	102.1	62.3	85.8
	n	30	25	30	n	52	38	99
Length (mm)	x	10.74	13.00	15.38	x	13.13	12.80	15.40
	s_x	0.12	0.14	0.15	s	1.39	1.41	1.22
Width (mm)	x	6.77	7.51	7.93	x	7.96**	7.63	8.21
	s_x	0.07	0.08	0.06	s	0.66	0.40	0.49
Thickness (mm)	x	4.78	5.45	6.00	x	5.48	5.51	5.87
	s_x	0.06	0.10	0.06	s	0.69	0.51	0.53
L/W[c]	x	1.59	1.73	1.94	x	1.64	1.67	1.88
	s_x	0.02	0.01	0.01	s	0.08	0.15	0.13
L/T	x	2.25	2.40	2.57	x	2.42	2.32	2.64
	s_x	0.02	0.05	0.03	s	0.20	0.21	0.21
W/T	x	1.42	1.39	1.32	x	1.46*	1.39	1.41
	s_x	0.02	0.03	0.01	s	0.15	0.12	0.13

[a] x = mean, s_x = standard error, s = standard deviation, n = sample size
[b] t-tests were used to compare the means of the two independent backcrosses; no significant differences were found between means that underlined
[c] L = length, W = width, T = thickness.
*, ** Significant differences at the 0.05 and 0.01 levels, respectively. Comparisons of BC_1A and F_2 using t-tests showed significant differences at the 0.001 level for all parameters listed

Table 2. Segregation analysis of enzyme loci backcross and F_2 progenies using the χ^2 test

Locus		Genotype			χ^2	χ^2	Locus		Genotype			χ^2	χ^2
		x/x[a]	x/c	c/c	1:1	1:2:1			x/x	x/c	c/c	1:1	1:2:1
Aco-1	BC_1A	32	20		2.77		*Est-2*	BC_1A	20	24		0.36	
	BC_1B	20	18		0.10			BC_1B	21	17		0.42	
	F_2	31	61	31		0.01		F_2	37	68			5.87*,b
Aco-2	BC_1A	26	23		0.18		*Got-2*	BC_1A	19	33		3.77	
	BC_1B	18	20		0.10			BC_1B	21	17		0.42	
	F_2	32	70	21		4.32		F_2	24	58	41		5.10
Adh-1	BC_1A	20	32		2.77		*Mdh-1*	BC_1A	25	27		0.08	
	BC_1B	21	17		0.42			BC_1B	20	18		0.10	
	F_2	25	59	39		3.39		F_2	25	58	40		4.06
Bnag	BC_1A	22	21		0.02		*Skdh*	BC_1A	18	23		0.61	
	BC_1B	20	18		0.10			BC_1B	13	25		3.79	
	F_2	33	57	33		0.66		F_2	38	58	27		2.36
Dia-1	BC_1A	26	23		0.18		*Pha*	BC_1A	29	23		0.69	
	BC_1B	17	21		0.42			BC_1B	18	20		0.10	
	F_2	33	65	25		1.44		F_2	31	59	29		0.07

[a] x/x and c/c are homozygous for 'XR-235' and 'Calima' alleles, respectively; x/c is heterozygous
[b] A 1:3 ratio was tested because the heterozygotes could not be distinguished from the homozygotes for 'Calima' alleles
* Significant at the 0.05 level

be compared with each other because they were obtained at the same time and under identical greenhouse conditions. One-tail t-test indicated that the means for seed weight, length, width, and thickness of the F_2 were significantly larger than those of the BC_1A (0.001 level). Finally, tests of normality using the Shapiro-Wilk statistic ($n<51$) (Shapiro an Wilk 1965) or the Kolomogorov statistic ($n>50$) (Stephens 1974) indicated that all of the seed weight and size parameters were normally distributed.

Table 3. Tests of independent assortment for linked isozyme loci using two-way contingency tables. Results from two independently obtained backcross progenies

Loci			Genotypes				χ^2	p[a]	$\bar{p}$
Y	Z		Parentals		Recombinants				
			$Y^{x/x}Z^{x/x}$	$Y^{x/c}Z^{x/c}$	$Y^{x/x}Z^{x/c}$	$Y^{x/c}Z^{x/x}$			
Aco-2	*Dia-1*	A[b]	19	16	7	7	8.91*	28.6	
		B	13	16	5	4	10.45**	23.7	26.1
Adh-1	*Got-2*	A	18	31	2	1	40.06**	5.8	
		B	21	17	0	0	38.00**	0.0	2.6
Est-2	*Pha*	A	20	20	4	0	30.56**	9.1	
		B	17	16	4	1	21.24**	13.1	11.1

[a] Percent recombination
[b] A and B indicate rows for data obtained with the BC_1A and BC_1B respectively
*, ** Significant at the 0.005 and 0.001 levels, respectively

Table 4. Calculation of linkage by the maximum likelihood method (Allard 1956) for three pairs of linked loci using F_2 data

		Dia-1			cM
		x/x	c/c	c/c[a]	
Aco-2	x/x	23	8	1	19.7±2.9
	x/c	8	49	13	
	c/c	2	8	11	
		Got-2			
Adh-1	x/x	24	1	0	2.0±0.9
	x/c	0	56	3	
	c/c	0	1	38	
		Pha			
Est-2	x/x	28	5	3	10.8±3.2
	_/c	1	43	22	

[a] Genotypes: x/x and c/c stand for homozygotes for 'XR-235' and 'Calima' alleles, respectively; x/c stands for heterozygotes; _/c stands for a mixture of homozygotes and heterozygotes

Segregation and linkage analysis of isozyme loci

Monogenic segregations of the isozyme loci were analyzed with χ^2 tests in the backcross and F_2 progenies (Table 2). All loci tested in the backcrosses segregated in the expected 1:1 ratio. Slight but not significant skewings were observed for *Aco-1*, *Adh-1*, and *Got-2* in BC_1A, and for *Skdh* in BC_1B. No significant deviations from Mendelian ratios were detected in the F_2, with the exception of *Est-2*, for which an excess of 'XR-235' alleles was detected; a 1:3 ratio was tested at this locus because the heterozygotes could not be distinguished from the homozygotes for the 'Calima' allele. In this progeny again, a slight but not significant overabundance of 'Calima' alleles was detected at the *Adh-1*, *Got-2*, and *Mdh-1* loci, whereas the opposite was true for the *Aco-2* and *Skdh* loci.

The linkage relationships between isozyme loci were also investigated. Tests of independent assortment for all pair combinations were conducted using two-way contingency tables. These tests revealed three linked pairs (Table 3). Significant deviations from independent assortment were detected in all three progenies under study for these pairs. The average recombination obtained with the two BC_1s was 26% for the *Aco-2–Dia-1* pair, while the maximum likelihood method using F_2 data gave a linkage intensity of 19.8±2.9 cM (Table 4). The average recombination between *Adh-1* and *Got-2* estimated with BC_1 data (2.6%) was in good agreement with the linkage intensity determined with F_2 data (2±0.9 cM). Finally, an average recombination of 11% between *Est-2* and *Pha* observed with BC_1 data was confirmed with the maximum likelihood estimate from F_2 data of 10.8±3.2 cM.

Detection of QTLs for seed size

Large-seeded 'Calima' was used as the donor parent to generate the BC_1s. Thus, genes of this genotype that control embryo size by overcoming maternal control would be segregating in these progenies. In order to investigate possible linkage between these genes and isozyme loci, one-tail *t*-test were used to compare the means of seed size parameters between homozygotes and heterozygotes at all segregating enzyme loci (Table 5). The means of all seed size parameters of heterozygotes were significantly higher than those of homozygotes at the *Adh-1* and *Got-2* loci. It has already been established that these loci are 2 cM apart. Therefore, it can be concluded that there is at least one locous affecting seed size in this chromosome region. These results were confirmed with one-way analyses of variance for seed size parameters of the F_2 progeny at each segregating locus. A *t*-test

Table 5. Detection of linkage between segregating isozyme loci and a locus that affects seed size. One-tailed *t*-test was used to compare the means of homozygotes (x/x) and heterozygotes (x/c) at each segregating locus

Locus		Genotype		Weight (mg)			Length (mm)			Width (mm)			Thickness (mm)		
		x/x	x/c	x/x	x/c	*t*-T	x/x	x/c	*t*-T	x/x	x/c	*t*-T	x/x	x/c	*t*-T
Aco-1	A	32	20	396.5	445.2	*	12.8	13.6	*	7.78	8.24	*	5.42	5.58	NS
	B	20	18	376.7	391.3	NS	12.4	13.1	NS	7.56	7.70	NS	5.53	5.48	NS
Aco-2	A	26	23	441.0	405.4	NS	13.6	12.9	NS	8.09	7.96	NS	5.66	5.37	NS
	B	18	20	390.6	377.3	NS	13.0	12.5	NS	7.63	7.62	NS	5.56	5.46	NS
Adh-1	A	20	32	329.5	468.9	***	11.9	13.9	***	7.39	8.31	***	4.96	5.81	***
	B	21	17	353.5	420.8	***	12.1	13.6	***	7.50	7.79	**	5.35	5.71	*
Bnag	A	22	21	414.2	438.6	NS	13.1	13.5	NS	7.96	8.08	NS	5.50	5.64	NS
	B	20	18	394.4	371.5	NS	13.1	12.3	NS	7.68	7.57	NS	5.56	5.45	NS
Dia-1	A	26	23	424.4	424.1	NS	13.3	13.2	NS	7.95	8.12	NS	5.63	5.41	NS
	B	17	21	380.0	386.5	NS	13.0	12.6	NS	7.62	7.64	NS	5.50	5.52	NS
Est-2	A	20	24	443.1	416.9	NS	13.4	13.3	NS	8.02	8.10	NS	5.74	5.40	NS
	B	21	17	367.7	403.2	*	12.4	13.2	NS	7.55	7.72	NS	5.44	5.59	NS
Got-2	A	19	33	324.5	467.5	***	11.8	13.9	***	7.33	8.31	***	4.99	5.77	***
	B	21	17	353.5	420.8	***	12.1	13.6	***	7.50	7.79	**	5.35	5.71	*
Mdh-1	A	25	27	407.9	422.1	NS	13.0	13.3	NS	7.91	8.00	NS	5.49	5.48	NS
	B	20	18	384.3	382.8	NS	13.0	12.5	NS	7.58	7.68	NS	5.50	5.52	NS
Skdh	A	18	23	424.6	425.6	NS	13.4	13.2	NS	8.04	8.00	NS	5.50	5.60	NS
	B	13	25	385.4	382.6	NS	12.8	12.7	NS	7.66	7.61	NS	5.44	5.49	NS
Pha	A	29	23	413.4	417.6	NS	13.0	13.3	NS	7.85	8.09	NS	5.55	5.41	NS
	B	18	20	377.3	389.2	NS	12.5	13.0	NS	7.56	7.70	NS	5.55	5.46	NS

*, **, *** Significant at the 0.05, 0.01, and 0.001 levels, respectively
NS – not significant

Table 6. Results of the analyses of variance and Duncan's tests (0.05 level) used to compare the means of the F_2 genotypes at each isozyme locus

Locus	Genotypes			Weight (mg)			Length (mm)			Width (mm)			Thickness (mm)		
	x/x	x/c	c/c	x/x	x/c	c/c	x/x	x/c	c/c	x/x	x/c	c/c	x/x	x/c	c/c
Aco-1	24	48	27	518	514	524	15.1	15.4	15.7	8.18	8.18	8.30	5.80	5.94	5.80
Aco-2	24	57	18	511	523	511	15.4	15.4	15.0	8.13	8.25	8.20	5.87	5.85	5.91
Adh-1	16	49	34	471[*,b]	517[a,b]	541[a]	14.9[*,b]	15.3[a,b]	15.8[a]	8.17	8.25	8.17	5.58[**,b]	5.81[a,b]	6.08[a]
Bnag	26	49	24	513	535	487	15.3	15.6	15.0	8.27	8.30	7.27	5.93	5.92	5.70
Dia-1	24	52	23	514	516	526	15.3	15.5	15.3	8.14	8.20	8.30	5.89	5.84	5.91
Est-2[d]	31	62		506	521		15.2	15.4		8.12	8.23		5.98	5.82	
Got-2	15	48	36	467[a,b]	514[a,b]	544[a]	14.8[*,b]	15.2[a,b]	15.8[a]	8.17	8.25	8.19	5.59[**,b]	5.79[b]	6.09[a]
Mdh-1	24	44	31	499	517	534	14.9[*,b]	15.4[a,b]	15.8[a]	8.14	8.23	8.24	5.94	5.82	5.87
Skdh	31	44	24	525	518	507	15.6	15.7	15.2	8.24	8.24	8.14	5.94	5.86	5.77

*, ** Significant at the 0.5 and 0.005 levels respectively; '*F*' tests for the main effect
[a,b] No significant differences were detected between means bearing the same letters
[d] A *t*-test was used at this locus because only two electrophoretic phenotypes could be distinguished

was used for the *Est-2* locus, because the heterozygotes could not be distinguished from the homozygotes for 'Calima' alleles. These results showed that the chromosome segment marked by *Adh-1*–*Got-2* had significant effects on all seed size parameters, with the exception of width (Table 6). Duncan tests were used to evaluate the difference between the genotype means at these loci. In both cases, the two homozygote classes were significantly different from each other, but each was not significantly different from the heterozygote class. However, the mean of the heterozygote class was always intermediate, suggesting additivity at this locus. Noteworthy are the contrasting sensitivities of the statistical tests for the BC_1 ($P=0.001$) and the F_2 ($P=0.05$) progenies.

Marginal significant differences were detected only in BC_1A at *Aco-1* for weight, length, and width. Similar results were obtained for the seed weight of BC_1B at the *Est-2* locus. However, these results were proven to be artifacts under closer scrutiny. Two-way factorial analyses of variance between *Got-2* and all the other loci showed that the chromosome segment marked by *Got-2* was the only source of variation; when the variation at the *Got-2* locus was accounted for, variation at the other loci was not detected. Furthermore, inspection of the two-way contingency tables showed slight deviations for the pair combinations of both *Aco-1* (in BC_1A) and *Est-2* (in BC_1B) with *Got-2*: these anomalies can account, at least in part, for the artifacts observed. Interestingly, the two-way factorial analysis between *Adh-1* and *Got-2* in BC_1A suggested that the latter was closer to the locus responsible for seed size. However, this observation cannot be considered conclusive. The tight linkage between these loci resulted in only one individual in one recombinant class and two in the other (Table 3); the excessive disproportion in class sizes seriously reduced the usefulness of this analysis of variance. Significance for seed length at the *Mdh-1* locus was detected in the F_2. Several interpretations are possible for this results. If a seed size gene in this chromosome segment had a detectable effect only in the F_2 seed and not in the BC_1 seed, then a recessive allele would be expected; however, the homozygous class for the 'Calima' allele was not significantly different from the heterozygous class. Another explanation is the possible interaction between the *Mdh-1*- linked locus with other factors not detected in this study. Finally, this result may be due to the observed underrepresentation of the double homozygote class for 'XR-235' alleles for *Got-2* and *Mdh-1*. At best, the role of this chromosome segment cannot be assessed conclusively in this study.

Discussion

The normal distribution detected in the BC_1 and F_2 progenies underscored the polygenic nature of the seed size character, as others have previously reported (Motto et al. 1978; Sax 1923; Yarnell 1965). Two lines of evidence indicate that the principal mode of action of these polygenes is additivity. Seeds of the F_1 progeny had intermediate weights between the progenitors. The mean seeds weight of the BC_1A was closer to the small-seeded recurrent parent than the mean seed weight of the F_2 progeny. Additive gene action for genes involved in seed size has also been reported previously (Hamblin and Morton 1977; Conti 1985).

The twofold difference in seed size between the progenitors suggested that either a large number of polygenes with similar effects or at least one major and a few others with minor effects would segregate in the progenies. With a limited number of markers, one can expect to detect minor genes with tight linkages or major genes with moderate linkage. The consistent results obtained with two independently obtained backcrosses and the F_2 strengthened the validity of the linkage between (*Adh-1*-2cM-*Got-2*) and the locus that contributes to seed size. Inspection of the mean seed weights of the three genotypes at either the *Got-2* or *Adh-1* loci indicates that the mode of action of this gene is additive. The mean seed weight of heterozygotes was intermediate to the two homozygote classes. This locus is assigned the symbol *Ssz-1* to indicate seed size. An interesting feature of this gene is its ability to overcome the maternal control of seed size, which is normally observed in beans.

The contrasting sensitivities of the statistical analyses observed between the BC_1 and F_2 progenies makes for an interesting observation. The much higher level of significance detected with the BC_1 indicates clearly that this gene can easily overcome the maternal control of the recurrent parent genotype (XR-235). On the other hand, the marginal level of significance observed with the F_2 suggests two nonexclusive explanations. First, a greater variation is expected in each class of the F_2 relative to each class of the BC_1. For a given number of genes (n) involved in a character, the number of genotypes per class at each locus will be a power base higher in the F_2 than in the BC_1 [$(F_2)\ 3^{n-1} > (BC_1)\ 2^{n-1}$]. The second and more interesting possibility is that this gene has a smaller effect when it is present in seeds coming from the hybrid genotype (XR-235 × Calima). The fact that there are not many reports on this type of gene action for seed size leads to the question of whether this effect would be more frequently observed in compatible crosses between the two gene pools of beans. These results also suggest that there may be specific interactions between genes expressed in the embryo and genes expressed in the maternal plant. These interactions could be more noticeable in wide crosses. We have begun to investigate this possibility.

Parental and progeny seed sizes were obtained from seeds produced under different greenhouse conditions and direct comparisons cannot be made. However, with the available data it is possible to obtain a first approximation of the contribution to seed weight by this gene. The smallest difference (67 mg) between heterozygotes and homozygotes at the *Got-2* locus was detected in BC_1B. This weight difference represents 50% of the variation in seed weight between the F_1 and the small-seeded progenitor. On the other hand, the difference between the two homozygous classes in the F_2 at the *Got-2* locus is about 33% of that of the parents. Therefore, this appears to be a major gene that can explain from 33 to 50% of the variation in seed size. These results also mean that there

must be other genes responsible for the remainder of the seed size variation besides environmental effects.

Tagging of *Ssz-1* provides a unique opportunity to study the mechanism by which this gene controls seed size. Does *Ssz-1* control the extent of cell division in the developing embryo or the final cell size? For instance, these results indicate that *Ssz-1* affects lenght to a greater extent than it does the other two dimensions, suggesting in turn that this gene also affects seed shape. What effect does this gene have on source-sink relationships during embryo development? The availability of tags for genes that control seed size can be useful in studies of bean evolution. In addition to phaseolin type, seed size is one of the major differentiating characters between mesoamerican and andean *Phaseolus vulgaris* beans (Gepts 1988). The availability of genetic markers for genes involved in seed size control in beans will certainly aid physiological and evolutionary studies centered around this important agronomic character.

Acknowledgements. This research was supported in part by USAID grant DPE-5542-G-SS-7060-00. This project is part of a collaborative program between the University of Florida and the International Center for Tropical Agriculture (CIAT). We would like to thank Mr. S. Camp for his technical assistance and Dr. M.J. Bassett for making his computer program available for the calculation of map distances. We are grateful to Dr. M.J. Bassett, Dr. L.C. Hannah, and Dr. G.A. Moore for their critical reading of this manuscript and for their valuable suggestions.

References

Allard RW (1956) Formulas and tables to facilitate the calculation of recombination values in heredity. Hilgardia 24:235–278

Al-Mukhtar FA, Coyne DP (1981) Inheritance and association of flower, ovule, seed, pod, and maturity characters in dry edible beans (*Phaseolus vulgaris* L.). J Am Soc Hortic Sci 106:713–719

Bassett MJ (1982) A dwarfing gene that reduces seed weight and pod length in common bean. J Am Soc Hortic Sci 197:1058–1061

Conti L (1982) Bean germ plasm evaluation from the collection at Minoprio (Como, Italy) in view of a breeding program for the improvement of the proteic content of the seed. Genet Agrar 35:375–392

Conti L (1985) Conclusive results of a selection program for obtaining a dwarf bean (*P. vulgaris*) resistant to some viruses and characterized by agronomical qualities. Genet Agrar 39:51–63

Coyne DP (1968) Correlation, heritability, and selection of yield components in field beans, *Phaseolus vulgaris* L. Proc Am Soc Hortic Sci 93:388–396

Freytag GF, Bassett MJ, Zapata M (1982) Registration of XR-235-1-1 bean germ plasm. Crop Sci 22:1268–1269

Gepts P (1988) Phaseolin as an evolutionary marker. In: Gepts P (ed) Genetic resources of *Phaseolus* beans. Kluwer Academic, Dordrecht, pp 215–241

Hamblin J, Morton JR (1977) Genetic interpretations of the effects of bulk breeding on four populations of beans (*Phaseolus vulgaris* L.). Euphytica 26:75–83

Laemmli UK (1970) Cleavage of structural proteins during assembly of the head of bacteriophage T4. Nature 227:680–685

Motto M, Soressi GP, Salamini F (1978) Seed size inheritance in a cross between wild and cultivated common beans (*Phaseolus vulgaris* L.). Genetica 49:31–36

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335:721–726

Sax K (1923) The association of size differences with seed coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics 8:552–560

Shapiro SS, Wilk MB (1965) An analysis of variance test for normality. Biometrika 52:591–611

Sprecher SL, Vallejos CE (1989) System/tissue/enzyme combination for starch gel electrophoresis of bean. Annu Rep Bean Improv Coop 32:32–33

Stephens MA (1974) EDF statistics for goodness of fit and some comparisons. J Am Stat Assoc 69:730–737

Tanksley SD (1984) Aconitase isozymes – new gene markers for tomato. Rep Tomato Genet Coop 34:16–18

Tanksley SD, Medina-Filho H, Rick CM (1982) Use of naturally occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. Heredity 49:11–25

Vallejos CE (1983) Enzyme activity staining. In: Tanksley SD, Orton TJ (eds) Isozymes in plant genetics and breeding, vol 1A. Elsevier, Amsterdam, pp 469–516

Weeden NF (1984) Distingushing among white-seeding bean cultivars by means of allozyme genotypes. Euphytica 33:199–208

Weeden NF (1986) Genetic confirmation that the variation in the zymograms of 3 enzyme systems is produced by allelic polymorphism. Annu Rep Bean Improv Coop 29:117–118

Yarnell SH (1965) Cytogenetic of the vegetable crops. IV. Legumes. Bot Rev 31:247–330

28

Theor Appl Genet (1991) 82:353–357

THEORETICAL AND APPLIED GENETICS

© Springer-Verlag 1991

Extended map for the phaseolin linkage group of *Phaseolus vulgaris* L*

C. E. Vallejos and C. D. Chase**

Department of Vegetable Crops, Institute of Food and Agricultural Science, 1255 Fifield Hall, University of Florida, Gainesville, FL 32611, USA

Received December 1, 1990; Accepted January 8, 1991
Communicated by R. Hagemann

Summary. The linkage relationship of 11 bean (*Phaseolus vulgaris*) seed proteins (including phaseolin), 9 enzyme loci, and the *P* locus were analyzed in backcross and F_2 progenies by use of the software package "Mapmaker." The progenies were obtained by crossing the breeding line 'XR-235-1' and the cultivar 'Calima'. Allelic differences for seed protein loci were detected with SDS-PAGE and those for enzyme loci with starch gel electrophoresis and activity stains. The seed coat color of 'Calima' is a red/beige mottled pattern and that of 'XR-235-1' is white. Segregation at the *P* locus was followed by recording the phenotype of the BC_1S_1 and F_3 seed. A linkage group comprising ca. 90 cM was detected with the following gene order: *Est-2* – 11 – *Pha* – 8 – (*Spe*/*Spg*) – 24 – *P* – 9 – (*Spa*/*Spb*) – 16 – *Spba* – 22 – *Mdh-1*. In addition, another linkage group was detected: (*Spd*/*Spf*/*Sph*) – 5 – *Spca*. Therefore, the seed proteins appear to be organized in clusters in the bean genome.

Key words: Common bean – Seed proteins – Isozymes – Mapmaker

Introduction

Limited linkage information is currently available for *Phaseolus vulgaris* L. Only nine small linkage groups have been identified and they are defined by as few as two or as many as six morphological marker loci (Bassett 1988). The paucity of linkage data in common bean is due to reliance on morphological markers provided by natural and induced mutations; bean geneticists have found it difficult to work with lines that carry more than two recessive mutations, because of reduced vigor and fertility (Bassett 1988).

The use of molecular markers (isozymes and other proteins) for linkage analysis in common bean has been introduced in more recent years. Brown et al. (1981 b) reported that the gene encoding the major seed storage protein (phaseolin, *Pha*) was linked to genes encoding other seed storage proteins. Weeden (1984) reported linkage between the small subunit of Rubisco and malic enzyme: later, Weeden and Liang (1985) detected linkage between *Est-2* and the *P* locus. The *P* locus is essential for the expression of pigmentation in common bean (Leakey 1988). We have recently reported (Vallejos and Chase 1991) three pairs of linked isozyme/protein loci: *Adh-1* – *Got-2* (2 cM), *Aco-2* – *Dia-1* (20 cM), and *Est-2* – *Pha* (11 cM). We report here an extended linkage map of the latter group, which spans ca. 90 cM and includes six seed storage proteins, two enzyme loci, and the *P* locus.

Materials and methods

Plant material

Two inbred lines of beans were used to generate the segregating progenies: 'Calima' and 'XR-235-1.' 'Calima' is a large-seeded cultivar which belongs to the phaseolin type C group. The seed coat has a red/beige mottled pattern; this line is homozygous dominant (*PP*) at the *P* locus. 'XR-235-1' is a breeding line derived from an interspecific cross with *P. coccineous* (Freytag et al. 1982); this is a small, white-seeded (*pp*) bean which belongs to the phaseolin type S group. This breeding line was used as the pistillate parent to produce both the F_1 and the backcross (BC_1) progeny; and F_2 progeny was also produced. All BC_1 and F_2

* Florida Agricultural Experiment Station, Journal Series No. R-01131

** To whom correspondence should be addressed

plants were grown to maturity and selfed to produce the BC_1S_1 and F_3 progenies, respectively.

Isozymes

Preparation of samples, electrophoretic separation in starch gels, and enzyme activity staining have been described elsewhere (Vallejos and Chase 1991). Segregation for the following enzyme loci was recorded: *Adh-1*, *Aco-1*, *Aco-2*, *Bnag*, *Dia-1*, *Est-2*, *Got-2*, *Mdh-1*, and *Skdh*.

Seed storage proteins

After removing the seed coat under the raphe, seed flour was obtained by carefully scraping the cotyledons with a scalpel. Five milligrams of seed flour was incubated with 50 µl of extraction buffer (0.0625 *M* TRIS · HCl, pH 6.8) for 30 min on ice. The samples were centrifuged for 2 min at 16.000 × *g* and 5 µl of the supernatant was mixed with 195 µl of SDS sample buffer (Laemmli 1970) and incubated at 65 °C for 1 h. Protein extracts were separated by SDS-PAGE (13%). Five microliter of each SDS sample was loaded in gels that were 16 cm long and 1 mm thick (Protean II; Bio-Rad, Richmond/CA). Phaseolin bands were visualized by staining with Coomassie brilliant blue. These were the only protein bands that were clearly visible in the gel at these protein concentrations. Under these conditions, it was possible to clearly identify the three *Pha* genotypes (*SS*, *SC*, and *CC*) in the F_2 progeny. In order to visualize the other seed proteins, the gels were destained and then silver stained using the procedure of Morrissey (1981), with the following modifications: gels were fixed overnight in ethanol (30%)/acetic acid (10%) and then incubated in glutaraldehyde (1.5%); finally, the gels were incubated in acetic acid (10%) after development of the silver stain.

Seed proteins of the parental genotypes were extracted and separated into G1, G2, albumin, and prolamine fractions according to the procedures of Brown et al. (1981a). The fractions were further separated by SDS-PAGE (Laemmli 1970) and visualized by silver staining as described above. This procedure allowed for further characterization of the protein bands that showed polymorphism between the selected parents.

Segregation and linkage analysis

Statistical analyses were carried out on a mainframe computer using the Statistical Analysis System (SAS Institute, Cary/NC). Linkage analysis was performed using the computer package "MAPMAKER" (Lander et al. 1987), which allows for a simultaneous multipoint linkage analysis. This program uses the Lander and Green (1987) algorithm to calculate the "best" map for a given order of loci.

Results

Polymorphisms

Mdh-1 alleles were detected as doublet bands. The 'Calima' allele was ca. 5 mm more advanced than the 'XR-235-1' allele, and the three possible genotypes could be distinguished in the F_2 progeny. On the other hand, the EST banding patterns of 'Calima' and XR-235 were almost identical, except that the second most anodal band of 'Calima' stained more intensely than that of 'XR-235-1.' It was not possible to distinguish between heterozygotes and 'Calima' homozygotes in the F_2. For this reason the 'Calima' allele was treated as dominant. Allelic differences in staining intensity between bands of identical electrophoretic mobility have been described in tomato (Tanksley and Rick 1980). All other enzyme loci were detected as single bands of different mobility, as described earlier (Vallejos and Chase 1991).

The seed coat is maternal tissue and therefore reflects the genotype of the mother plant. For this reason, the pigmentation pattern was recorded for seeds harvested from BC_1 and F_2 progenies. Seeds of the progenies were either white (*pp*) or pigmented (*Pp* for BC_1 and *P-* for F_2). In the pigmented class the patterns varied from plain beige to varying intensities of red mottling, indicating the presence of other genes involved in mottling.

Comparisons of the seed storage protein profiles revealed size polymorphisms for at least 11 different proteins (Table 1). The two parental alleles were identified for 5 of the 12 proteins, including phaseolin (*Spa*, *Spb*, *Spf*, *Sph*, and *Pha*). Full segregation information was recorded for these proteins in the F_2 progeny. Of the remaining protein bands, five were unique to 'Calima' and their corresponding alleles were unidentified in 'XR-235-1' (*Spba*, *Spc*, *Spd*, *Spe*, and *Spg*). The segregation of these proteins was recorded according to genotype in the BC_1 and according to phenotype in the F_2 progeny. One band was unique to 'XR-235-1' (*Spca*), and its allelic counterpart was not identified in 'Calima'; the segregation of this protein was recorded only in the F_2.

Monogenic segregations

Results of the tests for Mendelian ratios in the BC_1 (1:1) are depicted in Table 2. Significant deviations were detected for only two of the loci studied: *Spba* ($P = 0.05$) and *Spc* ($P = 0.005$). A significant deficiency of recurrent parent (XR-235) alleles was detected at each one of these

Table 1. Characterization of seed proteins that display allelic differences between 'Calima' and 'XR-235-1': molecular weights and protein fraction

Protein		Molecular weights (kD)	
Band	Fraction	Calima	XR-235
Spa	G1	82.8	79.1
Spb	G1	58.5	53.3
Spba	G2	52.1	
Pha	G1	45.3	
		44.0	44.5
		42.4	42.4
Spc	G2	37.9	
Spca	G2		34.8
Spd	G2	31.0	
Spe	G2	20.5	
Spf	G2	17.0	17.4
Spg	G1	16.0	
Sph	G2	14.8	15.4

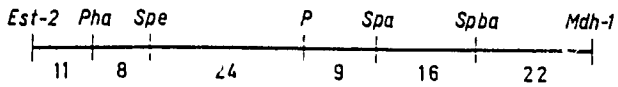


Fig. 1. Average map distances in centiMorgans for the *Pha* linkage group

Table 2. Analysis of monogenic segregation ratios in backcross (BC_1) and F_2 progenies

Locus	Backcross			F_2				
	Genotype		$\chi^2_{(1:1)}$	Genotype			$\chi^2_{(1:2:1)}$	$\chi^2_{(1:3)}$
	x/x[a]	x/c		x/x	x/c	c/c		
Est-2	41	41	0.00	37	68[b]			5.87*
Pha	47	43	0.18	31	59	29	0.07	
Spe, g	47	43	0.18	37	80[b]			2.74
P	31	37	0.37	22	101[b]			3.32
Spa, b	43	47	0.18	20	58	43	8.90*	
Spba	35	52	4.13*	30	92[b]			0.01
Mdh-1	45	45	0.00	25	58	40	4.06	
Spc	31	59	8.71**	21	101			3.94*
Spca					83	30		0.14
Spd, f, h	42	48	0.40	26	95			0.80

[a] Genotypes: x/x homozygous for XR-235 alleles; c/c homozygous for Calima alleles; x/c heterozygous
[b] When heterozygous plants could not be distinguished from homozygous for either Calima or XR-235 alleles, a 1 : 3 ratio was tested in the F_2 progeny
*, ** Significant deviation at the 0.05 and 0.005 levels, respectively

loci, although it was less pronounced for *Spba* than for *Spc*. This deficiency persisted for *Spc* in the F_2. In addition, in the F_2 progeny, two different loci displayed significant ($P=0.05$) deviations from expected ratios. One of them, *Est-2*, had an excess of homozygotes for 'XR-235' alleles, whereas the linked pair *Spa/Spb* showed a deficiency for the same class.

Linkage analysis

During the process of recording the segregation data, it was noticed that some of the seed proteins were inherited in clusters, i.e., only parental combinations could be detected. These clusters were made up as follows: (*Spa/Spb*), (*Spd/Spf/Sph*), and (*Spe/Spg*). One member of each group was included in the data set for linkage analysis. "Mapmaker" was used to analyze the linkage relationships among the segregating loci in both progenies (BC_1 and F_2). The linkage criteria for the analysis were such that linkage was considered only if the LOD scores were greater than 3.0 and the recombination value was less than 0.30. The maximum LOD score is the $\log_{10}$ of the ratio between the likelihood when the locus pair is at its maximum likelihood recombination fraction and when the loci are taken to be unlinked. Thus, a LOD score of 3.0 for a locus pair means that the pair is 1,000 times more likely to be linked than not linked. The analysis confirmed our previous report (Vallejos and Chase 1991) for three isozyme/protein locus pairs: (*Got-2* – *Adh-1*),

Table 3. Two-point linkage data for loci associated with the phaseolin (*Pha*) linkage group. The upper right corner contains data from the BC_1 and the lower left corner for the F_2

		Est-2	*Pha*	*Spe*	*P*	*Spa*	*Spba*	*Mdh-1*
Est-2	'p'	–	0.11	0.18	0.24	0.26	–	–
	cM	–	11.18	19.19	26.02	28.27	–	–
	LOD	–	12.36	7.74	4.17	4.42	–	–
Pha	'p'	0.11	–	0.07	0.19	0.22	–	–
	cM	10.87	–	6.74	20.47	23.86	–	–
	LOD	13.57	–	17.52	5.85	6.39	–	–
Spe	'p'	0.16	0.10	–	0.21	0.18	–	–
	cM	16.25	10.03	–	22.26	18.61	–	–
	LOD	9.75	15.29	–	5.25	8.80	–	–
P	'p'	–	0.18	0.25	–	0.05	0.17	–
	cM	–	18.50	28.14	–	4.51	17.37	–
	LOD	–	7.33	3.79	–	14.85	6.95	–
Spa	'p'	–	0.31	–	0.13	–	0.13	–
	cM	–	36.91	–	13.31	–	13.84	–
	LOD	–	4.04	–	10.26	–	11.51	–
Spba	'p'	–	–	–	0.26	0.18	–	0.21
	cM	–	–	–	29.37	18.61	–	22.75
	LOD	–	–	–	3.20	8.37	–	6.75
Mdh-1	'p'	–	–	–	–	0.29	0.21	–
	cM	–	–	–	–	33.88	22.14	–
	LOD	–	–	–	–	4.95	7.02	–

'p': recombination fraction; cM: centiMorgan distance; LOD: maximum LOD score

(*Aco-2* – *Dia-1*), and (*Est-2* – *Pha*). A linkage cluster was detected for a group of seed proteins: (*Spd*/*Spf*/*Sph*) – 5 cM – *Spca*. Furthermore, five additional seed proteins, the *P* locus, and *Mdh-1* were found in the linkage group of the *Pha* locus. The two-point linkage results from both progenies (BC_1 and F_2) for the members of this linkage group are shown in Table 3.

"Mapmaker" assigned seven loci (*Spa*, *Spe*, *Spba*, *Pha*, *P*, *Est-2*, and *Mdh-1*) to one linkage group based on two- and three-point linkage analyses. Inspection of these results led to the selection of the first six loci for determination of gene order. With the multipoint mapping capability, "Mapmaker" evaluates all possible gene orders and then selects the best 20 orders. To reduce the number of permutations and save computer time, the *Mdh-1* locus was not included. After the gene order was established, this locus was positioned based on the two-point linkage results. The best gene orders were identical in both the BC_1 and the F_2-*Est-2* – *Pha* – *Spe* – *P* – *Spa* – *Spba*. However, it should be pointed out that in the BC_1 this order was only 2.5 times more likely than the second order, in which the positions of *P* and *Spa* were reversed. The possibility of this order reversal was not apparent in the F_2 data. The distances between loci were calculated using the simultaneous multipoint linkage capability of "Mapmaker." The Kosambi (1944) function was selected to estimate map distances. Simlar distances between loci were calculated with both progenies; however, the major discrepancy was observed for the distance between *P* and *Spa* (see Table 3). The distances between loci were averaged for the two progenies (Fig. 1).

Discussion

Two interesting observations can be made about the seed proteins described in this project. First, the polymorphisms detected by SDS-PAGE suggest that insertions and/or deletions in the coding regions may be responsible for the size differences; however, one cannot disregard the possibility of critical amino acid changes, which can alter mobility in an SDS gel (Ferl 1985). Second, some of the seed proteins appear to be organized in clusters in the bean genome. These observations raise questions about the nature of these polymorphisms and the bases for cluster organization. Proteins from a single cluster may be related and may have arisen by tandem gene duplication and subsequent divergence, as proposed for the phaseolin multigene family (Talbot et al. 1984).

The BC_1 and F_2 differed in the loci that displayed significant deviations from Mendelian ratios. Although there was an excess of 'XR-235-1' allele homozygotes at the *Est-2* locus in the F_2, defficiencies of 'XR-235-1' homozygotes were detected for *Spba* in the BC_1 and for *Spa* in the F_2. These differences suggest the possibility of specific interaction between the genomes of the two bean groups. Noteworthy is the fact that the deviations in the same direction occurred in two neighboring loci. Deviations from Mendelian ratios have been detected previously in bean progenies resulting from crosses between accessions from middle America and those of Andean origin (Koenig and Gepts 1989). Furthermore, distorted ratios have been reported for interspecific progenies between *P. vulgaris* and *P. coccineous* (Smartt 1970). This type of deviation has also been detected in interspecific progenies in *Lycopersicon* (Rick 1969; Tanksley et al. 1982; Vallejos and Tanksley 1983). Thus, in this particular cross, some of the observed skewings may involve chromosome segments from the *P. coccineous* progenitor of 'XR-235-1,' whereas others may involve interactions between the two *P. vulgaris* gene pools.

The linkage group flanked by *Est-2* and *Mdh-1* and comprising ca. 80–100 cM is the largest identified for beans with molecular markers. The greater distances between loci observed in the F_2 suggests that perhaps recombination is not uniform in male and female gametes. The use of molecular markers and the multipoint mapping capability of "Mapmaker" have made it possible to define such a linkage group. This group appears to be that identified by Bassett (1988) as Group VII. This group is also the largest group that has been identified with morphological markers.

Acknowledgements. We would like to thank Dr. M. J. Basset (University of Florida) for the seed of the breeding line XR-235-1, and Drs. J. Thome and W. Roca (CIAT) for seed of the cultivar 'Calima.' We want to acknowledge the technical assistance of Mr. S. F. Camp. We would also like to thank Drs. M. J. Bassett, L. C. Hannah, and G. A. Moore for their valuable comments on the original manuscript. This work was supported in part by USAID grant DPE-5542-G-SS-7060-00, and is part of a collaborative program between the University of Florida and CIAT.

References

Bassett MJ (1988) Linkage mapping of marker genes in common bean. In: Gepts P (ed) Genetic resources of *Phaseolus* beans. Kluwer, Dordrecht, pp 329–353

Brown JWS, Ma Y, Bliss FA, Hall TC (1981 a) Genetic variation in the subunits of globulin-1 storage protein of French bean. Theor Appl Genet 59:83–88

Brown JWS, Bliss FA, Hall TC (1981 b) Linkage relationships between genes controlling seed proteins in French bean. Theor Appl Genet 60:251–259

Ferl RJ (1985) Discontinuities in the topography of alcohol dehydrogenase-sodium dodecyl sulphate complexes revealed by mutagenesis and proteolysis. Biochem J 226:853–858

Freytag GF, Bassett MJ, Zapata M (1982) Registration of XR-235-1-1 bean germplasm. Crop Sci 22:1268–1269

Koenig R, Gepts P (1989) Segregation and linkage of genes for seed proteins, isozymes, and morphological traits in common bean (*Phaseolus vulgaris*). J Hered 80:455–459

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Laemmli UK (1970) Cleavage of structural proteins during assembly of the head of bacteriophage T4. Nature 227:680–685

Lander ES, Green P (1987) Construction of multilocus linkage maps in humans. Proc Natl Acad Sci USA 84:2363–2367

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

Leakey CLA (1988) Genotypic and phenotypic markers in common bean. In: Gepts P (ed) Genetic resources of *Phaseolus* beans. Kluwer, Dordrecht, pp 245–327

Morrissey JH (1981) Silver stain for proteins in polyacrylamide gels: a modified procedure with enhanced uniform sensitivity. Electrophoresis 117:307–310

Smartt J (1970) Interspecific hybridization between cultivated American species of the genus *Phaseolus*. Euphytica 19:480–489

Rick CM (1969) Controlled introgressions of chromosomes of *Solanum pennellii* into *Lycopersicon esculentum:* segregation and recombination. Genetics 62:753–768

Talbot DR, Adang MJ, Slightom JL, Hall TC (1984) Size and organization of a multigene family encoding phaseolin, the major seed storage protein of *Physeolus vulgaris* L. Mol Gen Genet 198:42–49

Tanksley SD, Rick CM (1980) Genetics of esterases in species of *Lycopersicon*. Theor Appl Genet 56:209–219

Tanksley SD, Medina-Filho H, Rick CM (1982) Use of naturally occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. Heredity 49:11–25

Vallejos CE, Tank ·y SD (1983) Segregation of isozyme markers and cold tolerance in an interspecific cross of tomato. Theor Appl Genet 66:241–247

Vallejos CE, Chase CD (1991) Linkage between isozyme markers and a locus affecting seed size in *Phaseolus vulgaris* L. Theor Appl Genet 81:413–419

Weeden NF (1984) Linkage between the gene coding for the small subunit of ribulose bisphosphate carboxylase and the gene coding malic enzyme in *Phaseolus vulgaris* L. Annu Rep Bean Impr Coop 27:123–124

Weeden NF, Liang CY (185) Detection of a linkage between white flower color and *Est-2* in common bean. Ann Rep Bean Impr Coop 28:87–88

-33

Copyright © 1992 by the Genetics Society of America

A Molecular Marker-Based Linkage Map of *Phaseolus vulgaris* L.

C. Eduardo Vallejos,[1] Ney S. Sakiyama[2] and Christine D. Chase

Department of Vegetable Crops, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, Florida 32611

Manuscript received February 21, 1992
Accepted for publication March 31, 1992

ABSTRACT

A seed and flower color marker (*P*), nine seed protein, nine isozyme and 224 restriction fragment length polymorphism marker loci were used to construct a linkage map of the common bean, *Phaseolus vulgaris* L. ($n = 11$). The mapping population consisted of a backcross progeny between the Mesoamerican breeding line 'XR-235-1-1' and the Andean cultivar 'Calima'; the former was used as the recurrent parent. A bean *Pst*I genomic library enriched for single copy sequences (95%) was the source of DNA probes. Sixty percent of the probes tested detected polymorphisms betwen the parental genotypes with at least one of the four restriction enzymes used here (*Dra*I, *Eco*RI, *Eco*RV and *Hin*dIII). The computer software **Mapmaker** was used to determine the linkage relationships and linear order of segregating markers. These markers assorted into 11 linkage groups covering 960 cM of the bean genome. Partial linkage data were used to estimate the total length of the genome at 1200 cM. This estimate and that for the physical size of the genome yield an average ratio of 530 kb/cM. The relatively small size of the genome makes this crop species a good candidate for the isolation of genes via chromosome walking techniques.

PHASEOLUS vulgaris L., the common bean, is a diploid ($2n = 22$) legume with a relatively small genome [633 Mbp (0.66 pg)/1C] (ARUMUGANATHAN and EARLE 1991). This bean is an important source of dietary protein for over half a billion people in Africa and Latin America (PACHICO 1989), and hypocholesterolemic properties add to its dietary value (ANDERSON *et al.* 1984). However, in spite of the importance of this crop species, its genetics has been poorly characterized. For instance, only a few morphological and seed and flower color markers have been used to develop a rudimentary linkage map (BASSETT 1991). More recently, a few isozyme and protein markers have been added to the map (VALLEJOS and CHASE 1991a,b). Although cytogenetic studies have been hampered by the small size of the chromosomes (ZHENG *et al.* 1991), five primary trisomics have been characterized (ASHRAF and BASSETT 1987).

The construction of linkage maps based on DNA markers (restriction fragment length polymorphisms, RFLPs) has been accomplished for a number of species: *Arabidopsis* (CHANG *et al.* 1988; NAM *et al.* 1989), lettuce (LANDRY *et al.* 1987), maize (HELENTJARIS, WEBER and WRIGHT 1986), potato (BONIERBALE, PLAISTED and TANKSLEY 1988), rice (MCCOUCH *et al.* 1988), soybean (KEIM *et al.* 1990) and tomato (BERNATZKY and TANKSLEY 1986). Extensive polymorphism between the progenitors of the mapping population is a prerequisite for the construction of a comprehensive linkage map. Low levels of RFLPs have been detected within each of the gene pools of common bean (CHASE, ORTEGA and VALLEJOS 1991). This problem has been solved in other species through the development of mapping populations from interspecific crosses. However, this is not a viable option for the common bean. Although interspecific crosses between *P. vulgaris* and *P. coccineus* or *P. acutifolius* can be obtained (HUCL and SCOLES 1985), the resulting progenies are far from normal and of limited value for mapping purposes. However, moderate levels of polymorphism can be detected between the Mesoamerican and Andean gene pools of the common bean (CHASE, ORTEGA, and VALLEJOS 1991). In addition, the extent of detectable polymorphisms can be increased by the inclusion of breeding lines carrying introgressions from either *P. acutifolius*, or *P. coccineus*. We report the construction of a linkage map for *P. vulgaris* based mostly on RFLP markers.

MATERIALS AND METHODS

Parental genotypes: Single inbred representatives of each of the two major *P. vulgaris* gene pools were used to generate the backcross mapping population, *viz.*, 'XR-235-1-1' for the Mesoamerican gene pool and 'Calima' for the Andean gene pool. However, the Mesoamerican breeding line 'XR-235-1-1' (FREYTAG, BASSETT and ZAPATA 1982) carries some chromosome segments from *P. coccineus*. 'XR-235-1-1' was used as the pistillate parent to obtain the F_1 progeny. The resulting F_1 plants were used to pollinate the recurrent parent, 'XR-235-1-1.' The parental genotypes differ in a large number of morphological and molecular characters. These include seed size, seed pigmentation, plant morphol-

[1] To whom correspondence should be addressed.
[2] Present address: Conselho de Pesquisas, Universidade Federal de Vicosa, 36570 Vicosa, MG, Brazil.

©AP APRMV2 GEN7$$692G (2cm-2106/2120) qqdp. gene-7 159692 1k/xy 05-06-92 20-15-30

ogy, quantitative resistance to bacterial blight, phaseolin type and other seed proteins, isozymes, and DNA marker loci (VALLEJOS and CHASE 1991a,b; CHASE, ORTEGA and VALLEJOS 1991).

Seed color and protein markers: The *P* locus controls the development of pigmentation on the seed coat; the dominant allele is present in 'Calima,' the donor parent, and absent in 'XR-235-1-1.' Thus, the segregation at this locus was followed by recording the presence or absence of pigmentation on the seeds collected from BC_1 individuals. The segregation of seed proteins was followed using SDS-PAGE as described earlier (VALLEJOS and CHASE 1991b). The α-amylase inhibitor was monitored using a combination of sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis and Western analysis using polyclonal antibodies obtained from M. CHRISPEELS (MORENO and CHRISPEELS 1989).

RFLP marker analysis: A modified procedure of MURRAY and THOMPSON (1980) was used to isolate total DNA from 2–4 g of young leaf tissue. Briefly, a liquid nitrogen powder of this tissue was mixed with three volumes of extraction buffer (133 mM Tris·HCl (pH 7.8)/6.7 mM NaEDTA/0.95 M NaCl/1.33% Na Sarkosyl/1.33% mercaptoethanol) in a polypropylene tube and incubated at 65° for 30 min; the tubes were inverted gently every 5 min. The homogenate was then extracted with one volume of chloroform (CH_3Cl/octanol; 24:1) and the aqueous phase separated by centrifugation. DNA was precipitated from the aqueous phase by the addition of ⅔ volumes of isopropanol; the precipitate was transferred with a glass hook (made from a Pasteur pipet) to the first washing solution (76% ethanol/0.2 M Na acetate) and held for 30 min to a few hours. This washing step eliminated flavonoid pigments and some oligosaccharides that coprecipitated with the DNA. The DNA was finally transferred to the second washing solution (76% ethanol/10 mM ammonium acetate) for 30 sec. After eliminating most of the liquid by pressing the pellet against the tube walls, the pellet was transferred to a new tube where it was dried and the DNA was dissolved in TE buffer. DNA was quantitated by electrophoresing undigested aliquots in agarose gels along with undigested lambda DNA standards. Concentrations were estimated by comparing the fluorescence intensities of the ethidium bromide stained samples against those of the standards.

DNA samples were digested with restriction enzymes (*Dra*I, *Eco*RI, *Eco*RV and *Hin*dIII) under conditions recommended by the manufacturer (Life Technologies, Inc., Gaithersburg, Maryland). DNA restriction fragments (1.5–2.0 μg/lane) were separated in 0.9 or 1.0% agarose gels using a modified TAE buffer (100 mM Tris·acetate/1 mM NaEDTA/pH 8.1; MURRAY and THOMPSON 1980). DNA in the gels was stained with ethidium bromide and visualized under UV light. DNA in the gels was depurinated with 0.25 M HCl for 10 min, and then denatured in 0.4 N NaOH/1.5 M NaCl for 30 min. SOUTHERN (1975) blots were obtained by alkaline transfer (20 mN NaOH/0.6 M NaCl) overnight. Membranes (Nytran; Schleicher & Schuell, Keene, New Hampshire) were then washed in 2 × SSPE (1 × SSPE = 150 mM NaCl/10 mM NaH_2PO_4, pH 7.4 (NaOH)/1 mM EDTA), air dried for 30 min over filter paper, and then UV cross-linked in a stratabox at a setting of 1200 (Stratagene, La Jolla, California). This procedure allowed the reprobing of blots for at least 8–10 times.

Probes and hybridizations: A *Pst*I genomic library enriched for single copy sequences was used as the source of probes. The construction of this library has been described elsewhere (CHASE, ORTEGA and VALLEJOS 1991). The mungbean cDNA of the chlorophyll *a/b* binding protein was also included (THOMPSON *et al.* 1983).

Inserts for hybridizations were obtained via the polymerase chain reaction (SAIKI *et al.* 1988). Wells of a microtiter plate were loaded with 100 μl of LB containing ampicillin (50 μg/ml). Each well was inoculated with a different clone, and the plate was incubated overnight at 37° on a rotary shaker. Overnight cultures were transferred to 500-μl microcentrifuge tubes and the cells pelleted for 2 min at 11,000 × *g*. The supernate was removed and the cells were resuspended in 100 μl of distilled water. The cell suspension was freeze-thawed and then spun at 16,000 × *g* for 15 min. The aqueous supernate was transferred to a new tube and stored at −20°. Amplification of insert DNA was carried out in 100-μl reaction volumes containing: 10 mM Tris·HCl (pH 8.8), 50 mM KCl, 15 mM $MgCl_2$, 0.1% Triton X-100, 200 μM of each dNTP, 100 nM of each primer, 0.5 unit of *Taq* DNA polymerase (Promega, Madison, Wisconsin), and 5 μl of the aqueous supernate from the freeze-thawed bacterial cells. Primers were synthesized at the DNA Synthesis Core Laboratory, University of Florida, and corresponded to the 18-base sequences up- and downstream from the *Pst*I cloning site of the plasmid pTZ18R. Amplification was carried out in a thermocycler (Coy Laboratory Products, Inc., Ann Arbor, Michigan). The initial denaturation was at 94° for 2 min followed by 10 cycles of 1 min at 94°/1 min at 55°/3 min at 72°, 25 cycles of 1 min 94°/1 min at 55°/2.5 min at 72°, 10 cycles of 1 min at 94°/1 min at 55°/3 min at 72°. Amplified insert DNA was precipitated in ethanol and later resuspended in TE buffer. DNA concentration was estimated in ethidium bromide stained agarose gels. Insert DNA was labeled with [^{32}P]dCTP using a random primer kit (Boehringer Mannheim Biochemicals, Indianapolis, Indiana).

Hybridizations of Southern blots with ^{32}P-labeled probes were carried out in glass bottles mounted on the rotisserie of a Hyb-Aid oven (Hyb-Aid, Middlesex, United Kingdom). The hybridization solution contained 5 × SSPE, 5 × Denhardt's solution, 1% SDS, 100 ng/ml of carrier DNA, and 5–10 ng/ml labeled insert DNA (1 × Denhardt's solution = bovine serum albumin 0.02%/Ficoll 0.02%/polyvinylpyrrolidone 0.02%). Hybridizations were carried out at 65° for 16–20 hr. Blots were washed twice in 2 × SSPE/0.1% SDS for a few seconds followed by three washes with 2 × SSPE/0.1% SDS at 65° for 15 min each and a final wash with 0.5 × SSPE/0.1% SDS for 15 min at 65°. The membranes were wrapped with polyethylene film and the hybridizing fragments were visualized by autoradiography on X-Omat film (Kodak) using intensifying screens (Lightening Plus, Du Pont). Segregation data for the DNA markers was collected with high efficiency by carrying simultaneous hybridizations with three to five probes. Typically, 12 hybridizations were carried out in a single day with a total of 36 to 45 probes. For single probe hybridizations labeled DNA was used directly without purification. On the other hand, for multiple probe hybridizations, a mixture of selected probes were labeled in a single vial, and then cleaned by spun column dialysis (SAMBROOK, FRITSCH and MANIATIS 1989). This step was taken to avoid high background on the blots.

The library was screened to identify restriction enzyme-probe combinations that detected length polymorphisms between 'Calima' and 'XR-235-1-1.' Thus, test blots containing parental DNA digested with the four restriction enzymes were hybridized with each probe. Probes that detected polymorphic fragments with the same restriction enzyme on nonoverlapping areas of the blot were mixed to probe Southern blots of the segregating population. Multi-

©AP APRMV2 GEN7$$692G (2cm-2106/2120) qqdp. gene-7 159692 1k/xy 05-06-92 20-15-30

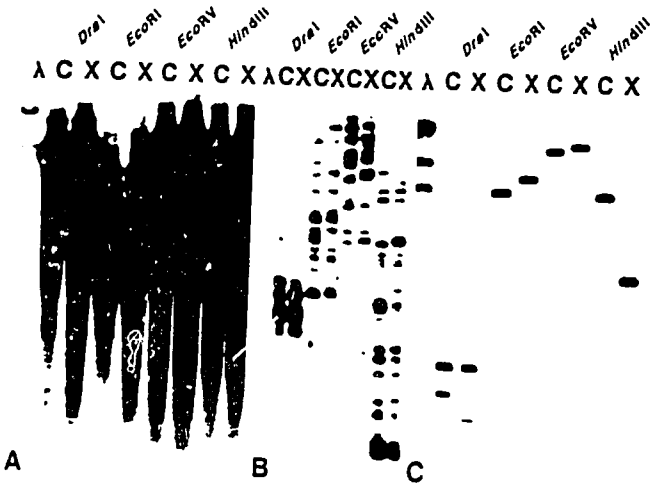


FIGURE 1.—Southern hybridizations of test blots used to screen the *Pst*I genomic library for single copy sequences that detected RFLPs between the parental genotypes. From left to right: lambda DNA (*Hin*dIII digest), 'Calima' (C) and 'XR-235-1-1' (X) digested with *Dra*I, *Eco*RI, *Eco*RV and *Hin*dIII. Hybridization patterns of: A, highly repetitive DNA; B, multigene family sequences; and C, single copy sequences.

point linkage analysis was conducted using **Mapmaker** (LANDER *et al.* 1987).

RESULTS

Library screen: Four restriction enzymes were used to screen 362 clones for their ability to detect polymorphisms between the parental genotypes. These clones had insert sizes ranging between 500 and 4000 bp. Eight (2.2%) of these clones produced hybridization patterns typical of highly repetitive DNA (Figure 1A), while nine (2.5%) clones yielded hybridization patterns expected of multigene families (Figure 1B). The hybridization pattern of the remaining 345 clones was typical of single copy sequences (Figure 1C). RFLPs between the parental genotypes were detected with 227 clones, although 11 of these clones were redundant. Redundancy was assumed when two clones yielded identical segregation, displayed identical hybridization patterns with the four enzymes, and had the same insert sizes. Thus, only 216 unique clones (60%) were used in this study. The four restriction enzymes differed in their ability to reveal polymorphisms between 'Calima' and 'XR-235-1-1': *Dra*I (42%), *Hin*dIII (53%), *Eco*RI (62%) and *Eco*RV (64%).

Monogenic segregation ratios: Ratios distorted from the expected 1:1 were detected at 19 loci in three different linkage groups (Table 1). The nine loci located centrally in group *C*, and a distal locus (*Bng35*) in group *K* displayed an excess of homozygotes. Examination of raw data indicated that the deviation detected for *Bng35* may be an artifact. In the four plants that were missing data for *Bng35*, all loci within 25 cM from *Bng35* were heterozygous. Thus, if we assume that these plants were also heterozygous at the *Bng35* locus, then, no significant deviations would be found at this locus. In contrast, the block of nine distal loci in group *H* showed a highly significant excess of heterozygotes.

Linkage analysis: The linkage relationships of the segregating markers were analyzed with the multipoint linkage analysis software **Mapmaker** version 1.9 (LANDER *et al.* 1987). Segregation data were collected efficiently by carrying simultaneous hybridizations with multiple probes (Figure 2). After obtaining all the two-point linkage data, the markers were sorted into distinct groups using an LOD of 4.0 and a distance of 25 cM (Haldane function) as default linkage criteria. This procedure resulted in the formation of 12 groups and six isolated markers; two of these markers showed no recombination between them and were linked to a third one. The three-point linkage data were obtained next. A **framework order** for each group was obtained using an LOD of 3.0 and a distance of 25 cM as linkage criteria. A three-point LOD exclusion threshold of −3 was also used for this process. Loci included in the framework are marked on the map (Figure 3). Linkage between the six isolated markers and the 12 linkage groups was tested relaxing the distance criterion while keeping the LOD criterion at 4.0. These markers were found to assort with three of the established linkage groups and were also assigned to the **framework** (*B*: *Bng71*, *Bng151*, *Bng160*; *E*: *Bng161*, *Bng162*; *I*: *Bng200*). The remaining loci were placed on the map using the **try** command. This command determines the relative likelihoods of maps where a marker is placed at different intervals of a gene order. A place was selected for a marker only when the log-liklihood for that order was at least 100 times better than the second best order. Markers for which an order could not be established with a log-likelihood greater than 2 were encased by brackets to indicate this fact (Figure 3). Tests for possible linkage between two of the 12 identified linkage groups were carried out because *P. vulgaris* is known to have 11 chromosomes. The **link** command was used with all possible group pair combinations to test for possible associations. Linkage was detected between two of the smallest linkage groups, although the LOD score was of only 2.22. *Bng7* and *Bng22* have been found to be *ca.* 15 cM (LOD = 4.4) apart in the analysis of a smaller backcross population ($n = 34$). For these reasons, these small groups have been assigned to the same linkage group (*F*). A thin line on the map was used to indicate the low LOD score (Figure 3). The distance between these subgroups, 36 cM, was calculated using the KOSAMBI (1944) function; this function is normally used with high recombination frequencies. The final linear orders were tested with the **ripple** command; this test consists of comparing the likelihood of the current map with those of maps obtained by permuting the orders of all adjacent triplets. The

©AP APRMV2 GEN7$$692G (2cm-2106/2120) qqdp. gene-7 159692 1k/xy 05-06-92 20-15-30

TABLE 1

List of loci displaying deviations from Mendelian ratios (1:1)

Group	Locus	x/x[a]	x/c	X^2	Group	Locus	x/x	x/c	X^2
C	*Bng124*	44	24	5.88*	H	*Bng65a*	24	44	5.88*
C	*Bng32*	44	24	5.88*	H	*Bng4*	24	44	5.88*
C	*Bng216*	43	25	4.76*	H	*Bng188*	22	46	8.47**
C	*Bng221*	43	25	4.76*	H	*Bng39*	21	47	9.94**
C	*Bng164*	44	24	5.88*	H	*Bng36*	22	46	8.47**
C	*Bng3*	43	24	5.39*	H	*Bng41*	20	48	11.53**
C	*Bng123*	43	25	4.76*	H	*Bng72*	23	45	7.12**
C	*Bng114*	44	24	5.88*	H	*Bng195*	20	48	11.53**
C	*Bng75*	43	25	4.76*	H	*Bng130*	19	49	13.24**
K	*Bng35*	40	24	4.00*					

[a] *x/x*: homozygous for XR-235-1-1 alleles; *x/c*: heterozygous.

ART
CP: insert 1 pt vert rule as marks.

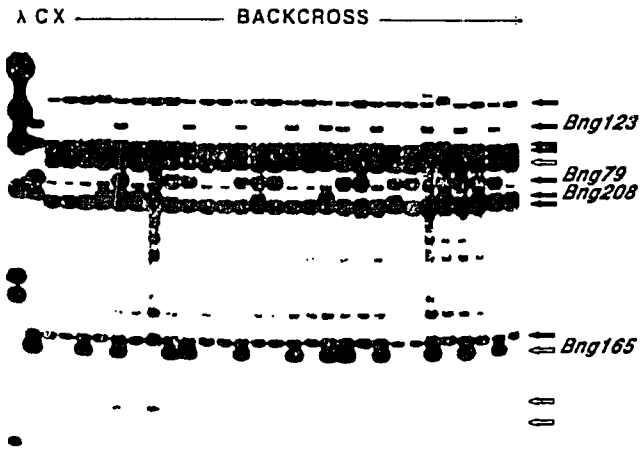


FIGURE 2.—Southern blot displaying the segregation of four RFLP loci. Lambda and bean genomic DNAs were digested with *Hin*dIII. From left to right: lambda fragments, 'Calima' (C), 'XR-235-1-1' (X) and 27 backcross individuals. The blot was probed simultaneously with radiolabeled inserts for *Bng79*, *Bng123*, *Bng165*, and *Bng208*. Restriction fragments detected with the same probe are marked with arrows of a single shading pattern. Arrows with a label are pointing at segregating fragments, whereas those without a label are pointing at recurrent or nonsegregating allelic fragments.

isozyme locus *Aco1* was assigned to group *D* and listed at the bottom of the group. This locus could not be placed at any specific location within this group because 8 out of the 68 BC_1 plants had missing data for this locus. However, data from a different population had previously placed *Aco1* between *Bng90* and *Bng180*.

The *P* locus, 17 protein, and 226 DNA marker loci presented in this report define 963 cM of the common bean genome. These markers have assorted into 11 linkage groups ranging in size from 56 to 104 cM: *A* (104.7), *B* (104), *C* (94.6), *D* (85.8), *E* (85.5), *F* (78), *G* (73.6), *H* (73.3), *I* (70.6), *J* (60.7) and *K* (56.2). The 244 markers fell into only 145 loci/locus clusters with the following distribution: 90 (62%) loci were occupied by a single marker, 40 (28%) by two markers, 8 (5.6%) by three markers, and seven clusters (4.8%) contained between 5 and 10 markers each. These clusters were located on four linkage groups and four of these clusters were detected in the smallest linkage group (*K*).

Five seed proteins (*Pha*, *Spa*, *Spb*, *Spba* and *Spe*), *Est2*, and the *P* locus were detected in the same linkage group (*A*) as previously reported (VALLEJOS and CHASE 1991b). In addition, we had also reported that three seed proteins named *Spd*, *Spf* and *Sph* were linked (VALLEJOS and CHASE 1991b). *Spf* and *Sph* have been renamed *AAI1* and *AAI2*, respectively, because they comigrated and cosegregated with two bands detected by western analysis with a polyclonal antibody produced against the bean α-amylase inhibitor (MORENO and CHRISPEELS 1989). The mung bean cDNA clone of the chlorophyll *a/b* binding protein (pMB123, THOMPSON *et al.* 1983) detected 10 *Eco*RV fragments in each of the parents, 'Calima' and 'XR-235-1-1'; these fragments ranged in size between 1.4 and 11 kb. Five of those fragments were common to both genotypes. Two fragments unique to 'Calima' (9.4 and 8.8 kb) cosegregated and were assigned to locus *Cab1* in group *B*. Three other fragments (6.4, 5.9 and 1.5 kb) also cosegregated, but assorted independently from the previous ones and were assigned to locus *Cab2* in group *D*. Eight of the 216 genomic clones detected two independent loci each and were given the letters *a* or *b* after the number to differentiate them. In all cases the homologous sequences were located on different linkage groups suggesting duplication/translocation events.

Genome size: The length (cM) of the *P. vulgaris* genome was estimated using the method of HULBERT *et al.* (1988). This method is based on calculating the probability [$P(T, G)$] that a pair of linked markers randomly chosen from a genome of size (G) has an LOD score ≥ T. A first-order approximation of $P(T, G)$ is given by the expression $2x/G$, where x is the distance (cM) between the markers; this distance is relatively small in relation to the size of the genome. The expected frequency (K) of locus pairs having an LOD ≥ T is $(N)P(T, G)$, where N is the number of

X

(N)

AUTHOR:

Phaseolus vulgaris Linkage Map

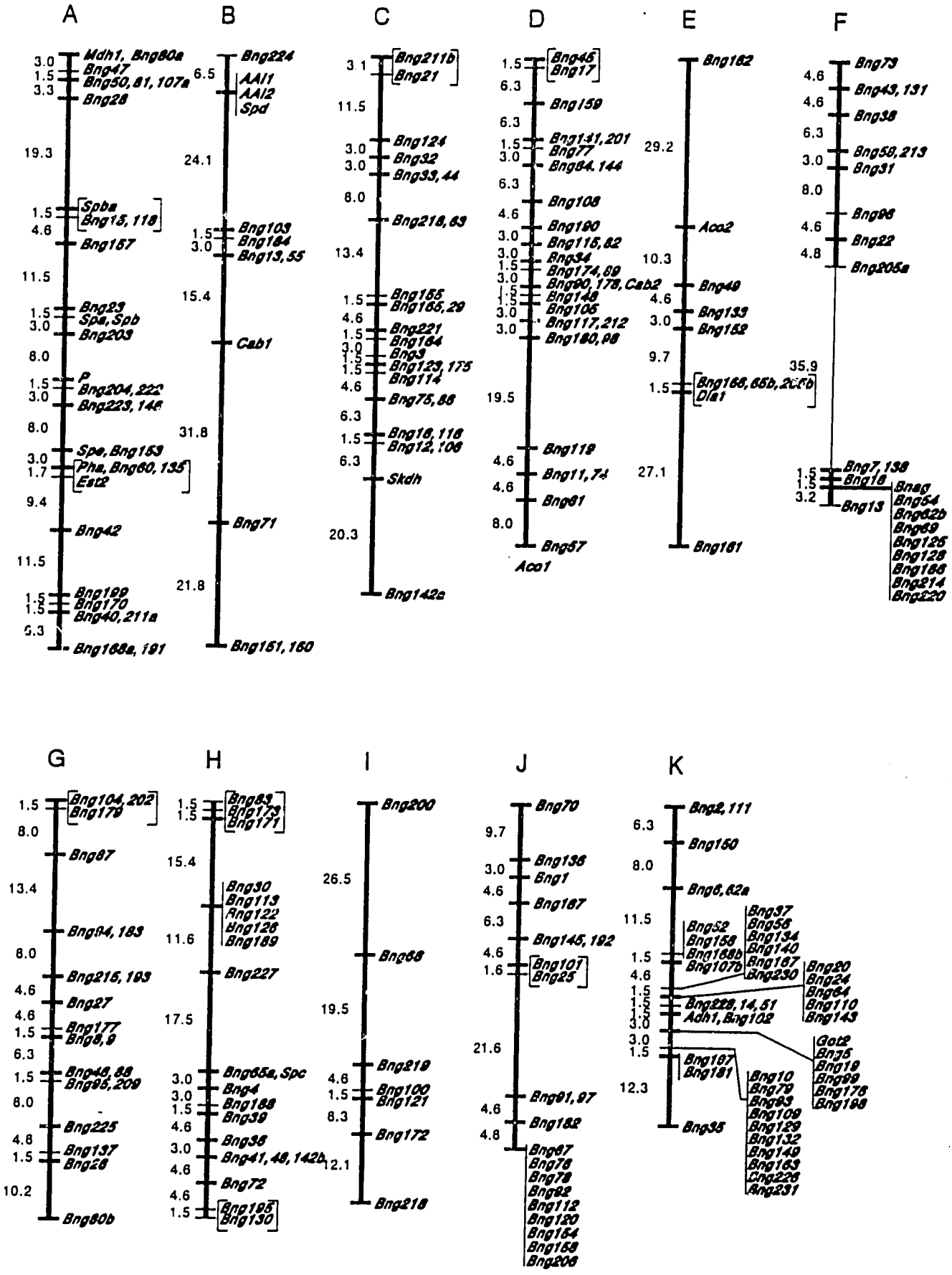


FIGURE 3.—Linkage map of *P. vulgaris* L., the common bean. With the aid of **Mapmaker** (LANDER *et al.* 1988) 244 markers were assigned to 145 loci/locus clusters on 11 linkage groups (*A–K*). The exact position of *Aco1* on group *D* could not be located due to incomplete data for this locus; *Aco1* is listed at the bottom of the group. A **framework** order was obtained after grouping the markers (LOD 4.0, 25 cM), and

©AP APRMV2 GEN7$$692G (2cm-2106/2120) qqdp. gene-7 159692 1k/xy 05-06-92 20-15-30

markers. **Mapmaker** was used to determine that there were 433 locus pairs each within 25 cM and with an LOD score ≥ 4.0. These pairs were extracted from the 145 loci identified in this study. Only a single member of a cluster was used in the calculations because inclusion of all cluster members would have artificially increased the probability estimate. With this information at hand and the relationship $G = (N)(2x/K)$, we have estimated the bean genome to be approximately 1,200 cM in length. Thus, the markers presented here cover 80% of the estimated bean genome.

DISCUSSION

Reassociation kinetics measurements have shown that 60% of the common bean genome contains single copy sequences (TALBOT *et al.* 1984). Thus, a genomic library of small fragments would be expected, in the best of cases, to have 60% of the clones from the slow reassociating class of DNA. However, the results from the library screen clearly demonstrated that the size selected *Pst*I fragments of bean genomic DNA were highly enriched in single copy sequences (95%). BURR *et al.* (1988) pointed out that small DNA fragments obtained with restriction enzymes sensitive to *C*-methylation, such as *Pst*I, would yield genomic libraries enriched in single copy sequences. This phenomenon has been observed previously in tomato (TANKSLEY *et al.* 1987) and maize (BURR *et al.* 1988), but not in rice (MCCOUCH *et al.* 1988). New flow cytometry measurements indicate that the bean genome is intermediate in size between rice and tomato (ARUMUGANATHAN and EARLE 1991). This observation suggests the question of whether there is a relationship between genome size and the extent of *C*-methylation in plant repetitive DNA.

We have constructed a linkage map with 145 loci/locus clusters. These loci/locus clusters are occupied by 244 markers. An interesting feature of this map is the uneven distribution of markers on the linkage groups; approximately 20% of the markers were found in 7 clusters. Groups *B*, *E* and *I* each have seven to eight loci which delimit 104, 85 and 70 cM, respectively. Groups *A* (104.7 cM) and *D* (85.8 cM) are more densely populated and contain approximately 30–33 markers each which are distributed among 20 loci/locus clusters. Finally, group *K* is the smallest linkage group (56.2 cM), but contains 44 markers, 27 of which are organized in 4 separate clusters. Several nonexclusive explanations could account for the disparity in the distribution of these markers. If all regions of the genome were to be evenly represented in our library, then one must assume that small *Pst*I fragments (500–4000 bp) containing single copy sequences are evenly distributed throughout the bean genome and that the number of clones used to construct the linkage map was large enough to give an unbiased sample. It has been estimated that about 60,000 genes are expressed during the life cycle of a plant (KAMALAY and GOLDBERG 1980). Thus, in comparison, the small number of markers used in this project may represent a biased sample of the bean genome and be in part responsible for the uneven distribution of markers. In addition, one must consider that the library was screened for probes that detected polymorphisms between the Mesoamerican breeding line 'XR-235-1-1' and the Andean cultivar 'Calima.' Therefore, densely populated chromosome segments may correspond to regions of the genome that are variable between the two gene pools of beans, whereas sparsely populated segments may carry more conserved sequences. Differences in recombination frequencies throughout the genome may also account for marker density differences among different chromosome segments. The fact that 'XR-235-1-1' has *P. coccineus* introgression must also be taken into account. Some bias was expected towards sequences that detect polymorphisms between *P. vulgaris* and *P. coccineus*, but only to the extent by which the latter has been introgressed into 'XR-235-1-1.' Small chromosome inversions can also act as suppressors of recombination and explain the presence of clusters. Finally, the recombination frequency changes from 10 to 50% in a stretch of chromosome 2 of tomato in two different interspecific crosses (PATTERSON *et al.* (1989). It remains to be determined whether differences in recombination values in certain chromosome regions are due to the presence of *P. coccineus* chromosome segment, to the combination of Andean and Mesoamenrican backgrounds, or both.

Strong distortions of Mendelian ratios detected for a distal section of group *H* have been observed previously for a marker from this group (*Spc*) in two different progenies of 'XR-235-1-1' and 'Calima' (VALLEJOS and CHASE 1991b). Significant skewings have been reported also in progenies obtained between Mesoamerican and Andean genotypes (KOENIG and GEPTS 1989), and for interspecific crosses between *P. vulgaris* and *P. coccineus* (SMARTT 1970). These distortions could be interpreted as the result of pre- or postsyngamic selection of certain allelic com-

then ordering each group (LOD 3.0, 25 cM, exclusion LOD −3.0). Markers that belong to the **framework** order have been marked with thicker crossbars. Loci for which an order could not be established (LOD < 2.0) have been enclosed by brackets. Numbers on the left of a linkage group represent map distances in cM. Loci at distances that exceed 25 cM in groups *B*, *E* and *I* have an LOD score > 4.0 and have been included in the **framework**. The distance (35.9 cM) between *Bng205a* and *Bng7* (group *F*) was calculated using the Kosambi function. Although an LOD of 2.22 was found for this linkage, other independent data supports the assignment to this group.

©AP APRMV2 GEN7$$692G (2cm-2106/2120) qqdp. gene-7 159692 1k/xy 05-06-92 20-15-30

binations. For instance, gametes are eliminated when certain allelic combinations are present at the *Ge* locus of tomato (RICK 1966). An alternative explanation could be found in chromosome inversions. CHENG, BASSETT and QUESENBERRY (1981) have presented evidence for chromosome inversions in interspecific crosses of *P. vulgaris* and *P. coccineus*. Indeed, inversions can also explain reduced recombination frequencies and the appearance of clusters. We are currently investigating these issues in progenies between pure *P. vulgaris* lines obtained within and between gene pools.

The linear order for the protein/isozyme loci in group *A* is the same as that found previously (VALLEJOS and CHASE 1991B), ALTHOUGH THE DISTANCES AROUND *Est2* are at variance with those previously obtained due mainly to missing data points for this locus. The distances for the locus pairs *Aco2-Dia1* (group *E*) and *Adh1-Got2* (group *K*) were essentially as previously reported (VALLEJOS and CHASE 1991a).

The Genetic Resources Unit of CIAT manages a large and well diversified bean germ plasm collection with over 40,000 accessions (HIDALGO 1991). It may be possible to tag many genes of agricultural importance found in that collection using the markers presented here. Many of these genes, such as those responsible for disease resistance, have eluded isolation via standard cDNA cloning approaches. Map-based cloning has been suggested as an alternative approach for the molecular cloning of these genes (TANKSLEY *et al.* 1989). The relatively small size of the bean genome makes this species a good candidate for gene isolation via chromosome walking techniques (ROMMENS *et al.* 1989). We have estimated the length of the bean genome to be approximately 1200 cM. This estimate and that of the physical size of the bean genome (637 Mbp) lead to an average estimate of 530 kb/cM. However, an accurate estimate of this relationship would be needed for a targeted region before any of these efforts are initiated in beans. For instance, a ratio of 14 kb/cM has been detected within the *bronze* locus of maize (DOONER 1986); this is about two orders of magnitude smaller than the average ratio. On the other hand, a ratio severalfold above the average has been detected around the *Tm-2a* locus of tomato (GANAL, YOUNG and TANKSLEY 1989).

We wish to thank SAMUEL F. CAMP and CASSIA C. SAKIYAMA for their excellent technical assistance in this project. We also thank M. J. BASSET (University of Florida) for providing seed of 'XR-235-1-1,' M. J. CHRISPEELS (University of California) for providing the serum for the α-amylase inhibitor, and W. F. THOMPSON (North Carolina State University) for providing pM2123, the clone for the chlorophyll *a/b*-binding protein. We extend our appreciation to M. J. BASSETT, L. C. HANNAH and Z. H. YU for their critical reviews of the manuscript and their valuable suggestions. This project represents a collaborative effort between the University of Florida and the Biotechnology Research Unit of the International Center for Tropical Agriculture (CIAT; Cali, Colombia). This project was partially funded by U.S. Agency for International Development Grant DPE-5542-G-SS-7060-00. The bean genomic clones (*Bngs*) will be available through the Biotechnology Research Unit of CIAT. Florida Agricultural Experiment Station Journal Series No. R-02226.

LITERATURE CITED

ANDERSON, J. W., L. STORY, B. SIELING, W-J. L. CHEN, M. S. PETRO and J. STORY, 1984 Hypocholesterolemic effects of oat-bran or bean intake for hypercholesterolemic men. Am. J. Clin. Nutr. **40:** 1146–1155.

ARUMUGANATHAM, K., and E. D. EARLE, 1991 Nuclear DNA content of some important plant species. Plant Mol. Biol. Rept. **9:** 208–218.

ASHRAF, M., and M. J. BASSETT, 1987 Five primary trisomics from translocation heterozygote progeny in common bean. *Phaseolus vulgaris* L. Theor. Appl. Genet. **74:** 346–360.

BASSETT, M. J., 1991 A revised linkage map of common bean. HortScience **26:** 834–836.

BERNATZKY, R., and S. D. TANKSLEY, 1986 Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genetics **112:** 887–899.

BONIERBALE, M., R. PLAISTED and S. D. TANKSLEY, 1988 Construction of comparative genetic maps of potato and tomato based on a common set of cloned sequences. Genetics **120:** 1095–1103.

BURR, B., F. A. BURR, K. H. THOMPSON, M. C. ALBERTSON and C. W. STUBER, 1988 Gene mapping with recombinant inbreds in maize. Genetics **118:** 519–526.

CHANG, C., J. L. BOWMAN, A. W. DEJOHN, E. S. LANDER and E. M. MEYEROWITZ, 1988 Restriction fragment length polymorphism linkage map for *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **85:** 6856–6860.

CHASE, C. D., V. M. ORTEGA and C. E. VALLEJOS, 1991 DNA restriction fragment length polymorphisms correlate with isozyme diversity in *Phaseolus vulgaris* L. Theor. Appl. Genet. **81:** 806–811.

CHENG, S. S., M. J. BASSETT and K. H. QUESENBERRY, 1981 Cytogenetic analysis of interspecific hybrids between common bean and scarlet runner bean. Crop Sci. **21:** 75–79.

DOONER, H. K., 1986 Genetic fine structure of the *bronze* locus in maize. Genetics **113:** 1021–1036.

FREYTAG, G. F., M. J. BASSETT and M. ZAPATA, 1982 Registration of XR-235-1-1 bean germ plasm. Crop Sci. **22:** 1268–1269.

GANAL, M. W., N. D. YOUNG and S. D. TANKSLEY, 1989 Pulsed field electrophoresis and physical mapping of large DNA fragments in the *Tm-2a* region of chromosome 9 in tomato. Mol. Gen. Genet. **215:** 395–400.

HELENTJARIS, T., D. F. WEBER and S. WRIGHT, 1986 Use of monosomics to map cloned DNA fragments in maize. Proc. Natl. Acad. Sci. USA **83:** 6035–6039.

HIDALGO, R., 1991 CIAT's world *Phaseolus* collection, pp. 163–197 in *Common Beans: Research for Crop Improvement*, edited by A. VAN SCHOONHOVEN and O. VOYSEST. C. A. B. International, Wallingford, U.K.

HUCL, P., and G. J. SCOLES, 1985 Interspecific hybridizations in the common bean: a review. HortScience **20:** 352–357.

HULBERT, S. H., T. W. ILOT, E. J. LEGG, S. E. LINCOLN, E. S. LANDER and R. W. MICHELMORE, 1988 Genetic analysis of the fungus, *Bremia lactucae*, using restriction fragment length polymorphisms. Genetics **120:** 947–958.

KAMALAY, J. C., and R. B. GOLDBERG, 1980 Regulation of structural gene expression in tobacco. Cell **19:** 935–946.

KEIM, P., B. W. DIERS, T. C. OLSON and R. C. SHOEMAKER, 1990 RFLP mapping in soybean: association between marker loci and variation in quantitative traits. Genetics **126:** 735–742.

KOENIG, R., and P. GEPTS, 1989 Segregation and linkage of genes

©AP APRMV2 GEN7$$692G (2cm-2106/2120) qqdp. gene-7 159692 1k/xy 05-06-92 20-15-30

for seed proteins, isozymes, and morphological traits in common bean (*Phaseolus vulgaris*). J. Hered. **80:** 455–459.

KOSAMBI, D. D., 1944 The estimation of map distances from recombination values. Ann. Eugen. **12:** 172–175.

LANDER, E. S., P. GREEN, J. ABRAHAMSON, A. BAARLOW, M. J. DALY, S. E. LINCOLN and L. NEWBURG, 1987 MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1:** 174–181.

LANDRY, B. S., R. V. KESSELI, B. FARRARA and R. W. MICHELMORE, 1987 A genetic map of lettuce (*Lactuca sativa* L.) with restriction fragment length polymorphisms, isozymes, disease resistance, and morphological markers. Genetics **116:** 331–337.

MCCOUCH, S. R., G. KOCHERT, Z. H. YU, Z. Y. WANG, G. S. KHUSH, W. R. COFFMAN and S. D. TANKSLEY, 1988 Molecular mapping of rice chromosomes. Theor. Appl. Genet. **76:** 815–829.

MORENO, J., and M. J. CHRISPEELS, 1989 A lectin gene encodes the α-amylase inhibitor of the common bean. Proc. Natl. Acad. Sci. USA **86:** 7885–7889.

MURRAY, M. G., and W. F. THOMPSON, 1980 Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. **19:** 4321–4325.

NAM, H.-G., J. GIRAUDAT, B. DEN BOER, F. MOONAN, W. D. B. LOOS, B. M. HAUGE and H. M. GOODMAN, ~~1991~~ 1989 Restriction fragment length polymorphism linkage map of *Arabidopsis thaliana*. Plant Cell **1:** 699–705.

PACHICO, D., 1989 Trends in world common bean production, pp. 1–8 in *Bean Production Problems in the Tropics*, edited by H. F. SCHWARTZ and M. A. PASTOR-CORRALES. CIAT Press, Cali, Colombia.

PATERSON, A. H., E. S. LANDER, J. D. HEWITT, S. PETERSON, S. E. LINCOLN and S. D. TANKSLEY, 1988 Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature **335:** 721–726.

RICK, C. M., 1966 Abortion of male and female gametes in the tomato determined by allelic interactions. Genetics **53:** 85–96.

ROMMENS, J. M., M. C. IANNUZZI, B-S. KEREM, M. L. DRUMM, G. MELMER, M. DEAN, R. ROZMAHEL, J. L. COLE, D. KENNEDY, N. HIDAKA, M. ZSIGA, M. BUCHWALD, J. R. RIORDAN, L-C. TSUI and F. S. COLLINS, 1989 Identification of the cystic fibrosis gene: chromosome walking and jumping. Sience **245:** 1059–265.

SAIKI, R. K., D. H. GELFAND, S. STOFFEL, S. J. SCHARF, R. HIGUCHI, G. T. KOPN, K. B. MULLIS, and H. A. ERLICH, 1988 Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239:** 487–491.

SAMBROOK, J., E. F. FRITSCH and T. MANIATIS, 1989 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

SMARTT, J., 1970 Interspecific hybridization between cultivated American species of the genus *Phaseolus*. Euphytica **19:** 480–489.

SOUTHERN, E. M., 1975 Detection of specific sequences among DNA fragments separated by electrophoresis. J. Mol. Biol. **98:** 503–517.

TALBOT, D. R., M. J. ADANG, J. L. SLIGHTON and T. C. HALL, 1984 Size and organization of a multigene family encoding phaseolin, the major seed storage protein of *Phaseolus vulgaris* L. Mol. Gen. Genet. **198:** 42–49.

TANKSLEY, S. D., J. MILLER, A. PATERSON and R. BERNATZKY, 1987 Molecular mapping of plant chromosomes, pp. 157–173 in *Chromosome Structure and Function*, edited by J. P. GUSTAFSON and R. A. APPLES. Plenum Press, New York.

TANKSLEY, S. D., N. D. YOUNG, A. H. PATERSON and M. W. BONIERBALE, 1989 RFLP mapping in plant breeding: new tools for an old science. Bio/Technology **7:** 257–264.

THOMPSON, W. F., M. EVERETT, N. O. POLANS and R. A. JORGENSEN, 1983 Phytochrome control of RNA levels in developing pea and mung bean leaves. Planta **158:** 487–500.

VALLEJOS, C. E., and C. D. CHASE, 1991a Linkage between isozyme markers and a locus affecting seed size in *Phaseolus vulgaris* L. Theor. Appl. Genet. **81:** 413–419.

VALLEJOS, C. E., and C. D. CHASE, 1991b Extended linkage map for the phaseolin linkage group of *Phaseolus vulgaris* L. Theor. Appl. Genet. **82:** 353–357.

ZHENG, J., M. NAKATA, H. UCHIYAMA, H. MORIKAWA and R. TANAKA, 1991 Giemsa C-banding pattern in several species of *Phaseolus* L. and *Vigna* Savi, fabacea. Cytologia **56:** 459–466.

Communicating editor: S. D. TANKSLEY

Phaseolus vulgaris

The common bean

C. Eduardo Vallejos

Department of Horticultural Sciences & Graduate Program in Plant Molecular and Cellular Biology. 1255 Fifield Hall. University of Florida. Gainesville, FL 32611. USA.

ABSTRACT

The common bean, *Phaseolus vulgaris* L., is a diploid organism (n=11) with a small genome (0.66 pg/1 C). Segregation data obtained from a backcross population between the mesoamerican breeding line 'XR-235-1-1' and the andean cultivar 'Calima' have been used to construct a linkage map based primarily on RFLP markers. A genomic library of *Pst*I size-selected fragments, enriched for single copy sequences, was the source of probes used for this purpose. The map encompasses 980 cM with a total of 250 markers assigned to 11 linkage groups. These markers correspond to 227 RFLP loci, two chlorophyll a/b binding protein loci, one pigmentation locus (*P*), nine isozymes, nine seed proteins, and two RAPD markers.

A brief discussion is included about some relevant problems in common bean research that can be effectively addressed with RFLP markers. These include the identification genetic factors that restrict gene flow between mesoamerican and andean gene pools and from the ancillary gene pools, QTL analysis of yield components, and the localization and genomic isolation of disease resistance genes.

INTRODUCTION

Gregor Mendel (1866) conducted the first genetic analysis of common beans. Mendel studied the inheritance of growth habit, and pod color and shape in a progeny between *P. vulgaris* and *P. nanus* (=*P. vulgaris*, bush type) in order to confirm his findings with peas. Unfortunately, further studies on the inheritance of flower and seed coat color were hampered by his use of interspecific hybrids between *P. nanus* and *P. multiflorus* (=*P. coccineus*), these are now known to yield aberrant ratios. Later on, Shaw and Norton (1918) used intraspecific crosses and determined that pigmentation and pigmentation patterns of the seed coat are controlled by multiple independent factors. A few years later Karl Sax (1923) began to identify the multiple components that determine the inheritance of these traits. A single factor was identified as responsible for pigmentation, while two linked factors were identified to control mottling; this appears to be the first report of linkage in beans. Furthermore, Sax (1923) was the first to report linkage between a Mendelian character (seed coat pigmentation) and a QTL (for seed size). Although the common bean was used as experimental material at the inception of genetics, its genetic characterization has lagged behind that of many other crop species.

44

The common bean is a diploid organism (n=11) with relatively small chromosomes (Zheng *et al.*, 1991) and a small genome estimated by flow cytometry to be 637 Mbp or 0.66 pg /1C (Arumuganathan & Earle, 1991). It has also been estimated, via DNA reassociation kinetics, that 60% of the genome comprises of single copy sequences (Talbot *et al.*, 1984). The chromosome number and genome size of *P. vulgaris* are very similar to those of *P. acutifolius* and *P. coccineus* (Arumuganathan & Earle, 1991), these two species are partially compatible with the common bean and represent an important source of germ plasm for plant improvement (Hucl & Scoles, 1985). A rudimentary linkage map has been developed through the years with mostly morphological markers and a few isozymes (Bassett, 1991; Vallejos & Chase, 1991a, b). Electrophoretic analysis of the major seed storage protein (phaseolin) and a group of isozymes has led to the identification of a mesoamerican and an andean gene pool (Gepts *et al.*, 1986; Koenig & Gepts, 1989). Moreover, preliminary survey of DNA restriction fragment length polymorphisms showed that DNA probes can be used to differentiate the two groups because low levels of polymorphism were detected within each gene pool, but moderate levels were found between the gene pools (Chase *et al.*, 1991).

45

CONSTRUCTION OF THE LINKAGE MAP

Recently, a linkage map, based mainly on RFLP markers, has been constructed using a backcross progeny between a mesoamerican breeding line and an andean cultivar (Vallejos *et al.*, 1992). A genomic library of size-selected (500-4000 bp) *Pst*I fragments was the main source of probes (Chase *et al.*, 1991). This library was enriched for single copy fragments as 95% of the 362 clones tested yielded hybridization patterns typical of single copy sequences. Sixty percent of the clones tested revealed polymorphisms between the parental genotypes ('XR-235-1-1' and 'Calima') with at least one of four restriction enzymes (*Dra*I, *Eco*RI, *Eco*RV, and *Hin*dIII). More recently, 28 of the clones that had not detected RFLPs with any of the previously used enzymes were tested with four new enzymes: *Bam*HI, *Bgl*II, *Kpn*I, and *Xba*I. About 50 % of these clones have revealed polymorphisms with at least one of these enzymes. These results bring the estimated polymorphism between the parental genotypes at 80 % with at least one of the 8 restriction enzymes. These enzymes differed in their ability to detect polymorphisms between the parental genotypes: *Kpn*I (40%); *Drab* (42%); *Bam*HI, *Bgl*II and *Hin*dIII (53%); *Eco*RI (62%); and *Eco*RV (64%).

46

Segregation data obtained from a backcross between the mesoamerican breeding line 'XR-235-1-1' and the andean cultivar 'Calima' have been used to assemble a linkage map for the common bean (Vallejos *et al.*, 1992) with the aid of the computer software **Mapmaker** (Lander *et al.*, 1987). This map (Fig. 1) was constructed by first establishing a framework order using stringent 'linkage criteria' (LOD 3.0, 25 cM, and three point LOD exclusion threshold -3.0). Additional loci have been added later using either the **try** or **place** commands. The linkage between *Bng205a* and *Bng7* in group *F* has an LOD score of 2.22; although weak, this linkage is supported by data from two different progenies (Vallejos *et al.*, 1992). The current map comprises 227 RFLP loci. Eight out of the 219 genomic clones listed in Table 1 hybridize to homologous sequences located in different linkage groups. The map also includes two loci of the chlorophyll a/b binding protein in groups *B* and *D* ; these loci were identified with pMB123, a mungbean cDNA clone (Thompson *et al.*, 1983). Also included are: one phenotypically identified pigmentation locus (*P*), nine isozymes and nine seed proteins (Table 2). The seed proteins are: phaseolin, the α-amylase inhibitor proteins (identified by western blots with antibody provided by M. Chrispeels (Moreno & Chrispeels, 1989), and other globulins (Vallejos & Chase, 1991b). Electrophoretic variation at the protein level could be due to either variation oat the DNA level and/or variation in post-translational events--protease processing and/or glycosylation. Nevertheless,

Southern analysis of genomic blots with a phaseolin clone (Sun *et al.*, 1981) has showed perfect co-segregation between a restriction fragment identified by this clone and phaseolin protein bands identified by SDS-PAGE. Thus, these results strongly suggest that variation at the protein level in phaseolin is due to variation at the DNA level and that the locus mapped with protein data corresponds to the structural gene of phaseolin. We have recently begun to use RAPD markers (Williams *et al.*, 1990) and have added two of these markers to the map: *OA1a* in *D* and *OA10a* in *F* (Random primers were obtained from Operon Technologies, Inc; Alameda, CA) (Z.H. Yu & C.E. Vallejos, unpubl.). In summary, a linkage map of the common been has been assembled that includes 250 markers assigned to 150 loci/locus clusters and comprise 980 cM, or approximately 82% of the bean genome. The development of the bean genomic clones at the University of Florida was financed in part by a grant from the Agency for International Development (AID). These clones have been transferred to the International Center for Tropical Agriculture (CIAT) in Cali, Colombia, and can be requested from its Biotechnology Research Unit.

SOME APPLICATIONS OF RFLP MARKERS IN COMMON BEAN

Although the current and potential uses of molecular markers have been treated in detail in the first chapter of this book, I would like to point out three targets of interest in common beans: identification of genetic factors that affect gene flow between gene pools, analysis of QTL that affect yield, and tagging genes for resistance to disease for plant breeding purposes and direct genomic cloning. Mesoamerica and the northern Andean region of South America have been identified as the two major centers of diversity (Gepts et al., 1986; Koenig & Gepts, 1989). In addition to the group of land races and modern cultivars that constitute the primary gene pool of common beans, additional genetic variation can be found in: the secondary gene pool that comprises the wild forms of *P. vulgaris*, the tertiary gene that corresponds to the *P. coccineus* complex, and finally the quaternary gene pool that includes *P. acutifolius* and other species (Hidalgo, 1991). All these pools represent a valuable source of genes of economic importance. Unfortunately, gene transfer between these pools can be hampered by intra- and interspecific genetic barriers. For instance, differential photoperiodic responses within the andean gene pool alone can present some difficulties for within-pool gene transfer (Brücher, 1988). Even more dramatic is the dwarf-lethal two gene system (DL_1 and DL_2) that restricts gene flow between some mesoamerican and andean accessions (Gepts & Bliss, 1985; Shii *et al.*, 1980; Singh & Gutierrez, 1984). The existence of some genetic factors that affect compatibility in interspecific crosses has been

49

suggested and it is supported by the identification of certain genotype combinations with increased compatibility (Hucl & Scoles, 1985; Parker & Michaels, 1986). The extent of structural similarities in the chromosomes of different gene pool members can also affect the effectiveness of gene transfer. Only a modest characterization of the cytogenetics of the *Phaseolus* group has been achieved due to the small size of the chromosomes. It is not known for instance whether the chromosomes of closely related species are homosequential. For example, Cheng *et al.* (1981) reported two chromosome inversions on different chromosomes in an interspecific hybrid between *P. vulgaris* and *P. coccineus*; however these findings have been contested by Shii *et al.*, (1982). Identification and molecular tagging of genetic factors that restrict gene flow between the gene pools of beans will permit the design of strategies to facilitate an effective gene transfer. The availability of the RFLP map in beans will also permit to compare the genomes of the secondary, tertiary and quaternary gene pools, a project that is currently underway in my laboratory. Comparisons of this kind have already been performed in the solanaceae (Bonierbale *et al.*, 1988; Tanksley *et al.*, 1988).

Common beans lend themselves as a good model system to study QTLs that affect yield. Wallace & Masaya (1988) have developed a "yield system

analysis" to investigate the genetic components of yield and their interactions with the environment. In addition, Hoogenboon *et al.* (1988) have developed a computer simulation model for common beans -- BEANGRO. The application of molecular markers to recombinant inbred lines (Burr *et al.*, 1989), generated from suitable contrasting genotypes, will be useful in the identification of both genetic factors that affect yield and the responses these factors have to different environments. Molecular markers have been used to detect a number of QTLs in tomato (Paterson *et al.*, 1988; See also QTL chapter in this book). This information can in turn be used to refine computer simulation programs that would take into account specific genetic factors. For instance, a negative correlation between seed size and yield has been reported for beans (Coyne, 1968). A genetic factor that affects seed size has been identified via isozyme linkage analysis (Vallejos & Chase, 1991a). Thus, tagging genes that affect the different components of yield can lead to the construction of specific genetic stocks carrying one or multiple combinations of these genes. These stocks can then be used to ask specific questions about the role certain genes play on different physiological processes such as sink-source relationships or photosynthate partitioning.

Finally, a large number of genes involved in disease resistance are available for

molecular tagging in beans. There are at least 21 genetically characterized monogenic virus resistances (Provvidenti, 1987), and a few other resistances to bacterial and fungal pathogens (Bassett, 1989). Molecular tags can facilitate the efficient pyramiding of appropriate resistances into single breeding lines tailored for specific environments. Furthermore, the relatively small size of the bean genome opens the possibility of molecular cloning of any of these resistances via chromosome walking techniques (Rommens *et al.*, 1989). The average ratio of physical distance to map distance has been estimated at 530 Kb/cM in beans (Vallejos *et al.*, 1992). Development of high density maps around a specific disease resistance gene will expedite the isolation of its genomic clone.

REFERENCES

Arumuganatham K & Earle ED (1991) Nuclear DNA content of some important plant species. Plant Molec. Biol. Rept. 9: 208-218.

Bassett MJ (1989) List of genes-*Phaseolus vulgaris* L. Annu. Rept. Bean Improv. Coop. 32: 1-7.

Bassett MJ (1991) A revised linkage map of common bean. HortScience 26: 834-836.

Bonierbale M, Plaisted R & Tanksley SD (1988) Construction of comparative genetic maps of potato and tomato based on a common set of cloned sequences. Genetics 120: 1095-1103.

Brücher H (1988) The wild ancestor of *Phaseolus vulgaris* in South America. In: Gepts P (ed.) Genetic resources of *Phaseolus* beans (pp. 185-214) Kluwer Academic Publishers, Dordrecht.

Burr B, Burr FA , Thompson KH, Albertson MC & Stuber CW (1988) Gene mapping with recombinant inbreds in maize. Genetics 118: 519-526.

Cheng SS, Bassett MJ & Quesenberry KH (1981) Cytogenetic analysis of interspecific hybrids between common bean and scarlet runner bean. Crop Sci. 21: 75-79.

Chase CD, Ortega VM & Vallejos CE (1991) DNA restriction fragment length

polymorphisms correlate with isozyme diversity in *Phaseolus vulgaris* L. Theor. Appl. Genet. 81: 806-811.

Coyne DP (1968) Correlation, heritability, and selection of yield components in field beans, *Phaseolus vulgaris* L. Proc. Amer. Soc. Hort. Sci. 93: 388-396.

Gepts P & Bliss FA (1985) F_1 hybrid weakness in the common bean: Differential geographic origin suggests two gene pools in cultivated germplasm. J. Hered. 76: 447-450.

Gepts P, Osborn TC, Rashka K & Bliss FA (1986) Phaseolin protein variability in wild forms and land races of the common bean (*Phaseolus vulgaris*): evidence for multiple centers of domestication. Econ. Bot. 40: 451-468.

Hidalgo R (1991) CIAT's world *Phaseolus* collection. In: van Schoonhoven A & Voysest O (eds.) Common beans: research for crop improvement (pp. 163-197) C. A. B. International, Wallingford, UK.

Hoogenboon G, Jones JW, White JW & Boote KJ (1988) BEANGRO V 1.0: Dry bean crop growth simulation model: user's guide. Department of Agricultural Engineering. University of Florida.

Hucl P & Scoles GJ (1985) Interspecific hybridizations in the common bean: a review. HortSci. 20: 352-357.

Koenig R & Gepts P (1989) Allozyme diversity in wild *Phaseolus vulgaris*:

further evidence for two major centers of genetic diversity. Theor. Appl. Gent. 78: 809-817.

Lander ES, Green P, Abrahamson J, Baarlow 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.

Mendel G (1866) Experiments in plant hybridisation. Translated by the Royal Horticultural Society of London in 1938. Harvard University Press, Cambridge, MA.

Moreno J & Chrispeels MJ (1989) A lectin gene encodes the α-amylasinhibitor of the common bean. Proc. Natl. Acad. Sci. USA 86: 7885-7889.

Parker JP & Michaels TE (1986) Simple genetic control of hybrid plant development in interspecific crosses between *Phaseolus vulgaris* L. and *P. acutifolius* A. Gray. Plant Breeding 97: 315-323.

Paterson AH, Lander ES, Hewit JD, Peterson TS, LINCOLN SE & Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335: 721-726.

Provvidenti R (1987) List of genes in *Phaseolus vulgaris* for resistance to viruses. Annu. Rept. Bean Improv. Coop. 30: 1-4.

Rommens JM, Iannuzzi MC, Kerem B-S, Drumm ML, Melmer G, Dean M,

Rozmahel R, Cole JL, Kennedy D, Hidaka N, Zsiga M, Buchwald M, Riordan JR, Tsui L-C & Collins FS (1989) Identification of the cystic fibrosis gene: chromosome walking and jumping. Science 245: 1059-265.

Shaw JK & Norton JB (1918) The inheritance of seed-coat color in garden beans. Massachusetts Agric. Exp. Sta. Bull. 185: 58-104.

Shii CT, Rabakoarihanta A, Mok MC, and Mok DWS (1982) Embryo development in reciprocal crosses of *Phaseolus vulgaris* L. and *P. coccineus* Lam. Theor. Appl. Genet. 62: 59-64.

Sun SM, Slightom JL & Hall TC (1981) Intervening sequence in a plant gene - comparison of the partial sequence of cDNA and genomic DNA of French bean phaseolin. Nature 289: 37-41.

Talbot, DR, AdangMJ, Slighton JL & Hall TC (1984) Size and organization of a multigene family encoding phaseolin, the major seed storage protein of *Phaseolus vulgaris* L. Molec. Gen. Genet. 198: 42-49.

Tanksley SD, Bernatzky R, Lapitan LL, & Prince JP (1988) Conservation of gene repertoire but not gene order in pepper and tomato. Proc. Natl. Acad. Sci. USA 85: 6419-6423.

Thompson WF, Everett M, Polans NO & Jorgensen RA (1983) Phytochrome control of RNA levels in developing pea and mungbean leaves. Planta 158: 487-500.

Vallejos CE & Chase CD (1991a) Linkage between isozyme markers and a

locus affecting seed size in *Phaseolus vulgaris* L. Theor. Appl. Genet. **81:** 413-419.

Vallejos CE & Chase CD (1991b) Extended linkage map for the phaseolin linkage group of *Phaseolus vulgaris* L. Theor. Appl. Genet. **82:** 353-357.

Vallejos CE, Sakiyama NS & Chase CD (1992) A molecular marker-based linkage map of Phaseolus vulgaris L. Genetics 131: (in press).

Wallace DH & Masaya PN (1988) Using yield trial data to analyze the physiological genetics of yield accumulation and the genotype X environment interaction effects on yield. Annu. Rept. Bean Improv. Coop. 31: vii-xxiv.

Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski, S.V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl. Acids Res. 18: 6531-6535.

Zheng J, Nakata M, Uchiyama H, Morikawa H & Tanaka R (1991) Giemsa C-banding pattern in several species of *Phaseolus* L. and *Vigna* Savi, fabacea. Cytologia 56: 459-466.

Table 1. List of loci identified with genomic clones (pBngN), their corresponding linkage group association, and relative position. The insert size of each clone is also included.

Locus	Linkage Group	Relative Position	Insert size(bp)	Locus	Linkage Group	Relative Position	Insert size(bp)
Bng1	*J*	3	1790	*Bng14*	*K*	9	3370
Bng2	*K*	1	1800	*Bng15*	*A*	6	3220
Bng3	*C*	12	1200	*Bng16*	*C*	16	2040
Bng4	*B*	9	1270	*Bng17*	*D*	2	1600
Bng5	*K*	12	1470	*Bng18*	*F*	10	1100
Bng6	*K*	3	1400	*Bng19*	*K*	12	2370
Bng7	*F*	9	2800	*Bng20*	*K*	7	1430
Bng8	*G*	8	800	*Bng21*	*C*	2	1980
Bng9	*G*	8	4510	*Bng22*	*F*	7	1780
Bng10	*K*	13	4190	*Bng23*	*A*	8	620
Bng11	*D*	19	3180	*Bng24*	*K*	7	1070
Bng12	*C*	17	2700	*Bng25*	*J*	7	1100
Bng13	*B*	5	1560	*Bng26*	*G*	13	1960

Locus	Linkage Group	Relative Position	Insert size(bp)	Locus	Linkage Group	Relative Position	Insert size(bp)
Bng27	*G*	6	1260	*Bng43*	*F*	2	1060
Bng28	*A*	4	2110	*Bng44*	*C*	5	970
Bng29	*C*	9	1120	*Bng45*	*D*	1	890
Bng30	*B*	4	1350	*Bng46*	*G*	9	2370
Bng31	*F*	5	1130	*Bng47*	*A*	2	2480
Bng32	*C*	4	950	*Bng48*	*B*	13	1790
Bng33	*C*	5	1100	*Bng49*	*E*	4	1320
Bng34	*D*	11	4330	*Bng50*	*A*	3	850
Bng35	*K*	16	1020	*Bng51*	*K*	10	1150
Bng36	*B*	12	2500	*Bng52*	*K*	4	1140
Bng37	*K*	6	1250	*Bng54*	*F*	11	1170
Bng38	*F*	3	1800	*Bng55*	*B*	5	1650
Bng39	*B*	11	2300	*Bng56*	*K*	6	2480
Bng40	*A*	21	800	*Bng57*	*D*	21	2090
Bng41	*B*	13	2780	*Bng58*	*F*	4	2400
Bng42	*A*	18	1500	*Bng60*	*A*	16	2290

Locus	Linkage Group	Relative Position	Insert size(bp)	Locus	Linkage Group	Relative Position	Insert size(bp)
Bng61	*D*	20	1580	*Bng76*	*J*	10	3120
Bng62a	*K*	3	2310	*Bng77*	*D*	5	2500
Bng62b	*F*	11	2310	*Bng78*	*J*	10	2270
Bng63	*C*	6	1560	*Bng79*	*K*	13	2610
Bng64	*K*	7	1620	*Bng80a*	*A*	1	1300
Bng65a	*B*	7	2080	*Bng80b*	*G*	14	1300
Bng65b	*E*	7	2080	*Bng81*	*A*	3	1090
Bng67	*J*	10	2170	*Bng82*	*D*	10	1480
Bng68	*I*	3	2530	*Bng83*	*B*	1	870
Bng69	*F*	11	1000	*Bng84*	*D*	6	2190
Bng70	*J*	1	1970	*Bng86*	*C*	15	1650
Bng71	*B*	7	2240	*Bng87*	*G*	3	3350
Bng72	*B*	14	3880	*Bng88*	*G*	9	2510
Bng73	*F*	1	2410	*Bng89*	*D*	12	840
Bng74	*D*	19	650	*Bng90*	*D*	13	1850
Bng75	*C*	15	1600	*Bng91*	*J*	8	1470

Locus	Linkage Group	Relative Position	Insert size(bp)	Locus	Linkage Group	Relative Position	Insert size(bp)
Bng92	*J*	10	1130	*Bng107b*	*K*	5	2560
Bng93	*K*	13	800	*Bng108*	*D*	8	1540
Bng94	*G*	4	1460	*Bng109*	*K*	13	1930
Bng95	*G*	10	1580	*Bng110*	*K*	7	1270
Bng96	*F*	6	1140	*Bng111*	*K*	1	1970
Bng97	*J*	8	1380	*Bng112*	*J*	10	820
Bng98	*D*	17	1750	*Bng113*	*B*	4	660
Bng99	*K*	12	1260	*Bng114*	*C*	14	1540
Bng100	*I*	5	2160	*Bng115*	*D*	10	1930
Bng101	*J*	6	670	*Bng116*	*C*	16	1930
Bng102	*K*	11	1400	*Bng117*	*D*	16	1430
Bng103	*B*	3	1410	*Bng118*	*A*	6	1760
Bng104	*G*	1	1830	*Bng119*	*D*	18	2220
Bng105	*D*	15	1260	*Bng120*	*J*	10	1520
Bng106	*C*	17	2680	*Bng121*	*I*	6	1710
Bng107a	*A*	3	2560	*Bng122*	*B*	4	1330

Locus	Linkage Group	Relative Position	Insert size(bp)	Locus	Linkage Group	Relative Position	Insert size(bp)
Bng123	*C*	13	1180	*Bng140*	*K*	6	1780
Bng124	*C*	3	1270	*Bng141*	*D*	4	1340
Bng125	*F*	11	1140	*Bng142a*	*C*	19	3470
Bng126	*B*	5	2510	*Bng142b*	*B*	13	3470
Bng128	*F*	11	2340	*Bng143*	*K*	8	950
Bng129	*K*	13	1710	*Bng144*	*D*	6	2220
Bng130	*B*	16	950	*Bng145*	*J*	5	1930
Bng131	*F*	2	1190	*Bng146*	*A*	14	1550
Bng132	*K*	13	1120	*Bng148*	*D*	14	2220
Bng133	*E*	5	3120	*Bng149*	*K*	13	820
Bng134	*K*	6	3830	*Bng150*	*K*	2	1940
Bng135	*A*	16	3000	*Bng151*	*B*	8	1030
Bng136	*J*	2	1190	*Bng152*	*E*	6	1070
Bng137	*G*	12	1710	*Bng153*	*A*	15	1900
Bng138	*F*	9	1320	*Bng154*	*J*	10	1050
Bng139	*F*	12	2540	*Bng155*	*C*	8	2810

Locus	Linkage Group	Relative Position	Insert size(bp)	Locus	Linkage Group	Relative Position	Insert size(bp)
Bng156	*K*	4	2500	*Bng172*	*I*	7	3120
Bng157	*A*	7	1310	*Bng173*	*B*	2	2050
Bng158	*J*	10	2310	*Bng174*	*D*	12	2600
Bng159	*D*	3	3210	*Bng175*	*C*	13	880
Bng160	*B*	8	2280	*Bng176*	*K*	12	3120
Bng161	*E*	9	1290	*Bng177*	*G*	7	930
Bng162	*E*	2	590	*Bng178*	*D*	13	1310
Bng163	*K*	13	2180	*Bng179*	*G*	2	2240
Bng164	*C*	11	1480	*Bng180*	*D*	17	2550
Bng165	*C*	9	2260	*Bng181*	*K*	15	1770
Bng166	*E*	7	2170	*Bng182*	*J*	9	1180
Bng167	*K*	6	520	*Bng183*	*G*	4	1210
Bng168a	*A*	22	2210	*Bng184*	*B*	4	530
Bng168b	*K*	4	2210	*Bng186*	*F*	11	1350
Bng170	*A*	20	1170	*Bng187*	*J*	4	1860
Bng171	*B*	3	1900	*Bng188*	*E*	10	1960

67

Locus	Linkage Group	Relative Position	Insert size(bp)	Locus	Linkage Group	Relative Position	Insert size(bp)
Bng189	*B*	4	2030	*Bng206*	*J*	10	1290
Bng190	*D*	9	1330	*Bng209*	*G*	10	1020
Bng191	*A*	22	2030	*Bng211a*	*A*	21	1810
Bng192	*J*	5	1300	*Bng211b*	*C*	1	1810
Bng193	*G*	5	2620	*Bng212*	*D*	16	1380
Bng195	*B*	15	2700	*Bng213*	*F*	4	2180
Bng197	*K*	14	2560	*Bng214*	*F*	11	2330
Bng198	*K*	12	2140	*Bng215*	*G*	5	1350
Bng199	*A*	19	1230	*Bng216*	*C*	6	2040
Bng200	*I*	1	2340	*Bng218*	*I*	8	2020
Bng201	*D*	4	1320	*Bng219*	*I*	4	580
Bng202	*G*	1	1620	*Bng220*	*F*	11	4250
Bng203	*A*	10	2990	*Bng221*	*C*	10	2100
Bng204	*A*	12	1420	*Bng222*	*A*	12	2480
Bng205a	*F*	8	1350	*Bng223*	*A*	13	2350
Bng205b	*E*	7	1350	*Bng224*	*B*	1	880

64

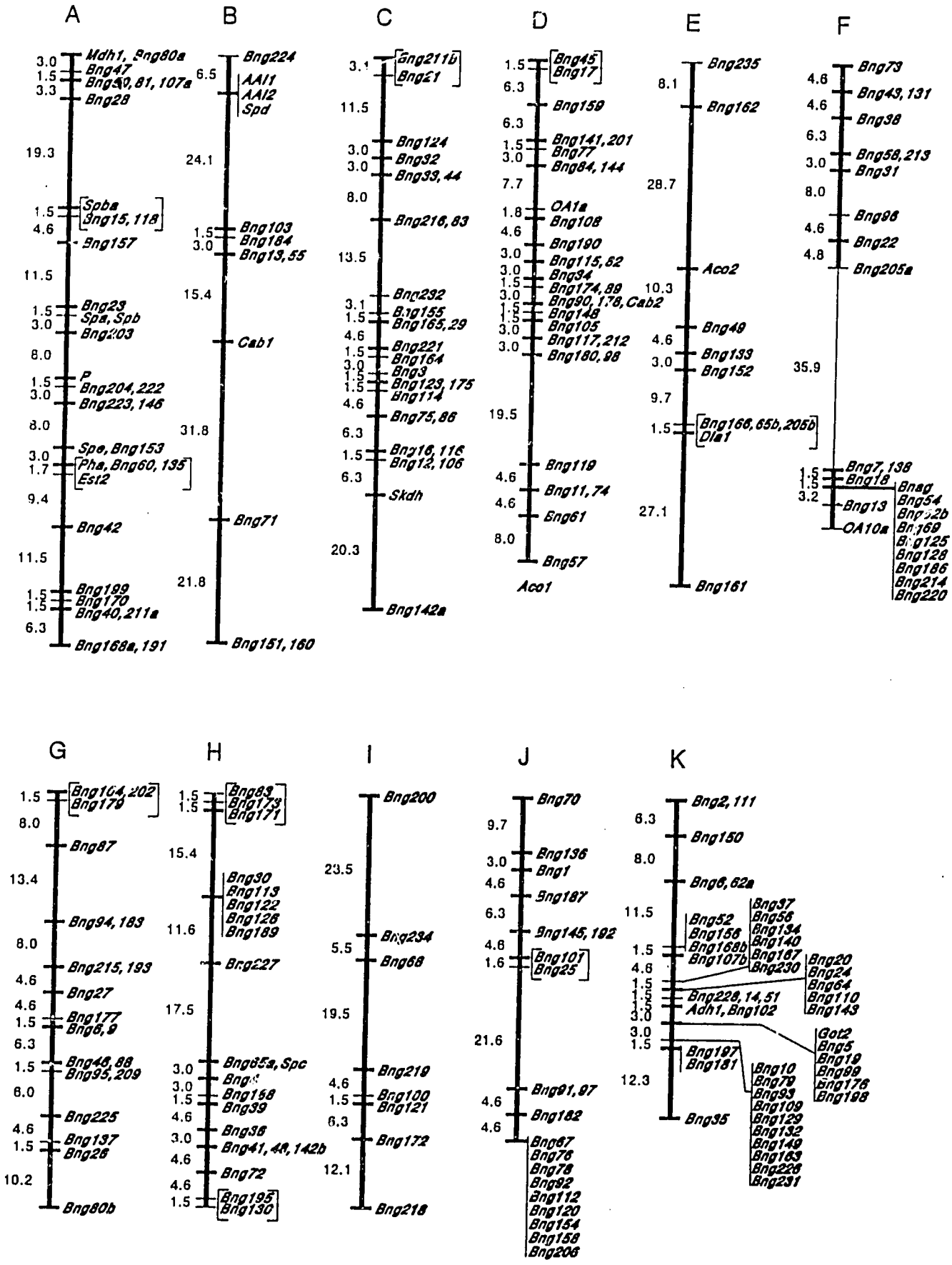
Locus	Linkage Group	Relative Position	Insert size(bp)	Locus	Linkage Group	Relative Position	Insert size(bp)
Bng225	*G*	11	1030	*Bng231*	*K*	13	2100
Bng226	*K*	13	2220	*Bng232*	*C*	7	1770
Bng227	*B*	6	700	*Bng234*	*I*	2	900
Bng228	*K*	9	1840	*Bng235*	*E*	1	2030
Bng230	*K*	6	500				

Table 2. List of Protein Marker Loci: Isozymes and Seed Storage Proteins.

Isozyme Locus	Linkage Group	Relative Position	References[1]	Seed Protein Locus	Link. Grp.	Relat. Posit.	References
Aco1	*D*	-	1, 3	*AAI1*	*B*	2	3
Aco2	*E*	3	1, 3	*AAI2*	*B*	2	3
Adh1	*K*	9	1, 3	*Pha*	*A*	15	2, 3
Bnag	*F*	11	1,3	*Spa*	*A*	9	2, 3
Dia1	*E*	8	1, 3	*Spb*	*A*	9	2, 3
Est2	*A*	16	1, 3	*Spba*	*A*	5	2, 3
Got2	*K*	10	1, 3	*Spc*	*H*	6	2, 3
Mdh1	*A*	1	1, 3	*Spd*	*B*	2	2, 3
Skdh	*C*	18	1, 3	*Spe*	*A*	14	2, 3

[1]1, Vallejos & Chase, 1991a; 2, Vallejos & Chase, 1991b; 3, Vallejos *et al.*, 1992.

Figure 1. *Phaseolus vulgaris* L. linkage map. **Mapmaker** (Lander *et al.* 1987) Was uused to group and order 250 markers. These marekers have been assigned to 150 loci/locus clusters on 11 linkage groups (*A-K*). The exact position of *Aco1* on *group D* could not be located due to incomplete data for this locus; *Aco1* is listed at the bottom of the group. A **framework** order was obtained after grouping the markers (LOD 4.0, 25 cM), and then ordering each group (LOD 3.0, 25 cM, three point exclusion LOD -3.0). Markers that belong to the **framework** order have been marked with thicker crossbars. Loci for which an order could not be established (LOD < 2.0) have been enclosed by brackets. Numbers on the left of a linkage group represent map distances in cM. Loci at distances that exceed 25 cM in groups *B*, and *E*, have an LOD score > 4.0 and have been included in the **framework**. The distance (35.9 cM) between *Bng205a* and *Bng7* (*group F*) was calculated using the Kosambi function. Although an LOD of 2.22 was found for this linkage, other independent data support their assignment to this group.



Phaseolus vulgaris Linkage Map