

# Molecular Analysis and Cloning of *Malus* Ribosomal DNA

C.J. Simon<sup>1</sup> and N.F. Weeden<sup>2</sup>

Department of Horticultural Sciences, New York State Agricultural Experiment Station, Cornell University, Geneva, NY 14456

**Additional index words.** crab apple, genetic diversity, gene structure

**Abstract.** The ribosomal genes of the two crab apple (*Malus*) genotypes 'White Angel' and 'Robusta 5' were characterized to determine the extent of between- and within-genotype heterogeneity. Initial investigations with a cloned sequence of soybean rDNA failed to detect some *Malus* intergenic spacer region fragments. An alternative probing method that used electrophoretically purified *Malus* rDNA was developed. Double-digests of total genomic DNA with combinations of 13 restriction endonucleases identified the positions of 35 restriction sites. Restriction site polymorphism was observed both between and within the crab apple genotypes. Ribosomal DNA from 'White Angel' was cloned in phage and plasmid vectors and mapped with 11 enzymes. The region of the spacer causing length heterogeneity was identified. These clones should be useful as genetic markers and for examining population dynamics and systematic of *Malus* and closely related taxa.

The genes encoding the 45S ribosomal RNA (rDNA) are particularly useful as genetic markers because of their unique structure and conserved nature. The genes exist as one or more arrays of tandemly repeated sequences with on the order of  $\geq 1000$  copies per haploid plant genome (Long and Dawid, 1980). Within an array, the units are usually identical. Digestion of this array with an endonuclease that recognizes one site per repeat unit will generate several hundred identical segments the size of the repeat unit, because the "head-to-tail" profile of the tandem repeats resembles a "circular map" (Appels and Honeycutt, 1986). This feature of the rDNA enables researchers to perform restriction mapping of the genes using total genomic DNA preparations, which allows examination of the rDNA profile of all sequences in the genome. This possibility contrasts to single or low copy number, or nontandem high copy number sequences, which may only be restriction mapped in their entirety as cloned sequences.

A second feature makes the array useful for restriction fragment length polymorphism (RFLP) analysis. The repeat unit consists of highly conserved segments that code for the ribosomal RNA sequences and highly variable "spacer" regions (Appels and Dvorak, 1982). The high number of copies and the highly conserved coding sequences permit heterologous probes to be used across wide taxonomic distances, while the polymorphic spacer regions often show variation among individuals within an interbreeding population (Rogers and Bendich, 1987). The usual source of variation of repeat unit size is found in the spacer region separating the coding regions of the repeat units. In such cases, a small tandem subrepeat located in the spacer is found in differing numbers. Well-characterized subrepeats range in size from 115 base pairs (bp) in barley (Saghai-Maroo

et al., 1984) to 334 bp in broad bean (Yakura et al., 1984). Thus, RFLP analysis, whether it is being used for population analysis, systematic studies, genome characterization, or mapping, is often initiated with an examination of the rDNA genes of the organism (Appels and Dvorak, 1982; Doyle and Beachy, 1985; Polaris et al., 1986; Saghai-Maroo et al., 1984; Systma and Schaaf, 1985).

Despite their utility and importance, ribosomal genes have not been examined in many horticultural crops. As part of our effort to apply molecular techniques to apple breeding and the characterization of *Malus* germplasm, we investigated the 45S ribosomal repeat unit. Initially we used heterologous clones from pea (*Pisum sativum* L.) and soybean (*Glycine max* (L.) Merr.) in Southern hybridization analysis. However, as our studies progressed, it became evident that a homologous rDNA probe was desirable. We report herein the isolation of several ribosomal DNA clones from *Malus* and the restriction mapping of 45 S-rDNA repeat units in the crab apple genotypes 'White Angel' and 'Robusta 5'.

## Materials and Methods

**Plant material.** The crab apple genotypes 'White Angel' and 'Robusta 5' were obtained at the New York State Agricultural Experiment Station in Geneva, N.Y. 'Robusta 5' is a selection of *Malus*  $\times$  *robusta* that is used in rootstock breeding. 'White Angel' is of uncertain taxonomic status but appears to be derived from *M. sargentii* or *M. seiboldii* (Simon and Weeden, 1991). Leaf material was taken in early August from mature individual trees. The leaves were soaked in 10% bleach (0.525% Na-hypochlorite) solution for 10 min and then rinsed well with water. The midveins were removed, and the leaf lamina were aliquoted into 20-g lots and stored immersed in liquid nitrogen.

**DNA Isolation.** DNA was isolated from the frozen samples by a modification of the procedure of Murray and Thompson (1980). Single 20-g samples of frozen leaf blades were first powdered in a stainless steel blender jar in the presence of liquid nitrogen. The powder was transferred to a glass blender jar containing 150 ml of ice cold nuclei extraction buffer (pH 7.5) that contained 2.85 M sorbitol, 0.1 M Tris, 5 mM EDTA, and 20 mM sodium bisulfite. The mixture was blended briefly to homogeneity and filtered through two layers of cheesecloth and one of Miracloth into a centrifuge bottle on ice. The bottle was spun for 15 min at  $500 \times g$  in a Sorvall GSA (DuPont, Wilmington, Del.) rotor at 4C. The pellet was gently resuspended

Received for publication 20 Dec. 1990. Accepted for publication 7 Sept. 1991. We thank J.J. Doyle and R.D. Way for helpful suggestions, support, and guidance during this research and J.C. Sanford and B.I. Reisch for critical reading and discussion of earlier versions of this manuscript. This work was supported in part by a grant from the Cornell Biotechnology Program, which is sponsored by the New York State Science and Technology Foundation, a consortium of industries, the U.S. Army Research Office, and the National Science Foundation. The cost of publishing this paper was defrayed in part by the payment of page charges. Under postal regulations, this paper therefore must be hereby marked advertisement solely to indicate this fact.

<sup>1</sup>Present address: USDA, ARS, 215 Johnson Hall, Washington State Univ., Pullman, WA 99164.

<sup>2</sup>To whom correspondence should be addressed.

into 5 ml of fresh extraction buffer and transferred into a 50-ml centrifuge tube. To this, 5 ml of nuclei lysis buffer [0.2 M Tris, 50 mM EDTA, 2 M NaCl, 2% hexadecyltrimethylammonium bromide (CTAB)] was added, followed by 2 ml of 5% sarkosyl, and the solution was shaken to mix well. The mixture was incubated at 65°C for 20 min and then extracted with 15 ml of chloroform/isoamyl alcohol (24: 1). Leaves collected late in the season produced a large amount of cloudiness, presumably carbohydrate, in the aqueous phase. The cloudiness was removed by up to six chloroform extractions with centrifugation at  $10,000 \times g$ . The DNA in the cleared aqueous layer was precipitated with two volumes of cold ethanol. Precipitate was spooled out on a glass rod and dissolved for 1 h at 65°C in 3.5 ml TE buffer (10 mM Tris, pH 7.5, 1 mM EDTA). Insoluble were spun down and the supernatant was collected and stored at 4°C.

**Molecular mapping.** Aliquots (5 mg) of each of the DNAs were subjected to single and double digestions with numerous restriction enzymes. Enzymes whose sites mapped to the *Malus* rDNA included *Apa* I, *Bam* HI, *Bgl* II, *Dra* I, *Eco* RI, *Eco* RV, *Hind* III, *Kpn* I, *Sac* I, *Sph* I, *Xba* I, *Xho* I, and *Xmn* I. Digestions were performed overnight using 10 units of each enzyme in a 25- $\mu$ l reaction volume at the conditions recommended by the supplier. RNase A (100 mg/ml-1) was included in the reaction to remove RNA. Double digests were performed in two steps with some of the enzyme combinations to optimize the salt concentration of the reactions. The enzymes *Eco* RI and *Xba* I were chosen as standard reference members in the double digestion scheme because they cleave sites that are well distributed across the apple ribosomal repeat unit, and because they were first used to orient the coding and spacer regions of the repeat units.

Restricted DNAs were separated on 0.7% agarose slab gels run in  $1 \times$  TBE buffer (0.089 M Tris, 20 mM EDTA, 0.089 M boric acid, pH 7.6) at low voltage (1 V·cm<sup>-1</sup>). DNA in the gels was transferred to nylon membranes (GeneScreen Plus, NEN; Du Pont) by the blotting technique described by Southern (1975).

Three DNAs were used to probe the filters. A plasmid containing the entire repeat unit of the pea ribosomal gene (Polaris et al., 1986) was nick-translated with DNA polymerase I/DNase I according to the manufacturer's instructions (Bethesda Research Labs, Bethesda, Md.) and used for the initial mapping. The membranes were then stripped, and a cloned soybean rDNA repeat unit (Doyle and Beachy, 1985) obtained from J. Doyle was hybridized with the same filters to determine if the spacer could be better resolved. A homologous probe for the spacer region was obtained by gel isolating relatively pure spacer sequence following triple digestion of 'White Angel' DNA with *Eco* RI, *Eco* RV, and *Bcl* I. After running  $\approx 250 \mu$ g of triply digested DNA on a 0.7% agarose preparative gel at low voltage (1 V·cm<sup>-1</sup>) and staining it with ethidium bromide, a 6-kb rDNA fragment, which had a higher molecular weight than other DNA in the gel, was isolated and extracted from the gel slice by the powdered glass method (Vogelstein and Gillespie, 1979) using a kit obtained from Bio 101 (La Jolla, Calif.). The 6-kb fragment was nick-translated and hybridized to the same filters previously tested with the legume probes.

Initial hybridizations were performed in the presence of 40% formamide at 42°C (Maniatis et al., 1982). Later, better results were obtained using a buffer system lacking formamide, but with incubation at 65°C. The buffer used in this procedure consisted of 0.75 M NaCl, 125 mM citric acid trisodium dihydrate, 0.6% SDS, 50 mM Na<sub>2</sub>HPO<sub>4</sub> (adjusted to pH 7.2-7.4 with 1 N HCl),  $5 \times$  Denhardt's solution, 2.5 mM EDTA, 5% dextran sul-

fate, and 100  $\mu$ g denatured calf thymus DNA/ml. Washes of the membranes took place at 65°C in 2  $\mu$  SSC, 0.1% SDS. After one 30-min wash and two 1-h washes, the membranes were wrapped in a thin plastic sheet and placed against X-ray film at -70°C with a DuPont Cronex intensifying screen.

**Molecular cloning of White Angel' rDNA.** About 400  $\mu$ g of 'White Angel' DNA was cut to completion with 200 units each of *Eco* RI and *Bcl* I. The restriction product was electrophoresed in a 0.7% agarose slab gel at a low voltage (1 V·cm<sup>-1</sup>). The gel was stained with ethidium bromide and destained in distilled water in darkness for several hours to improve the definition of the individual bands, and also to dilute the TBE, particularly the boric acid, in the gel. The gel slab was then placed on plastic wrap and briefly viewed with ultraviolet light (302 nm) to be photographed. The resulting photograph was examined closely to locate the bands corresponding to the ribosomal gene fragments, whose sizes were known to be 4 kb and 6 kb, based on previous mapping studies. Gel regions containing the rDNA sequence bands were excised with a new razor blade. The DNA in the excised gel regions were then extracted from the gel by the powdered glass method (Vogelstein and Gillespie, 1979).

Purified insert DNA of the 4-kb and 6-kb fragment size classes were mixed at two concentrations each with 0.5  $\mu$ g of commercially obtained dephosphorylated  $\lambda$ gt10 bacteriophage arms (Promega, Madison, Wis.). Relative DNA concentrations were  $\approx 1:1$  and  $4:1$  ( $\lambda$ gt10 : insert) by weight. The ligation reaction consisted of the DNAs, two units T<sub>4</sub> DNA ligase (Stratagene, La Jolla, Calif.), 50 mM Tris-HCl pH 8.0, 7 mM MgCl<sub>2</sub>, 1 mM DTT, and 0.5 mM ATP in a volume of 10  $\mu$ l overnight at 4°C. Ligated DNA was assembled into phage with commercially obtained packaging extracts (Packagene, Promega). Dilutions of the reaction were used to infect freshly cultured strain C600hfl *E. coli*, and were plated in a 0.7% LB top agar to obtain approximate titers of the reactions (Maniatis et al., 1982).

DNA from the plaques was lifted as suggested by the manufacturer of the membranes (colony/plaque screen, NEN). The probe for screening the filters was prepared by nick-translating gel purified insert DNA from the plasmid containing the soybean ribosomal gene. The filters were hybridized and washed as described above. Plaques corresponding to a positive signal on the autoradiograms were picked with a pasteur pipet, as described by Maniatis et al. (1982), and were re-plated for confirmation. High density plates of the confirmed positives were used for DNA mini-preps (Maniatis et al., 1982), for final verification by Southern analysis. Isolates containing appropriately sized inserts that exhibited homology to the soybean rDNA probe were selected for subcloning into pUC13 plasmid.

Subsequent analysis of the *Malus* rDNA focused on the 6-kb insert containing the intergenic spacer region. For ease of handling and analysis, the 6-kb *Eco* RI inserts were removed from phage containing positive clones and subcloned into the plasmid pUC13. About 2  $\mu$ g of the mini-prep DNA isolated from the phage was restricted with *Eco* RI and run on a 0.7% preparative agarose gel. The insert was excised from the gel and was extracted using the ground glass method as described above. This DNA was ligated into 100 ng of commercially obtained dephosphorylated pUC13 vector (Pharmacia, Piscataway, N.J.) overnight in 20  $\mu$ l under the conditions described above. One microliter of  $5 \times$  diluted ligation product was transformed into competent DH5a bacterial cells (BRL) and plated on LB plates containing 75  $\mu$ g·ml<sup>-1</sup> ampicillin, 20  $\mu$ g·ml<sup>-1</sup> X-gal, and 0.1 mM isopropyl (3-D-thiogalactopyranoside (IPTG). White colonies were picked the next morning and grown in 5 ml of liquid



RI fragment. *Bcl* I has no cut sites in 'White Angel' rDNA, and also has the property of digesting apple DNA very well. By performing the triple digestion, the 6-kb *Eco* RI ribosomal sequence was unusually large, relative to the rest of the 'White Angel' DNA. Upon electrophoresis of the digestion product, the 6-kb fragment was separated from almost all the rest of the apple DNA. As noted, these probes resolved *Malus* spacer sequence well, as later confirmed by the use of cloned *Malus* sequence.

An additional feature of the gel-isolated probes that facilitated fragment ordering is that the probe was comprised of only the 6-kb *Eco* RI region of the gene, which primarily contains spacer and 18S sequence. This probe did not detect fragments specific to the 3.9-kb *Eco* RI fragment (encoding the 5.8S and 25S rRNAs). Therefore, in hybridizations involving the 6-kb probe, lack of signal confirmed the location for fragments mapped to the 3.9-kb region with the legume probes. A final tool used in interpreting the fragment order was the consideration of signal intensity using the two probing systems. Bands showing strong hybridization to the legume probes always mapped to the coding regions, whereas spacer-derived fragments generally were faint on autoradiograms. The gel isolated probe clarified the placement of spacer-derived fragments, because it signaled spacer-derived and coding-derived regions with equal intensity. This approach can be useful for similar analyses of the 45S rDNA repeat in other taxa.

The compiled map (Fig. 2) of the three rDNA sequences cloned from 'White Angel' (clones pAR72, pAR73, and pAR77) corresponded well with the results of the genomic DNA mapping except for the *Barn* HI site found in the intergenic spacer. This site was missing on all three clones analyzed. The three clones differed slightly, however, enabling a closer examination of the nature of the polymorphisms. The clones pAR72 and pAR73 differed in their *Xho* I and *Sph* I maps. In contrast, pAR77 was smaller than the other two, being  $\approx 200$  base pairs shorter. The nature of the size polymorphism is partially elucidated by the restriction enzyme *Sph* I. Genomic DNA mapping revealed that the sum of the *Sph* I fragments seen did not equal the length of the entire sequence. This, coupled with the evidence of length polymorphism often seen in the region of the *Sph* I site just 3' to the 25S coding region, led us to suspect that the postulated subrepeat responsible for the length polymorphisms may contain an *Sph* I site. Small fragments of DNA generated by *Sph* I are visible on 3% gels (Fig. 3). Two of the smallest *Sph* I fragments have a combined size of  $\approx 200$  bp, indicating that they may represent the subrepeat, and that the subrepeat contains two *Sph* I sites. Sequence analysis in the region of these sites will be necessary to resolve precisely how the subrepeats are organized.

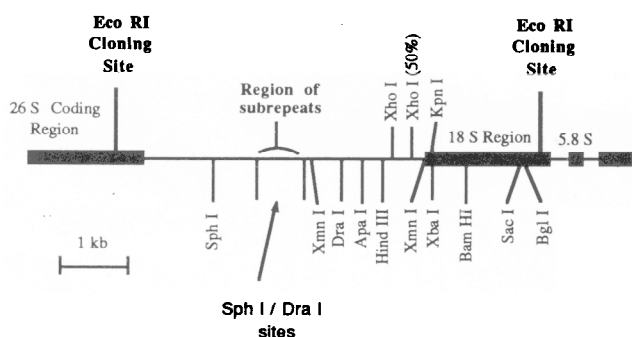


Fig. 2. Restriction map of generalized 'White Angel' cloned DNA, indicating region of subrepeats responsible for length polymorphism.

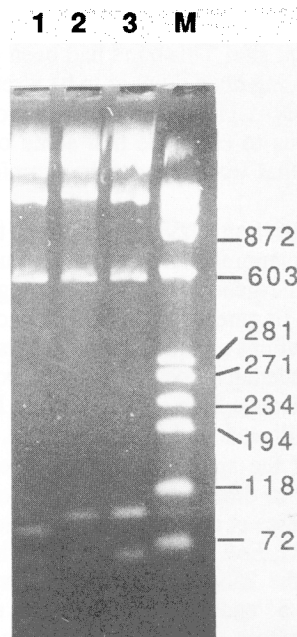


Fig. 3. *Sph* I digests of clones pAR72, pAR73 and pAR77, respectively, run on a 3% agarose gel. The molecular weight standard is Hae III  $\phi$  X174. The sum of the sizes of the two smallest fragments is  $\approx 190$  bp, similar to the size of the postulated subrepeat.

*Dra* I restriction site results provided further evidence for heterogeneity in the subrepeat region. Although this enzyme mapped simply with genomic DNA, restriction maps of the clones revealed an additional region of the digestion and also an apparent loss of DNA due to very frequent cleavage sites (data not shown). This situation resembles the genomic DNA results for *Sph* I. Analysis of the cloned sequences revealed a 190-bp fragment generated upon digestion with *Dra* I, suggesting that each subrepeat sequence possessed a single *Dra* I site. The clone pAR72 again exhibited a slightly different sized subrepeat fragment, as it did with *Sph* I digestion. This situation provides further evidence that the *Dra* I and *Sph* I patterns are related.

Two inconsistencies were observed with the restriction enzyme *Barn* HI. Within the 25S coding region there was a *Barn* HI site that appeared to be polymorphic within both crab apple genotypes. However, Flavell (1985) summarized reports indicating that rDNAs are more extensively methylated than average total DNA. Most restriction enzymes are sensitive to DNA methylations of one type or another. The *Barn* HI site in the middle of the 25S region is present in all copies of cloned rDNA, where methylation is not a factor, but is resistant to digestion in genomic DNA (Siegel and Kolacz, 1983). For the *Malus* 45S rDNA map, an asterisk rather than parentheses is used to denote this "polymorphism," because it quite likely that the site is present but is physically resistant to cleavage. It is important to recognize that estimates of genetic diversity based on RFLP data may be inflated by such pseudo-genetic polymorphism generated by differences in methylation patterns among repeat units.

The second inconsistency between the genomic DNA map and the map of the rDNA clones was the *Barn* HI site in the intergenic spacer. The *Barn* HI site was present in the genomic DNA map but missing in all three clones. Repetition of the initial genomic DNA mapping analysis confirmed that the *Barn* HI site existed; however, a faint band was observed on the

genomic DNA autoradiograms at a position expected for sequences lacking that site. This band had been disregarded in the initial analysis, having been thought to be a result of incomplete digestion of the DNA. However, the data from the cloned sequences have led us to conclude that some of the repeat units lack this site and that we happened to examine only clones of such sequences.

For all other restriction enzymes analyzed, the map generated by genomic DNA mapping matched that produced for the cloned segments. Thus, the genomic DNA mapping approach, using homologous DNA fragments to probe for spacer sequences, appears to be an excellent method for the analysis of rDNA polymorphisms. This technique does not require expertise in cloning, only the ability to isolate high quality nuclear DNA and perform single and double digests and Southern blots.

The structure of the apple 45S rDNA repeat is very similar to that determined in other angiosperms (Choumane and Heizmann, 1988; Cordesse et al., 1990; Rogers and Bendich, 1987). Not only are *Eco* RI, *Eco* RV, and *Xba* I restriction sites present at the 3' ends of the 25S and 18S sequences, within the 5.8S region, and at the 5' end of the 18S coding sequence, respectively, but the structure of the spacer also is similar. The position of the subrepeats near the 3' end of the 25S coding sequence and the apparent size of the subrepeat (190 bp) are typical.

Our primary interest in examining the ribosomal genes of *Malus* was to assess the utility of this gene family for future studies of *Malus* genome organization and analysis of genetic diversity. The present study demonstrates that the prerequisite genetic variation does exist among crab apple genotypes and that the 45S rDNA gene family can be used for such applications. The map developed in this study depicts 35 mapped restriction sites in the *Malus* rDNA, representing  $\approx 2\%$  of the entire ribosomal sequence. This information can be used to interpret restriction patterns of other *Malus* genotypes. An understanding of the structure of the ribosomal DNA can assist in the interpretations of restriction patterns obtained in population or taxonomic studies or for linkage mapping analyses.

#### Literature Cited

- Appels, R. and J. Dvorak. 1982. The wheat ribosomal DNA spacer region: Its structure and variation in populations and among species. *Theor. Applied Genet.* 63:337-348.
- Appels, R. and R. Honeycutt. 1986. rDNA Evolution over a billion years, p. 81-135. In: S.K. Dutta (ed.). *DNA systematic*, vol. II: plants. CRC, Boca Raton, Fla.
- Choumane, W. and P. Heizmann. 1988. Structure and variability of nuclear ribosomal genes in the genus *Helianthus*. *Theor. Applied Genet.* 79:481-489.
- Cordesse, F., G. Second, and M. Delseny. 1990. Ribosomal gene spacer length variability in cultivated and wild rice species. *Theor. Applied Genet.* 79:81-88.
- Doyle, J.J. and R.N. Beachy. 1985. Ribosomal gene variation in soybean (*Glycine*) and its relatives. *Theor. Applied Genet.* 70:369-376.
- Eckenrode, V. K., J. Arnold, and R.B. Meagher. 1985. Comparison of the nucleotide sequence of soybean 18 S rDNA with the sequence of other small-subunit rRNAs. *J. Mol. Evolution* 21:259-269.
- Flavell, R.B. 1985. Repeated sequences and genome change, p. 139-156. In: B. Miflin (ed.). *Oxford surveys of plant molecular and cell biology*. Oxford Univ. Press, Oxford, England.
- Long, E.O. and I.B. David. 1980. Repeated genes in eukaryotes. *Arm. Rev. Biochem.* 49:727-764.
- Maniatis, T., E. Fritsch, and J. Sambrook. 1982. *Molecular cloning*. A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Murray, M.G. and W.F. Thompson. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8:4321-4325.
- Polaris, N. O., N.F. Weeden, and W.F. Thompson. 1986. Distribution, inheritance and linkage relationship of ribosomal DNA spacer length variants in pea. *Theor. Applied Genet.* 72:289-295.
- Rogers, S.O. and A.J. Bendich. 1987. Ribosomal RNA in plants: variability in copy number and in the intergenic spacer. *Plant Mol. Biol.* 9:509-520.
- Saghai-Maroo, M.A., K.M. Soliman, R.A. Jorgenson, and R.W. Allard. 1984. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proc. Natl. Acad. Sci. USA* 81:8014-8018.
- Siegel, A. and K. Kolacz. 1983. Heterogeneity of pumpkin ribosomal DNA. *Plant Physiol.* 72:166-171.
- Simon, C.J. and N.F. Weeden. 1991. Elucidation of crab apple lineage by direct examination of rDNA sequences. *Malus* (In press.)
- Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.
- Systma, K.J. and B.A. Schaal. 1985. Phylogenetics of the *Lisianthus skinneri* (*Gentianaceae*) species complex in Panama utilizing DNA restriction fragment analysis. *Evolution* 39:594-608.
- Vogelstein, B. and D. Gillespie. 1979. Preparative and analytical purification of DNA from agarose. *Proc. Natl. Acad. Sci., USA* 76, 615.
- Wilimzig, M. 1985. LiCl-boiling method for plasmid mini-preps. *Trends in Genet.* 1:158.
- Yakura, K., A. Kate, and S. Tanifuji. 1984. Length heterogeneity in the large spacer of *Vicia faba* rDNA is due to the differing number of 325 bp repetitive sequence elements. *Mol. Gen. Genet.* 193:400-405.