This is to certify that the

dissertation entitled

Mitochondrial genome diversity in Phaseolus with special reference to P. vulgaris L.

presented by

Mireille Khairallah

has been accepted towards fulfillment of the requirements for

Ph. D. degree in Plant Breeding and Genetics (Crop and Soil Science dpt)

M. W. Adams
Major professor

Date Sept 25, 1990

MITOCHONDRIAL GENOME DIVERSITY IN PHASEOLUS, WITH SPECIAL REFERENCE TO P. VULGARIS L.

By

Mireille Michel Khairallah

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Plant Breeding and Genetics Program
Department of Crop and Soil Sciences

1990

ABSTRACT

MITOCHONDRIAL GENOME DIVERSITY IN PHASEOLUS, WITH SPECIAL REFERENCE TO P. VULGARIS L.

By

Mireille Michel Khairallah

Intraspecific mitochondrial DNA (mtDNA) diversity was determined in 23 Phaseolus vulgaris genotypes. Twenty of the lines were collected from Malawian landraces; the other three were pure-bred cultivars. The mtDNAs were digested with eight restriction endonucleases, revealing complex banding patterns. Southern hybridization showed a considerable amount of uniformity of the mtDNA banding patterns. However, five restriction fragment length polymorphisms (RFLPs) divided the bean lines into two groups corresponding to the previously known small-seeded Mesoamerican and large-seeded Andean gene pools of P. vulgaris.

Mitochondrial DNA was then used to compare the amount of diversity in wild beans to that in the cultivated, and to understand how and when the mitochondrial genomes of the gene pools became distinct. The mtDNA of six wild bean accessions from Central and South America were digested with nine endonucleases and analysed by Southern hybridization. Twenty RFLPs were detected, demonstrating a significantly higher amount of mtDNA variability in wild than in cultivated beans. Results indicated that in four out of the five inter-gene pool RFLPs, the polymorphisms arose soon after domestication, two in the Mesoamerican gene pool and two in the Andean one. The fifth RFLP must have occurred before domestication since the locus was polymorphic in the wild beans. The distribution of mtDNA RFLPs among wild beans strongly supports the concept of two different domestication events for P. vulgaris.

Restriction patterns of mtDNA were also examined from three Phaseolus species to estimate their genome sizes and to determine the level of interspecific variability and relatedness. Three endonucleases were used to determine the genome size of the species. The estimates were 456, 324, and 400kb, respectively, for P.vulgaris, P. coccineus, and P. acutifolius. Many RFLPs differentiated the species when the mtDNAs were digested with seven endonucleases and analyzed by Southern hybridization. Proportions of shared restriction fragments between species pairs were computed, and demonstrated that P. vulgaris and P. coccineus are more related to each other than either is to P. acutifolius. The latter had a similar degree of relationship to the other two species.

To my parents, Hoda and Michel,

for all their support, love, and confidence.

ACKNOWLEDGMENTS

I would like to acknowledge the many people who have helped me in one way or another during my doctoral studies. My sincere appreciation goes to my two advisors, Drs. Wayne Adams and Barbara Sears without whom this whole work would not have taken shape. Their guidance and help throughout this study as well as those of the two other members of my committee, Drs. Jim Kelly and Jim Hancock are gratefully acknowledged. Many thanks to Dr. Dave Douches for reviewing my dissertation and accepting to replace Jim Kelly during the defense.

I would like to thank all members of the Sears Lab and the Bean Group for various help in the lab, stimulating discussions and interactions throughout the course of this work. I am particularly indebted to Ruthy Wolfson and PyungOk Lim for their thorough concern, advice and encouragement especially during the frustrating moments of this research. The true interest of Susan Sprecher, Jorge Acosta, Joe Tohme, and Krista Shellie in my work and their unselfish help in various ways are very much appreciated. Un grand merci to Marie-Claude Fortin for her special friendship and support.

Thanks go to Sally Mackenzie for her advice on the mitochondrial DNA isolation procedure, and to Chritine Chase for providing the bean mtDNA clones. The help of people at the Centro Internacional de Agricultura Tropical, particularly Joe Tohme and Daniel Debouck, in growing the Malawian bean lines for seed increase, determining their phaseolin types and providing seed samples of wild bean accessions is gratefully acknowledged. Thanks also go to Jorge Acosta for providing seed samples of the scarlet runner bean and the tepary bean. I appreciate the help of Rabi' Mohtar in writing the computer program for determining DNA fragment sizes, and of Walid

Shayya in presenting the tabulated data in the fine shape they are in now, and in the final printing of this dissertation.

The initial financial support of the Hariri Foundation and the ensuing support of the Bean/Cowpea Collaborative Research Support Program -Malawi Project- and the Fares Foundation are greatly appreciated.

Special thanks go to Hoda, Hind, Nayla, Thalia, Nohma, Jorge, and many other family members and friends who through their confidence, prayers, thoughts and letters have helped keep me going and made me feel that I am still part of their world despite the distance separating us.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

GENERAL INTRODUCTION

The genetic resources of crop species consist of old and diverse landraces as well as wild relatives, but this natural treasure is quickly disappearing. Genetic diversity is the raw material for plant breeders. Its unavailability or infrequent use can result in crops with a narrow genetic base which increases their vulnerability to pests, diseases, and adverse environmental conditions. The epidemics of the potato late blight in Ireland in 1846 and the southern corn leaf blight in the United States in 1970 are only two famous examples of the disastrous effect of the genetic uniformity of crop plants. Several other cases have caused severe economic losses (Plucknett et al. 1987, Chapter 1). The threat of possible recurrences of such epidemics and the fear of ever-increasing genetic erosion caused by expanding modern civilizations, have prompted worldwide attention toward germplasm collection and conservation. The last twenty years or so have witnessed various national and international efforts at collecting and storing cultivated and wild germplasm of our most important crops. The task remains, however, to evaluate, enhance, and utilize the available genetic resources.

The analysis and characterization of genetic diversity is of fundamental importance not only to plant breeders but also to population geneticists in order to evaluate the course of evolution in a species and to understand the forces shaping its genetic diversity. The methods for measuring genetic diversity have expanded over the years from the analyses of discrete morphological and cytological variation, to the statistical analyses of quantitative variation, to assays of biochemical attributes, and finally, to molecular assays (Clegg 1990).

The use of molecular techniques to examine DNA has provided an unprecedented view of the kind and amount of plant genetic diversity. Since both coding and non-coding sequences can be analyzed, one can examine segments of DNA that are relatively conserved due to selective pressure, as well as those that are presumably free of the constraints of natural selection. The main approaches used to study the variability at the DNA level consist of sequencing particular regions of DNA for comparison, and surveying restriction site differences. The first approach is the most informative, however, until recent technological innovations, it was not practical for the large number of samples needed in detailed surveys of intraspecific diversity. This is also true when generating physical maps either of specific DNA regions (e.g. various single-copy genes in Drosophila, see Clegg 1990) or of the whole genome (e.g. animal mtDNA, DeSalle et al. 1986). The analysis of restriction fragment length polymorphisms (RFLPs), although not providing the most accurate picture of DNA variation, is less tedious and more appropriate when large numbers are to be analyzed. RFLPs have been extensively used as molecular markers in chromosome mapping efforts, evolutionary studies, and as aids to crop improvement programs (Beckmann and Soller 1986, Song et al. 1990).

In contrast with the nucleus, organelles have much smaller genomes and show non-Mendelian modes of transmission. While most animal and plant mitochondria are strictly maternally inherited, chloroplasts are transmitted from the maternal parent (in most angiosperms), the paternal side (in conifers and in alfalfa), or biparentally (e.g. Pelargonium, Oenothera) (Whatley 1982, Neale and Sederoff 1988, Schumann and Hancock 1989). The study of organelle DNA variation has allowed the tracing of maternal lineages in allopolyploid species (Palmer et al. 1983, Bland et al. 1985, Murai and Tsunewaki 1986), clarified phylogenetic relationships among related species and genera (Berthou et al. 1983, DeBonte et al. 1984, McClean and Hanson 1986), and revealed the extent of cytoplasmic diversity of crop species (Sisson et al. 1978,

Holwerda et al. 1986, Rines et al. 1988).

The plant chloroplast genome consists of a single closed circular molecule and has been shown to be quite conserved in size (120-217 kb), gene structure and DNA sequence (Palmer 1990). On the other hand, the plant mitochondrial genome is larger and more variable in size (200-2500 kb) than any other organelle genome; it is characterized by frequent rearrangements but has slow rates of nucleotide substitution (Palmer 1990). Several studies of organelle DNA variability have shown that mitochondrial genome diversity is more extensive than that of the chloroplast genome (Timothy et al. 1979, DeBonte et al. 1984, Holwerda et al. 1986, McClean and Hanson 1986, Murai and Tsunewaki 1986, Terachi and Tsunewaki 1986, Ichikawa et al. 1989). Examination of mitochondrial DNA (mtDNA) variation thus seems more appropriate when probing cytoplasmic diversity within plant species.

Some mtDNA studies have revealed informative levels of intraspecific variability. Timothy et al. (1979) have looked at organelle DNA variation within the maize:teosinte species complex using two restriction endonucleases. The differences observed in banding patterns approximated the biosystematic relationships of the taxa. Later on, Kemble et al. (1983) and Weissinger et al. (1983) studied the patterns of mtDNA variation in maize races indigenous to Mexico and Latin America, respectively, and reported considerable variation within Zea mays. The first group examined the mitochondrial genome of 25 maize races by agarose gel electrophoresis of uncut DNA followed by hybridization with probes that contained unique DNAs associated with cytoplasmic male sterility (CMS). They further assessed mtDNA variability by BamHI digestion, fractionation and hybridization. Although their work was more directed towards CMS factors, they reported variability in the mtDNA of accessions belonging to the same race, and of plants from a single accession. They also presented evidence for a non-random geographical distribution of mtDNA variants within Mexico, wherein one type, found in the oldest races, appeared to be widely dispersed, while another, less

common type, appeared to be confined to coastal regions. Weissinger et al. (1983) examined 93 maize races by electrophoresis of undigested and BamHI and EcoRI cleaved mtDNAs. They were able to distinguish 18 groups on the basis of restriction patterns and the presence of plasmid-like mtDNAs. Their racial grouping by cytoplasmic distinctions provided several insights into maize evolution and migration. Both studies reported substantial agreement between mtDNA groups and racial affinities based on morphological and cytological data. The only other report to date of high intraspecific mtDNA variation is in Daucus carota. Ichikawa et al. (1989) have used four endonucleases to compare mtDNA restriction patterns between 13 carrot cultivars. These clustered into three groups that differed significantly in their banding profiles (only 51-59% of the fragments were shared), suggesting that the mtDNA diversity existed before the domestication of carrots.

In contrast, investigations of mtDNA diversity in soybean (Sisson et al. 1978, Grabau et al. 1989), barley (Holwerda et al. 1986), tomato (McClean and Hanson 1986), millet (Chowdhury and Smith 1988), oat (Rines et al. 1988), Brassica (Palmer 1988) and Beta vulgaris (Ecke and Michaelis 1990) have revealed high degrees of intraspecific uniformity of mtDNA restriction patterns, with only one to two band differences in most cases.

Using two restriction endonucleases, Sisson et al. (1978) have examined the mtDNAs of eight soybean lines that form the maternal ancestors of most soybean cultivars grown in the United States. The high degree of homology encountered is an alarming indication of the cytoplasmic uniformity of the US soybean cultivars. In a more extensive study, Holwerda et al. (1986) surveyed 28 accessions of wild and cultivated barley for organelle DNA variation. The lines were collected in eight countries in the Near East region, believed to be the primary center of diversity of the species. Limited variation in the mtDNAs of both wild and cultivated barley was obvious. Of 860 mtDNA restriction fragments generated by 16 enzymes, 50 (or 5.8%)

were variant, and thought to have risen from 14 independent changes, with no direct evidence of nucleotide substitutions. In some cases, an affinity existed between landraces and wild accessions collected from the same locality.

The relatively small mitochondrial genome of the Brassicas (208-242 kb) makes it a relatively easy system in which to study the amount and kind of mtDNA variation. Palmer (1988) carried out detailed restriction site mapping of mtDNA molecules in 2-5 lines from each of eight Brassica species, using five endonucleases. The limited variation encountered within each species consisted entirely of structural changes; out of 140 restriction sites surveyed, none had mutated. The observation that sequence rearrangements rather than nucleotide substitutions play a more important role in plant mitochondrial genome evolution has been reported in maize and teosinte (Sederoff et al. 1981).

MtDNA restriction and hybridization patterns have also been used to determine levels of diversity and phylogenetic relationships among species (Berthou et al. 1983, Debonte et al. 1984, Baatout et al. 1985, Terachi and Tsunewaki 1986, McClean and Hanson 1986, Chowdhury and Smith 1988). Significantly higher degrees of variability contrast with the mtDNA homogeneity of intraspecific comparisons. Due to the large sizes and high rates of rearrangements of plant mitochondrial genomes, measures of mtDNA relatedness between plant species have been limited to the estimation of the proportion of shared restriction fragments between any two species. Plant phylogenies have been constructed using such indices and following the phenetic approach (McClean and Hanson 1986, Terachi and Tsunewaki 1986).

This dissertation reports on mitochondrial genome diversity within Phaseolus vulgaris, both domesticated and wild forms, and this comparison is extended to two other cultivated bean species. The genus Phaseolus originated in the American continent and comprises about thirty species (Maréchal et al. 1978), of which

approximately 90% occur in Mexico or adjacent areas of the USA and Central America (Smartt 1985). Four species have been most prominent as cultivated food crops: P. vulgaris L. (common bean, dry bean, snap bean), P. lunatus L. (lima bean, sieva bean), P. coccineus L. (runner or scarlet runner bean), and P. acutifolius Gray (tepary bean). Of these four, however, P. vulgaris makes up more than 90% of the cultivated crop worldwide (Singh 1989).

The primary center of origin of the common bean was held in dispute by different authors. Based on botanical, genetic, and archaeological data, Miranda-Colín (1967) and Gentry (1969) suggested that P. vulgaris was first domesticated in Mesoamerica while Berglund-Brücher and Brücher (1976) believed in a South American origin. The possibility of independent origins for North and South American common beans was discussed by Kaplan (1981) although previously suggested by several authors (Heiser 1965, Gentry 1969, Harlan 1971, Evans 1973).

Recent evidence from botanical explorations, genetic studies and biochemical assays have left no doubt that a non-centric origin applies to P. vulgaris, since multiple centers of domestication are apparent (reviewed by Gepts 1988a). Extensive botanical explorations indicate an almost continuous distribution of wild P. vulgaris from Mexico (State of Chihuahua) in the north, to Argentina (Province of San Luis) in the south (Debouck and Tohme 1989). Examination of phaseolin, the major seed storage protein of beans, in cultivated and wild P. vulgaris accessions, suggest three sites of domestication: one in Mesoamerica, leading to small-seeded (<25 g/100 seeds) cultivars with 'S' phaseolin types; one in the southern Andes giving rise to large-seeded (>40g/100 seeds) cultivars with 'T', 'C', 'H' and 'A' phaseolin; and a minor area in the northern Andes (Colombia) resulting in small-seeded cultivars with a 'B' phaseolin pattern (Gepts and Bliss 1986, Gepts et al. 1986).

Two major gene pools of cultivated common bean are thus recognized as the small-seeded Mesoamerican pool and the large-seeded Andean one. In addition to

their seed size and phaseolin types, they are distinguished by their environmental adaptation, disease resistance, growth habits, and isozyme variants (reviewed by Gepts 1988b). Moreover, the two germplasm pools are separated by partial fertility barriers; two such systems have been reported. In the first, two complementary dominant genes, DL_1 and DL_2 (for Dosage Lethal, Shii et al. 1980), are predominant in the small-seeded and the large-seeded germplasm, respectively, and produce hybrid weakness symptoms in F_1 and F_2 individuals from crosses between the gene pools (Singh and Gutierrez 1984, Gepts and Bliss 1985). In the second system, male sterility symptoms have been observed among gene pool recombinants identified by isozyme analysis (Sprecher 1988).

The variability of P. vulgaris is not limited to the differences between the major gene pools; it is illustrated by the multitude of seed classes and the remarkable variation in seed colors and patterns. Using seed size and some vegetative characters, Evans (1973,1976) grouped the genetic diversity of the common bean into five races. Vanderborght (1987) analyzed the diversity of landraces from Mexico and South America using multivariate statistical methods and found four natural groups that correspond to the different types of growth habits of the common bean. Singh (1989) proposed ten "working gene pools" for dry beans (six from the Mesoamerican and four from the Andean centers of domestication) and two for snap beans to facilitate germplasm management and utilization in breeding programs. The groups were defined as a result of the evaluation of 18,300 accessions from the germplasm collection of the Centro Internacional de Agricultura Tropical (CIAT), for morphological, phenological and seed characteristics. Specific patterns of variation were observed according to geographical distributions of bean types from the American continent and from other areas where beans are grown.

The variability of P. vulgaris is probably greater in, though not confined to, its primary centers of domestication and cultivation. The wide dissemination of the

common bean in Europe, Asia and Africa has led to the establishment of secondary centers of diversity (Smartt 1985). Martin and Adams (1987) have demonstrated the existence of a "vast store of genetic variability" in bean landraces from Malawi, in East Africa. They examined the variation among and within 15 landraces (which sometimes are deliberate mixtures of various seed types) collected in northern Malawi. They analyzed 25 lines in each landrace for morphological, phenological and agronomic traits using multivariate statistics. The large amount of variability they observed contrasts with the results of Sprecher (1988) who tested the same bean lines for allozyme diversity. Sprecher (1988) found the lines to cluster into two major groups that had alternate alleles at six isozyme loci, and these groups corresponded to the small-seeded Mesoamerican and the large-seeded Andean gene pools. Although large-seeded beans predominate and are preferred by consumers in East Africa, small-seeded types are also grown, often on the same farm, and even in the same mixture.

The objectives of the first chapter of this dissertation were to examine the diversity of a selected sample of the Malawian lines at the DNA level and compare and contrast it to the morphoagronomic and allozyme variation. The mitochondrial genome was chosen in order to consider the cytoplasmic diversity, and because of better chances of encountering intraspecific variability relative to the chloroplast genome (Lee 1988). The results show an appreciable degree of uniformity of the bean mitochondrial genome, however, five RFLPs differentiate again the Malawian lines into the small-seeded and the large-seeded groups.

In the second chapter, the mtDNA diversity was examined in six wild P. vulgaris accessions from across the natural distribution range in order to understand how and when the mitochondrial genomes of the two gene pools became distinct. The results reveal higher mtDNA variation in the wild than in the cultivated material and provide information on the timing and direction of four of the mtDNA mutations that resulted in the inter-gene pool RFLPs.

Finally, in the third chapter, the amount of mtDNA variation among three Phaseolus species is examined and compared to that within P. vulgaris. In addition, the mitochondrial genome size is estimated for the three species.

LIST OF REFERENCES

Baatout H, Marrakchi M, Mathieu C, Vedel F (1985) Variation of plastid and mitochondrial DNAs in the genus Hedysarum. Theor Appl Genet 70:577-584

Beckmann JS, Soller M (1986) Restriction fragment length polymorphisms in plant genetic improvement. Oxford Surveys Plant Mol Cell Biol 3:196-250

Berglund-Brücher O, Brücher H (1976) The south American wild bean (Phaseolus aborigineus Burk.) as ancestor of the common bean. Econ Bot 30:257-272

Berthou F, Mathieu C, Vedel F (1983) Chloroplast and mitochondrial DNA variation as indicator of phylogenetic relationships in the genus Coffea L. Theor Appl Genet 65:77-84

Bland MM, Matzinger DF, Levings CS III (1985) Comparison of the mitochondrial genome of Nicotiana tabacum with its progenitor species. Theor Appl Genet 69:535-541

Chowdhury MKU, Smith RL (1988) Mitochondrial DNA variation in pearl millet and related species. Theor Appl Genet 76: 25-32

Clegg MT (1990) Molecular diversity in plant populations. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) Plant population genetics, breeding, and genetic resources. Sinauer Ass Inc Publ, Sunderland, Massachusets, pp 98-115

DeBonte LR, Matthews BF, Wilson KG (1984) Variation of plastid and mitochondrial DNAs in the genus Daucus. Amer J Bot 71: 932-940

Debouck DG, Tohme J (1989) Implications for bean breeders of studies on the origins of common beans, Phaseolus vulgaris L. In: Beebe S (ed) Current topics in breeding of common bean. Working Document No. 47. Bean Program, CIAT; Cali, Colombia pp 3-42

DeSalle R, Val Giddings L, Templeton AR (1986) Mitochondrial DNA variability in natural populations of Hawaiian Drosophila. I. Methods and levels of variability in D. silvestris and D. heteroneura populations. Heredity 56:75-85

Ecke W, Michaelis G (1990) Comparison of chloroplast and mitochondrial DNA from five morphologically distinct Beta vulgaris cultivars: sugar beet, fodder beet, beet root, foliage beet, and Swiss chard. Theor Appl Genet 79:440-442

Evans AM (1973) Genetic improvement of Phaseolus vulgaris. In: Milner M (ed) Nutritional improvement of food legumes by breeding. Protein Advisory Group, UN, New York, pp 107-115

Evans AM (1976) Beans. In: Simmonds NW (ed) Evolution of crop plants. Longman, London, pp 168-172

Gentry HS (1969) Origin of the common bean, Phaseolus vulgaris. Econ Bot 23:55-69

Gepts P (1988a) Phaseolin as an evolutionary marker. In: Gepts P(ed) Genetic resources of Phaseolus beans. Kluwer Acad Publ Dordrecht, pp 215-241

Gepts P (1988b) A middle American and an Andean common bean gene pool. In: Gepts P (ed) Genetic resources of Phaseolus beans. Kluwer Acad Publ, Dordrecht, pp 375-390

Gepts P, Bliss FA (1985) F_1 hybrid weakness in the common bean: differential geographic origin suggests two gene pools in cultivated bean germplasm. 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

Grabau EA, Davies WH, Gengenbach BG (1989) Restriction fragment length polymorphism in a subclass of the 'Mandarin' soybean cytoplasm. Crop Sci 29:1554-1559

Harlan JR (1971) Agricultural origins: centers and non-centers. Science 174:468-474

Heiser CB Jr (1965) Cultivated plants and cultural diffusion in nuclear America. Amer Anthropol 67:930-949

Holwerda BC, Jana S, Crosby WL (1986) Chloroplast and mitochondrial DNA variation in Hordeum vulgare and Hordeum spontaneum. Genetics 114:1271-1291

Ichikawa H, Tanno-Suenaga L, Imamura J (1989) Mitochondrial genome diversity among cultivars of Daucus carota (ssp sativus) and their wild relatives. Theor Appl Genet 77:39-43

Kaplan L (1965) Archaeology and domestication in American Phaseolus (Beans). Econ Bot 19:358-368

Kaplan L (1981) What is the origin of the common bean? Econ Bot 35:240-254

Kemble RJ, Gunn RE, Flavell RB (1983) Mitochondrial DNA variation in races of maize indigenous to Mexico. Theor Appl Genet 65:129-144

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

Koenig RL, Singh SP, Gepts P (1990) Novel phaseolin types in wild and cultivated common bean (Phaseolus vulgaris, Fabaceae). Econ Bot 44:50-60

Lee PC (1988) Studies of chloroplast transmission and F2 hybrid variegation in the common bean Phaseolus vulgaris. MS thesis, Michigan State University, East Lansing, MI.

Maréchal R, Mascherpa J-M, Stainier F (1978) Etude taxonomique d'un groupe complexe d'espèces des genres Phaseolus et Vigna (Papilionaceae) sur la base de données morphologiques et polliniques, traitées par l'analyse informatique. Boissiera 28:1-273

Martin GB, Adams MW (1987) Landraces of Phaseolus vulgaris (Fabaceae) in Northern Malawi. I. Regional variation. Econ Bot 41:190-203

McClean PE, Hanson MR (1986) Mitochondrial DNA sequence divergence among Lycopersicon and related Solanum species. Genetics 112:649-667

Miranda-Colín S (1967) Origen de Phaseolus vulgaris L. (frijól comun). Agrociencia 1:99-109

Murai K, Tsunewaki K (1986) Phylogenetic relationships between Avena species revealed by the restriction endonuclease analysis of chloroplast and mitochondrial DNAs. In: Lawes DA Thomas H (ed) Proc 2nd Int Oats Conf, Aberystwyth, Wales. 15-18 July 1985. Martinus Nijhoff Publ, Nordrecht, Netherlands. pp 34-38

Palmer JD (1988) Intraspecific variation and multicircularity in Brassica mitochondrial DNAs. Genetics 118:341-351

Palmer JD (1990) Contrasting modes and tempos of genome evolution in land plant organelles. Trends in Genetics 6:115-120

Palmer JD, Shields CR, Cohen DB, Orton TJ (1983) Chloroplast DNA evolution and the origin of amphidiploid Brassica species. Theor Appl Genet 65:181-189

Plucknett DL, Smith NJH, Williams JT, Murthi Anishetty N (1987) Gene banks and the World's food. Princeton Univ. Press, Princeton, New Jersey. 247 pages

Rines HW, Gengenbach BG, Boylan KL, Storey KK (1988) Mitochondrial DNA diversity in oat cultivars and species. Crop Sci 28:171-176

Schumann CM, Hancock JF (1989) Paternal inheritance of plastids in Medicago sativa. Theor Appl Genet 78:863-866

Sederoff RR, Levings CS, Timothy DH, Hu WWL (1981) Evolution of DNA sequence organization in mitochondrial genomes of Zea. Proc Natl Acad Sci USA 78:5953-5957

Shii CT, Mok MC, Temple SR, Mok DWS (1980) Expression of developmental abnormalities in hybrids of Phaseolus vulgaris L.: Interaction between temperature and allelic dosage. J Hered 71:218-222

Singh SP (1989) Patterns of variation in cultivated common bean (Phaseolus vulgaris, Fabaceae). Econ Bot 43:39-57

Singh SP, Gutierrez JA (1984) Geographical distribution of the DL_1 and DL_2 genes causing hybrid dwarfism in Phaseolus vulgaris L., their association with seed size, and their significance to breeding. Euphytica 33:337-345

Sisson VA, Brim CA, Levings CS III (1978) Characterization of cytoplasmic diversity in soybeans by restriction endonuclease activity. Crop Sci 18:991-996

Smartt J (1985) Evolution of grain legumes. IV. Pulses in the genus Phaseolus. Expl Agric 21:193-207

Song K, Osborn TC, Williams PH (1990) Brassica taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs) 3. Genome relationships in Brassica and related genera and the origin of B. oleracea and B. rapa (syn. campestris). Theor Appl Genet 79:497-506

Sprecher SL (1988) Allozyme differentiation between gene pools in common bean (Phaseolus vulgaris L.), with special reference to Malawian germplasm. PhD Dissertation, Michigan State University, East Lansing, MI. 207 pages

Terachi T, Tsunewaki K (1986) The molecular basis of genetic diversity among cytoplasms of Triticum and Aegilops 5. Mitochondrial genome diversity among Aegilops species having identical chloroplast genomes. Theor Appl Genet 73:175-181

Timothy DH, Levings CS III, Pring DR, Conde MF, Kermicle JL (1979) Organelle DNA variation and systematic relationships in the genus Zea:Teosinte. Proc Natl Acad Sci USA 76:4220-4224

Vanderborght T (1987) The study of common bean (Phaseolus vulgaris L.) variability by the use of multivariate statistical methods applied to a data base. Diss Abstr 47:3190B-3191B

Whatley JM (1982) Ultrastructure of plastid inheritance: green algae to angiosperms. Biol Rev 57:527-569

Weissinger AK, Timothy DH, Levings CS III, Goodman MM (1983) Patterns of mitochondrial DNA variation in indigenous maize races of Latin America. Genetics 104:365-379

CHAPTER ONE

MITOCHONDRIAL DNA POLYMORPHISMS OF MALAWIAN BEAN LINES: FURTHER EVIDENCE FOR TWO MAJOR GENE POOLS

ABSTRACT. Intraspecific mitochondrial DNA (mtDNA) diversity was determined in 23 Phaseolus vulgaris genotypes, and compared to previously observed variability of morphoagronomic characters and isozyme loci. Twenty of the lines were collected from Malawian landraces; the other three were pure-bred cultivars. The mtDNAs were digested with eight restriction endonucleases, revealing complex banding patterns. Southern hybridization using cosmid clones covering about 200 kb of the genome showed a considerable amount of uniformity of the mtDNA banding patterns. However, five restriction fragment length polymorphisms (RFLPs) were detected, dividing the bean lines into two groups corresponding to the previously known Mesoamerican and Andean gene pools of P. vulgaris. The light red kidney cultivar 'Mecosta' was separated from the rest of the lines by an additional RFLP. At least two out of the six RFLPs are believed to be due to basepair mutation events. Our results provide the first evidence that the cytoplasms of the two major germplasm pools of beans are distinct.

INTRODUCTION

The common bean, Phaseolus vulgaris L., is native to the American continent where the wild ancestral forms are still distributed from West Central Mexico to Northwestern Argentina. Botanical, archaeological, and biochemical evidence indicate two major and independent domestication events in Mesoamerica and the Andes (reviewed by Gepts 1988a; Debouck and Tohme 1989), which resulted in two primary centers of genetic diversity. Consequently, two major gene pools are recognized in the common bean (for a review, see Gepts 1988b) . These are distinguished by their seed sizes and growth habits (Evans 1976), environmental adaptation (Ghaderi et al. 1982; Kelly et al. 1987), disease resistance (Gepts and Bliss 1985), isozyme variants (Bassiri and Adams 1978; Sprecher 1988), and phaseolin types (Gepts et al. 1986). In addition, the two germplasm pools are separated by partial fertility barriers (Singh and Gutierrez 1984; Gepts and Bliss 1985; Sprecher 1988).

As beans from these previously separate gene pools spread to the Old World, secondary centers of diversity were formed. One such area is Eastern Africa where Andean beans predominate, although strains from Mesoamerica are also grown. Recent germplasm collections in Malawi have drawn attention to the abundant variability in farmers' fields (Adams 1983, Martin 1984). Studies on 375 landrace lines collected from 15 sites in northern Malawi revealed extensive morphological, phenological and agronomic variation (Martin and Adams 1987a and b), but showed that the lines cluster into only two major groups in terms of variability at isozyme loci (Sprecher 1988). These correspond to the small-seeded Mesoamerican and the large-seeded Andean gene pools. In this investigation, we selected a subsample of these lines to assess the variability at an additional level, that of the mitochondrial DNA (mtDNA). Our objectives were (1) to compare and contrast the amount and pattern of mtDNA diversity with the morphoagronomic and isozyme variation; and (2) to explore the possibility that the two major gene pools of beans have evolved distinct mitochondrial genomes.

An increasing number of studies have compared organelle DNA restriction patterns as a measure of diversity and to assess phylogenetic relationships within and among plant species. Because of its higher rate of rearrangement, mtDNA appeared to be more useful than chloroplast DNA for studies below the species level. Indeed, studies in maize and teosinte had revealed a considerable level of mtDNA intraspecific variability (Timothy et al. 1979; Kemble et al. 1983; Weissinger et al. 1983). However, in most other species examined subsequently, limited variation has been observed (e.g. Holwerda et al. 1986; Rines et al. 1988; Palmer 1988), rendering intraspecific comparisons of mtDNA less informative. In P. vulgaris, as in other species, mtDNA restriction fragment differences were observed when cytoplasmic male sterile lines and male fertile ones were compared (Mackenzie et al. 1988). Another problem is that the mitochondrial genome is relatively complex and RFLPs can be difficult to resolve. In our analysis of mtDNA variability from 23 bean lines, we addressed these difficulties by using eight restriction endonucleases and by Southern hybridizations to mtDNA probes covering almost half the genome.

Our results show that the mitochondrial genome of Phaseolus vulgaris is characterized by an appreciable degree of uniformity. In addition, we provide the first evidence that the cytoplasms of the two major bean gene pools are distinct as defined by five mtDNA RFLPs.

MATERIALS AND METHODS

Bean lines

Bean landraces were collected in 1983 from 15 farm sites in northern Malawi. Twenty-five lines were extracted from each site by allowing plants derived from 25 individual seeds picked at random from the on-site collection to self pollinate (Martin and Adams 1987a). Of those 375 lines, our investigation used 20, originating at two sites in the Misuku Hills, Farms # 4 and 5 in Martin and Adams (1987a). Based on

genetic distances derived from a principal component analysis of morpho-agronomic characters (Martin 1984, Martin and Adams 1987a) both "related" and "distant" lines were included. The lines differ in seed characteristics, and include six isozyme genotypes (Sprecher 1988), and three phaseolin types (Table 1.1).

Three pure-line cultivars, 'Mecosta', 'Sanilac', and 'Tendergreen', were included as controls. The latter two are the cultivar types for the phaseolin variants 'S' and 'T', respectively (Brown et al. 1981).

The Malawian lines were grown in the field at Michigan State University for two summers in order to obtain enough seeds for the mtDNA extractions.

MtDNA isolation

Using 10-21 day old seedlings grown in the dark, mtDNA was isolated according to Mackenzie et al. (1988), with some modifications. 100-200 g of etiolated tissue (including roots) were homogenized in a Waring blender using 5-7 volumes (w/v) cold buffer 'A' (0.5M sucrose, 0.05M Tris, 0.005M EDTA, 0.1% BSA, pH 7.5, and 0.005 M mercaptoethanol added just before use). All subsequent steps were carried out at 4^{o}C or on ice unless otherwise specified. The homogenate was filtered through a 100 micron mesh nylon screen, then through two layers of Miracloth (Calbiochem) and centrifuged for 10 min at 1,000 x g to pellet nuclei and chloroplasts. The resulting supernatant was centrifuged at 15,900 x g for 10-15 min and the pellet resuspended with a paint brush in buffer 'G' (0.3M sucrose, 0.05M Tris, pH 7.5) using 20 ml/ 100 g starting material. After another 10 min spin at 1,000 x g the supernatant was brought to 10mM $MgCl_2$ and 20 ug/ml or 200 Kunitz units/ml DNase I (Sigma Chemical Co., Type IV), and incubated with gentle mixing at 4^{o}C for 60-90 minutes. The suspension was then underlaid with 20 ml shelf buffer per 10-15 ml (0.6M sucrose, 0.01M Tris, 0.02M Na_2EDTA, pH 7.5) and centrifuged for 20 min at 12,000 x g. The same buffer was used to resuspend the pellet (10 ml/50 g), and the mitochondria were pelleted by centrifuging for another 20 min at 15,900 x g. They were then lysed in 5 ml lysis

Table 1.1. Characteristics of the bean lines assayed for mtDNA variability.

Bean Line	Isozyme Genotype[a]	Gene Pool[a]	Phase-olin[b]	Bean Line	Isozyme Genotype[a]	Gene Pool[a]	Phase-olin[b]
4-1 	1	A	T	5-1 	7	M	S
4-4 	7	M	S	5-2 	7	M	S
4-6 	1	A	T	5-4 	7	M	S
4-7 	1	A	T	5-5 	7	M	S
4-10 	0	H	S	5-6 	1	A	T
4-11 	7	M	S	5-7 	9	A	T
4-20 	4	A	T	5-10 	1	A	T
4-22 	1	A	T	5-14 	1	A	C
Mecosta 	-	A	S	5-16 	1	A	C
Sanilac 	7	M	S	5-18 	7	M	S
Tendergreen 	1	A	T	5-20 	11	A	C
				5-25 	0	H	T

a From Sprecher (1988), based on the analysis of six isozyme loci.
A, M, H denote Andean, Mesoamerican, and heterozygous isozyme genotypes.

b Determined at the Centro Internacional de Agricultura Tropical
T = 'Tendergreen'; S = 'Sanilac'; and C = 'Contender' banding patterns (Brown et al. 1981).
T and C - Andean gene pool; S - Mesoamerican gene pool (Gepts 1988b).

buffer per 100 g tissue (0.1M Tris, 0.05M EDTA, 0.1M NaCl, 1% SDS, pH 8.0) and 100 ug/ml proteinase K at 65°C for 30 minutes. The protein-carbohydrate complexes were precipitated by adding one third the volume 5M K-acetate, incubating on ice for one hour with periodic shaking, and centrifuging for 20 min at 25,000 x g. The supernatant was passed through one layer of Miracloth into a siliconized Corex tube and the mtDNA precipitated with 1/20 vol. 5M ammonium acetate and 1/2 vol. ice cold isopropanol at -20°C.

The mtDNA was pelleted by centrifuging at 10,000 x g for 45 min; the pellets were air dried, and then dissolved in 0.7 ml $T_{50}E_{10}$ (50mM Tris, 10mM EDTA, pH 8.0). The solution was transferred to a microfuge tube and spun for 10 minutes. The supernatant was transferred to a clean tube and the mtDNA was precipitated with ammonium acetate and isopropanol as before. The mtDNA was washed twice with 70% ethanol, vacuum dried for 10-15 min, and allowed to dissolve in TE (10 mM Tris, 1 mM EDTA, pH 8.0) overnight at 4°C. Only mtDNAs that did not cut well with restriction endonucleases were phenol-chloroform extracted.

Restriction endonuclease analysis

The mtDNAs were digested with individual restriction endonucleases for 4-7 hours at 37°C using the buffers recommended by the manufacturers (Bethesda Research Lab, Inc., Boehringer Mannheim Biochemicals, and New England Biolabs). The fragments were separated by electrophoresis in 0.7% (DraI, PstI, SalI, XhoI), 0.8% (BamHI), 0.9% (EcoRI, HindIII) or 1.3% (HaeIII) agarose gels (Sigma, Type I: low EEO) using the TAE buffer system (0.04M Tris, 0.02M sodium acetate, 0.001M EDTA, pH 8.0, and 0.5 ug/ml ethidium bromide). Molecular size markers were obtained by digesting Lambda DNA with HindIII alone and in combination with EcoRI. Gels were run at room temperature for 18-22 hours at constant voltage, 35-45 volts.

After the gels were photographed, the restriction fragments were transferred to nylon hybridization membranes (0.45 micron, Micron Separations Inc.) as described by Maniatis et al. (1982). The DNA was fixed on the membrane by either overnight baking at 80°C or UV crosslinking for 3-5 minutes.

The membranes were consecutively hybridized to 3-4 different probes after stripping away the previous radioactive probe by incubating the filters in 0.5M NaOH at 50°C for one hour (Nugent and Palmer 1988).

Nick translation and hybridization

The probes used in this study (651-12-C2, -C3, -C4, -C6 and -C8) are random DNA inserts of the bean mitochondrial genome cloned into the cosmid pHC79, kindly provided by Dr. C.D. Chase (University of Florida, Gainesville). The insert sizes are as follows: C2: 37.9kb, C3: 33.2kb, C4: 35.5kb, C6:27.7kb, and C8:34.6kb.

The clones were nick-translated using ^{32}P-labeled nucleotides (New England Nuclear) based on the procedure described by Maniatis et al. (1982) except that the nick translation buffer was that of Rigby et al. (1977). The Southern filters were hybridized at 65°C overnight and then washed of excess label according to Maniatis et al. (1982). However, in order to facilitate the later stripwashing of the probe, the membranes were not allowed to dry before exposing them. We also gradually decreased the stringency of the third wash for the second, third, and fourth probe by lowering either the temperature or the time of incubation, or a combination of both. Filters were used for autoradiography with Kodak XAR-5 X-ray film in cassettes containing one or two intensifying screens at -70°C.

RESULTS

The mitochondrial genome of beans is large and its restriction patterns complex

Mitochondrial DNAs from the 23 bean lines were digested with the eight restriction endonucleases chosen for this study. The patterns of fragments resulting after electrophoresis were complex with at least 30 bands in the simplest profile (SalI) and a minimum of 60 bands with HindIII. The patterns of the different lines were very similar and it was hard to identify RFLPs with certainty in the stained gels. However, we were able to detect a few differences in the digestion patterns from two enzymes, DraI (Figures 1.1A and 1.3A) and EcoRI (Figure 1.2A). We ruled out the possibility of partially digested fragments by always adding excess enzyme and allowing long digestion times. Furthermore, the same RFLPs were seen in at least six gels for each enzyme.

Besides the six fragment differences indicated on the examples of the ethidium bromide stained gels (four in DraI and two in EcoRI), no other RFLPs were seen. Recognizing that because of the complexity of the restriction patterns, additional variability could go undetected in the stained gels, we decided to screen for more RFLPs by Southern analysis.

Little variation exists in the mitochondrial genome of bean

By transferring mtDNAs from the gels to nylon membranes for Southern hybridizations, we were able to more clearly examine specific regions of the mitochondrial genome. The screening was started with the use of cloned mitochondrial coding sequences from maize (cytochrome oxidase subunit II gene region provided by T.D. Fox, Cornell University, Ithaca, and the 26S rDNA, and 5S-18S rDNA obtained from C.S. Levings III, North Carolina State University, Raleigh). With these relatively small probes, one to four fragments per enzyme hybridized and no RFLPs were detected. We then decided to use larger cosmid clones (29 - 38kb) consisting of random inserts of the bean mitochondrial genome. Each clone hybridized to 3 - 17

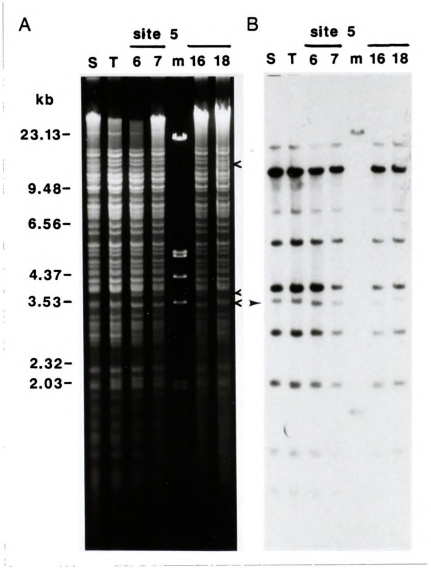

Figure 1.1. DraI digestion patterns of six bean mtDNAs. S=Sanilac, T=Tendergreen and four bean lines from site 5. (A) Ethidium bromide stained gel, m=molecular size markers. (B) Southern blot hybridized with clone C6. Polymorphic bands in the gel are indicated with open arrows (from the top, RFLPs # 1,2,3) and in the blot by a closed one (RFLP # 3).

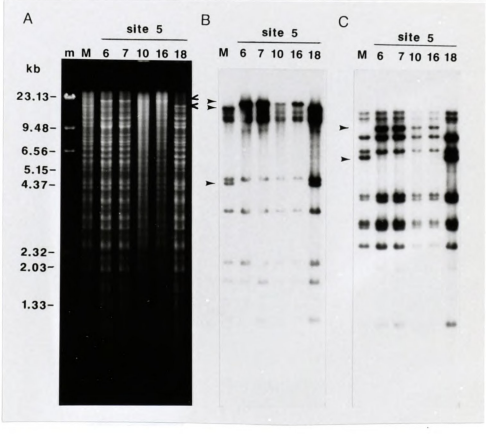

Figure 1.2. EcoRI digestion patterns of six bean mtDNAs. M=Mecosta and five bean lines from site 5. (A) Ethidium bromide stained gel, m=molecular size markers. (B,C) Southern blots hybridized respectively with clones C2 and C4. Polymorphic bands in stained gel are indicated with open arrows (part of RFLP # 4) and in the blots by closed arrows (RFLP # 4 in B and # 5 in C).

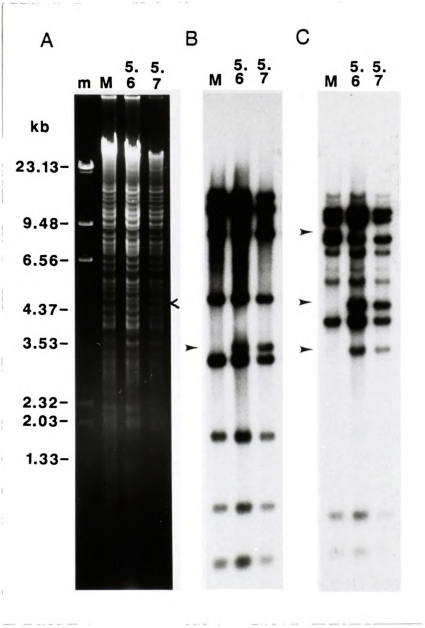

Figure 1.3. DraI digestion patterns of three bean mtDNAs. M=Mecosta and two bean lines from site 5. (A) Ethidium bromide stained gel, m=molecular size markers. (B,C) Southern blots hybridized respectively with clones C2 and C4. Polymorphic bands in the gel are indicated with an open arrow (part of RFLP # 7) and in the blots by closed arrows (RFLP # 6 in B and # 7 in C).

fragments per enzyme, with a total of 29 - 64 fragments per enzyme being revealed with five cosmid clones. These numbers were adjusted to 20 - 46 fragments per enzyme by accounting for possible overlapping sequences between the different clones (Table 1.2).

The Southern hybridizations showed very uniform patterns among the bean lines, and as Table 1.2 indicates, only two enzymes, DraI and EcoRI produced polymorphic fragments (Figures 1.1B, 1.2B,C, and 1.3B,C). Four of the polymorphic bands seen in the stained gels hybridized to the probes. The two non-hybridizing fragments and the ones revealed by hybridization total 13 polymorphic fragments, representing seven RFLPs (Table 1.3). Among all 23 bean lines, only two alternative patterns were observed for each of the seven RFLPs, pattern #1 and #2 defined in Table 1.3. For both RFLP 3 and 7, two different clones revealed the same polymorphic bands.

In order to estimate the proportion of the bean mitochondrial genome that was covered by the cosmid clones, the size of the fragments hybridizing to each clone was determined and added up for individual enzymes, overlapping fragments being counted only once. The estimates obtained were 200.0 kb, 198.6 kb, and 203.5 kb for DraI, EcoRI, and PstI, respectively.

Two distinct mitochondrial genomes are present in the Malawian beans, representing the two major gene pools.

Most bean lines were analyzed for all enzyme x probe combinations and the results are summarized in Table 1.4. In contrast to RFLPs 6 (DraI with C2) and 7 (DraI with C4 or C8), which differentiate the cultivar Mecosta from all other lines, RFLPs 1 to 5 separate the bean lines into two groups. Among the Malawian lines, the eight exhibiting pattern #1 (Table 1.3) are small-seeded (average= 25g/100 seeds), have S phaseolin, and show the Mesoamerican isozyme allele complement (Table 1.1). The eleven Malawi lines with pattern #2 are medium to large-seeded (average=

Table 1.2. Number of bean mtDNA fragments to which the five cosmid clones hybridize.

Enzyme	Recognition Sequence	Total Number of Fragments	Adjusted Number of Fragments[a]	Number of Polymorphic Fragments[b]
Hae III	GGCC	61	ND	0
Bam HI	GGATCC	54	38	0
Dra I	TTTAAA	48	35	6
Eco RI	GAATTC	62	46	5
Hind III	AAGCTT	63	46	0
Pst I	CTGCAG	37	23	0
Sal I	GTCGAC	29	20	0
Xho I	CTCGAG	44	32	0

a Adjusted by substracting the number of overlapping fragments between clones.
b Detected by Southern hybridization.
ND = not determined.

Table 1.3. Intraspecific RFLPs in bean mtDNA following digestion with eight restriction endonucleases.

RFLP Number	Enzyme	Revealed by	Definition of Pattern # 1	Definition of Pattern # 2	Number of Polymorphic Fragments
1	DraI	stained gel	12.7 kb	no 12.7 kb	1
2	DraI	stained gel	no 3.7 kb	3.7 kb	1
3	DraI	clones C3,C6[a]	3.45 kb	3.37 kb	2
4	EcoRI	clone C2[a]	18.6 + 4.3 kb	22.3 kb	3
5	EcoRI	clone C4	6.2 kb	11.2 kb	2
6	DraI	clone C2	3.3 kb	no 3.3 kb	1
7	DraI	clones C4,C8[a]	4.5 + 3.3 kb	8.5 kb	3

a Also detected in the stained gel.

Table 1.4. Summary of results obtained when mtDNA from 23 bean lines were digested with eight restriction endonucleases and probed with five different mtDNA clones (C2,3,4,6,8).

Detected in Gel or Probe		Bean Line[a]																							
		Farm Site 4								Farm Site 5													Cultivar[b]		
	Enzyme[c]	1 A	4 M	6 A	7 A	10 H	11 M	20 A	22 A	1 M	2 M	4 M	5 M	6 A	7 A	10 A	14 A	16 A	18 M	20 A	25 H	M A	S M	T A
Gel	Dra I[1]	2	1	2	2	1	1	2	2	1	1	1	1	2	2	2	2	2	1	2	2	2	1	2
	Dra I[2]	2	1	2	2	1	1	2	2	1	1	1	1	2	2	2	2	2	1	2	2	1	1	1
C_2	Hae III	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Bam HI	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Dra I[6]	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1
	Eco RI[4]	2	1	2	2	1	1	2	2	1	1	1	1	2	2	2	2	2	1	2	2	1	1	1
	Hind III	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Pst I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Sal I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Xho I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C_3	Hae III	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Bam HI	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Dra I[3]	2	1	2	2	1	1	2	2	1	1	1	1	2	2	2	2	2	1	2	2	1	1	1
	Eco RI	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Hind III	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1
	Pst I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Sal I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Xho I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C_4	Hae III	1	1	-	1	1	1	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Bam HI	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Dra I[7]	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1
	Eco RI[5]	2	1	2	2	1	1	2	2	1	1	1	1	2	2	2	2	2	1	2	2	1	1	1
	Hind III	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1
	Pst I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Sal I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Xho I	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	-	1	1	1	1
C_6	Hae III	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Bam HI	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Dra I[3]	2	1	2	2	1	1	2	2	1	1	1	1	2	2	2	2	2	1	2	2	1	1	1
	Eco RI	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Hind III	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Pst I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Sal I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Xho I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C_8	Hae III	1	1	-	1	1	1	-	1	1	1	1	1	1	1	1	1	1	1	1	1	-	1	1
	Bam HI	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Dra I[7]	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1
	Eco RI	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Hind III	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Pst I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	1	1	1	1
	Sal I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Xho I	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Numbers indicate the presence of either pattern #1 or #2 for each probe x enzyme combination (see Table 1.2).

a A, M, and H = Andean, Mesoamerican, and heterozygous isozyme genotypes, respectively (Sprecher 1988).

b Me = Mecosta, S = Sanilac, T = Tendergreen.

c superscripts indicate the RFLP number from Table 1.2.

- not determined.

40g/100seeds), have T or C phaseolin, as well as the Andean isozyme genotypes. This shows that the two major gene pools of beans, the Mesoamerican small-seeded and the Andean large-seeded, have differentiated mitochondrial genomes as defined by five RFLPs.

Two of the Malawi lines, 4-10 and 5-25, considered as products of recent gene pool recombination based on isozyme heterozygosity (Sprecher 1988), showed patterns #1 and #2, respectively, for all five RFLPs. The mtDNAs of the three control cultivars fit the gene pool affiliation seen among Malawian lines only with respect to RFLP 1. Sanilac is a small-seeded Mesoamerican cultivar with pattern #1, while Mecosta and Tendergreen are large-seeded Andean types, with pattern #2. However, for each of the other four RFLPs, all three cultivars exhibited pattern #1, associated with small seed size and Mesoamerican origin.

Nature of the mutations giving rise to the RFLPs

RFLPs 1 and 2 were detected only in the stained gels of the DraI digests (Figure 1.1A). Since they did not hybridize to any of the probes, we do not know whether they are related to each other or what kind of mutation event caused them.

Both clones C3 and C6 hybridized to a DraI polymorphism, RFLP 3, consisting of a slight difference in migration of one fragment. This RFLP is likely to be due to a small insertion/deletion event on the order of 100 bp (Figure 1.1). Such an event could possibly go undetected in digest patterns from other enzymes, if part of the larger fragments.

RFLP 4 appeared in the EcoRI digest as a difference in migration of the largest fragment that could be interpreted as a large insertion or deletion event (Figure 1.2A). Furthermore, when the filter was hybridized with clone C2 (Figure 1.2B), a third polymorphic fragment of 4.3 kb appeared in the lanes of Mecosta and line 5-18, which, if added to the 18.6 kb fragment in those lanes, sums approximately to the 22.3 kb fragment present in the other lanes (lines 5-6, 5-7, 5-10, 5-16). This RFLP is thus most

probably due to a basepair mutation in an EcoRI recognition site within the 22.3 kb sequence.

RFLP 5 was revealed by the hybridization of clone C4 to a 6.2 and an 11.2 kb EcoRI fragment in patterns #1 and 2, respectively (Figure 1.2C). These fragment differences were not detectable in the stained gel (Figure 1.2A). The polymorphism is probably not due to a large insertion/deletion event, or it would have been evident in DNA cut with the other enzymes, when hybridized with the same probe. It is possible that this RFLP also results from a base mutation in an EcoRI site, whereby the 11.2 kb fragment would give a 6.2 and a 5.0 kb fragment and clone C4 does not extend into the 5.0 kb section, thus not hybridizing to it.

We believe that RFLPs 6 and 7, which differentiate Mecosta from the rest of the lines, are due to a single mutation event (Figure 1.3). In RFLP 6, Mecosta lacks a 3.3 kb fragment and in RFLP 7 (most probably due to a base mutation), it lacks both a 3.3 and a 4.5 kb fragment apparent in the other lines, but it shows a unique 8.5 kb band. The 3.3 kb fragment seen with probe C2 (RFLP 6) appears to be the same as that detected by C4 and C8 (RFLP 7). If the sequences contained in clone C2 do not extend into the area of the 4.5 kb fragment, C2 should still hybridize to an 8.5 kb band in Mecosta. As shown in Figure 1.3B, all lines hybridize to a fragment of that size, but this common fragment may mask the polymorphic one.

DISCUSSION

Low levels of bean mtDNA variation

These analyses, using eight restriction endonucleases to digest mtDNA and cosmid probes covering about 200 kb of random mtDNA sequences, showed a high level of mitochondrial genome homogeneity among all 23 lines tested, in spite of their variability for seed types and sizes, growth habits and other morpho-agronomic

characters, as well as seed storage protein patterns and isozymes (Table 1.1, Martin 1984, Sprecher 1988). Similar low levels of mitochondrial intraspecific variability have been reported in soybean (Sisson et al. 1978, Grabau et al. 1989), barley (Holwerda et al. 1986), tomato (McClean and Hanson 1986), oat (Rines et al. 1988), pearl millet (Chowdhury and Smith 1988), Beta vulgaris (Ecke and Michaelis 1990) and eight Brassica species (Palmer 1988).

The kind of mutations revealed in this study contrast with the findings of Palmer (1988) who investigated the level of intraspecific mtDNA variation in eight Brassica species. The limited variation he found among 2-5 lines per species consisted entirely of structural changes, two deletions and an inversion. Out of 140 mapped restriction sites (or 840 bp), no basepair mutation was detected. In our case, out of 247 restriction sites (excluding HaeIII fragments) representing 1482 bp, at least two sites have undergone a base mutation (RFLPs 4 and 7). This gives an estimate of 0.14% nucleotide substitution within P. vulgaris. Holwerda et al. (1986) and McClean and Hanson (1986) reported estimates of nucleotide divergence of 0.098% in barley and 0.37% in tomato, respectively. These estimates were determined using the shared fragment method that assumes that all fragment changes are due to base mutations, however, in both cases this assumption was not met.

Apart from sequence comparisons, the most accurate method to study nucleotide divergence is that followed by Palmer (1988). By mapping all restriction sites, he was able to determine the exact nature of every polymorphism. The larger size of the P. vulgaris mitochondrial genome (probably double that of the Brassicas) makes this approach difficult in beans. It may also explain the higher estimate of nucleotide divergence since the larger bean mitochondrial genome must contain more non-coding sequences than that of the Brassica species. Non-coding sequences are in general more prone to mutations, they also may be richer in AT content relative to coding sequences as is the case in the yeast mitochondrial genome (Gray 1982). This may be relevant to

our findings since only two out of the eight enzymes revealed polymorphisms, and these two are EcoRI and DraI which have a stretch of four and six ATs in their respective recognition sequences. The HindIII recognition site also contains four ATs but these are divided by a GC pair in the middle (see Table 1.2).

Whether structural rearrangements play a more important role than sequence divergence in the evolution of the mitochondrial genomes of plants (Palmer and Herbon 1988), cannot be confirmed from this study. Without mapping, the hybridization analysis cannot reveal possible inversions, and the clones we used did not cover all of the genome. Considering the bean mt genome to be close to 450 kb in size (Chapter 3), and since the five cosmid clones have covered about 200 kb, we expect that we have explored about 44% of the genome. This proportion should be representative of the whole genome, since the cosmid clones were randomly selected from a bean mtDNA library (CD Chase, personal communication).

Comparison of Malawian bean mtDNA variation and genetic diversity determined by other methods

The considerable homogeneity of the mitochondrial genome of the Malawian bean lines used in this study contrasts with the appreciable variability observed in the analysis of phenological, morphological and agronomic traits (Martin and Adams 1987a) but not with the paucity of variation at six isozyme loci (Sprecher 1988). Out of the 375 lines used in both previous studies, the twenty analysed here were from two farm sites, and included related as well as distant lines, as determined by a principal component analysis (PCA) of 21 quantitative traits. The variability in mtDNA did not correlate with the PCA distances between the bean lines and did not effectively differentiate them based on seed types per se, but rather on the basis of their gene pool affiliation, the small-seeded Mesoamerican, and the large-seeded Andean. Similarly, isozyme analysis characterized two major gene pools by the presence of contrasting alleles at six isozyme loci (Sprecher 1988). Both studies suggest that divergence

between the gene pools preceded domestication. This hypothesis has been supported by investigations of phaseolin variability (Gepts et al. 1986) and more recently of allozyme diversity (Koenig and Gepts 1989) among wild bean populations. In both cases, the wild beans of Mesoamerica were distinguished from those of the southern Andes by their different alleles and/or allelic frequencies at the protein loci investigated.

The value of examining diversity of beans from Malawi comes from the fact that lines from both gene pools have been co-cultivated for at least three centuries. The lack of inter-gene pool recombination in the mt genome can be explained by its strict maternal inheritance. Two of the lines used in this study were isozyme heterozygotes, i.e. recent gene pool recombinants (Sprecher 1988). Lines 4-10 (phaseolin S) and 5-25 (phaseolin T) exhibited, respectively, the Mesoamerican pattern #1 and the Andean #2 for all five mtDNA RFLPs. This suggests that in the original hybridizations, the female parents were small-seeded and large-seeded, respectively, and it demonstrates the usefulness of mitochondrial RFLPs in tracing maternal lineages.

In the same context, the control cultivars Mecosta and Tendergreen, both large-seeded beans, unexpectedly followed the Mesoamerican pattern #1 for four out of five of the inter-gene pool RFLPs. This strongly indicates that they must have had a small-seeded maternal parent in their pedigree. Mecosta does indeed have a Mesoamerican Great Northern (GN) bean in its background, [(GN x Commercial light red kidney) x Charlevoix kidney], however, the direction of the first cross is unknown. Moreover, Mecosta exhibited the S phaseolin pattern typical of small-seeded beans (Brown et al. 1981). The snap bean cultivar Tendergreen was developed in the mid 1940's, and although we do not know its entire pedigree, the data suggest that a small-seeded line was used as a maternal parent in its development.

Our interpretation of why both Mecosta and Tendergreen showed the large-seeded pattern for RFLP 1 is that the mutation that gave rise to pattern #1 for RFLP 1

occurred in the small-seeded gene pool at a later stage of domestication than the other mutations seen here. Strict maternal inheritance of mitochondria would rule out the possibility of a recombination event. Thus, this RFLP does not truly separate the gene pools.

Our results consolidate the concept of separate and presumably independent domestication events in P. vulgaris, leading to two major gene pools as previously defined by morphological, adaptive, and biochemical traits. This study contributes evidence that the cytoplasms of the gene pools are also differentiated. This has important implications for the taxonomy and genetics of the species and the genus, and on efforts to recombine the two gene pools. Nuclear-cytoplasmic incompatibilities may be part of the reason for the presence of male sterility among gene pool recombinants observed by Sprecher (1988) in the Malawian lines. Such incompatibilities may contribute, in addition to the action of the nuclear Dosage Lethal (DL_1 and DL_2) complementary genes (Shii et al. 1980), to the F_1 and F_2 hybrid weakness symptoms resulting from inter-gene pool crosses and observed by many bean breeders. Knowledge of the exact kind and location of the mitochondrial genome mutations differentiating the gene pools, and a better understanding of nuclear-cytoplasmic interactions in P. vulgaris will help resolve these questions.

LIST OF REFERENCES

Adams MW (1983) Genetic, agronomic and socio-cultural analysis of diversity among bean landraces in Malawi. The Bean/Cowpea Collaborative Research Support Program Annual Report pp. 91-99

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

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

Chowdhury MKU, Smith RL (1988) Mitochondrial DNA variation in pearl millet and related species. Theor Appl Genet 76: 25-32

Debouck DG, Tohme J (1989) Implications for bean breeders of studies on the origins of common beans, Phaseolus vulgaris L. In: Beebe S (ed) Current topics in breeding of common bean. Working Document No. 47. Bean Program, CIAT; Cali, Colombia pp 3-42

Ecke W, Michaelis G (1990) Comparison of chloroplast and mitochondrial DNA from five morphologically distinct Beta vulgaris cultivars: sugar beet, fodder beet, beet root, foliage beet, and Swiss chard. Theor Appl Genet 79:440-442

Evans AM (1976) Beans. In: Simmonds NW (ed) Evolution of crop plants. Longman, London, pp 168-172

Gepts P (1988a) Phaseolin as an evolutionary marker. In: Gepts P (ed) Genetic resources of Phaseolus beans. Kluwer Acad Publ Dordrecht, pp 215-241

Gepts P (1988b) A middle American and an Andean common bean gene pool. In: Gepts P (ed) Genetic resources of Phaseolus beans. Kluwer Acad Publ, Dordrecht, pp 375-390

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

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

Ghaderi A, Adams MW, Saettler AW (1982) Environmental response patterns in commercial classes of common bean (Phaseolus vulgaris L.). Theor Appl Genet 63:17-22

Grabau EA, Davies WH, Gengenbach BG (1989) Restriction fragment length polymorphism in a subclass of the 'Mandarin' soybean cytoplasm. Crop Sci 29:1554-1559

Gray MW (1982) Mitochondrial genome diversity and the evolution of mitochondrial DNA. Can J Biochem 60:157-171

Holwerda BC, Jana S, Crosby WL (1986) Chloroplast and mitochondrial DNA variation in Hordeum vulgare and Hordeum spontaneum. Genetics 114:1271-1291

Kelly JD, Adams MW, Varner GV (1987) Yield stability of determinate and indeterminate dry bean cultivars. Theor Appl Genet 74:516-521

Kemble RJ, Gunn RE, Flavell RB (1983) Mitochondrial DNA variation in races of maize indigenous to Mexico. Theor Appl Genet 65:129-144

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

Mackenzie SA, Pring DR, Bassett MJ, Chase CD (1988) Mitochondrial DNA rearrangement associated with fertility restoration and cytoplasmic reversion to fertility in cytoplasmic male sterile Phaseolus vulgaris L. Proc Natl Acad Sci USA 85:2714-2717

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, New York, 545 pages

Martin GB (1984) Genetic diversity of bean landraces in Northern Malawi. MS Thesis, Michigan State University, East Lansing, MI. 161 pages

Martin GB, Adams MW (1987a) Landraces of Phaseolus vulgaris (Fabaceae) in Northern Malawi. I. Regional variation. Econ Bot 41:190-203

Martin GB, Adams MW (1987b) Landraces of Phaseolus vulgaris (Fabaceae) in Northern Malawi. II. Generation and maintenance of variability. Econ Bot 41:204-215

McClean PE, Hanson MR (1986) Mitochondrial DNA sequence divergence among Lycopersicon and related Solanum species. Genetics 112:649-667

Nugent JM, Palmer JD (1988) Location, identity, amount and serial entry of chloroplast DNA sequences in crucifer mitochondrial DNAs. Curr Genet 14:501-509

Palmer JD (1988) Intraspecific variation and multicircularity in Brassica mitochondrial DNAs. Genetics 118:341-351

Palmer JD, Herbon LA (1988) Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. J Mol Evol 28:87-97

Rigby PW, Dieckmann M, Rhodes C (1977) Labeling deoxyribonucleic acid to high specific activity in vitro by nick-translation with DNA-Polymerase I. J Mol Biol 113:237-251

Rines HW, Gengenbach BG, Boylan KL, Storey KK (1988) Mitochondrial DNA diversity in oat cultivars and species. Crop Sci 28:171-176

Shii CT, Mok MC, Temple SR, Mok DW (1980) Expression of developmental abnormalities in hybrids of Phaseolus vulgaris L.: Interaction between temperature and allelic dosage. J Hered 71:219-222

Singh SP, Gutierrez JA (1984) Geographical distribution of the DL_1 and DL_2 genes causing hybrid dwarfism in Phaseolus vulgaris L., their association with seed size, and their significance to breeding. Euphytica 33:337-345

Sisson VA, Brim CA, Levings CS III (1978) Characterization of cytoplasmic diversity in soybeans by restriction endonuclease activity. Crop Sci 18:991-996

Sprecher SL (1988) Allozyme differentiation between gene pools in common bean (Phaseolus vulgaris L.), with special reference to Malawian germplasm. PhD Dissertation, Michigan State University, East Lansing, MI. 207 pages

Timothy DH, Levings CS III, Pring DR, Conde MF, Kermicle JL (1979) Organelle DNA variation and systematic relationships in the genus Zea:Teosinte. Proc Natl Acad Sci USA 76:4220-4224

Weissinger AK, Timothy DH, Levings CS III, Goodman MM (1983) Patterns of mitochondrial DNA variation in indigenous maize races of Latin America. Genetics 104:365-379

CHAPTER TWO

MITOCHONDRIAL RESTRICTION FRAGMENT LENGTH POLYMORPHISMS IN WILD PHASEOLUS VULGARIS L.

ABSTRACT. Previous examination of intraspecific mitochondrial DNA (mtDNA) diversity in common bean, Phaseolus vulgaris, showed that five restriction fragment length polymorphisms (RFLPs) distinguish the mitochondrial genomes of the two major gene pools, the Mesoamerican and the Andean (Chapter 1). In this study, mtDNA was used to compare the amount of diversity in the wild beans to that in the cultivated, and to understand how and when the mitochondrial genomes of the gene pools became distinct. The mtDNA of six wild bean accessions from Central and South America were digested with nine restriction endonucleases and analysed by Southern hybridization. A total of twenty RFLPs were detected demonstrating a significantly higher amount of mtDNA variability in wild than in cultivated beans. All of the wild beans had the same mtDNA pattern for four out of the five inter-gene pool RFLPs, indicating that the polymorphisms arose soon after domestication, two in the gene pool of the cultivated Mesoamerican beans and two in the gene pool of the cultivated Andean beans. The fifth RFLP must have occurred before domestication since the locus was also polymorphic in the wild beans. Wild beans from the south Andes were distinct and less variable than wild accessions of the north Andes and Mesoamerica. The distribution of mtDNA RFLPs among the wild beans strongly supports the concept of two distinct domestication events for P. vulgaris.

INTRODUCTION

In recent years, restriction fragment length polymorphisms (RFLPs) have been used increasingly as molecular markers to construct chromosome maps of crop plants, fingerprint crop cultivars, and mark quantitative trait loci (Beckmann and Soller 1986). In addition, RFLPs are useful in understanding genome evolution, studying plant origins and domestication, and probing genetic diversity.

Organellar DNA RFLPs allow maternal lineages to be followed, since in most plant species, chloroplasts and/or mitochondria are inherited exclusively from the maternal parent (e.g. Palmer et al. 1983). Moreover, plant organelle RFLPs can provide an indication of the cytoplasmic diversity of crops, an important issue if genetic vulnerability, due in part to uniform cytoplasms, is to be avoided. Primarily for this reason, several studies have investigated mitochondrial DNA (mtDNA) variability within various crop species such as soybean (Sisson et al. 1978, Grabau et al. 1989), barley (Holwerda et al. 1986) and oat (Rines et al. 1988). With the exception of annual teosinte (Timothy et al. 1979), maize (Kemble et al. 1983, Weissinger et al. 1983), and carrot (Ichikawa et al. 1989), high degrees of uniformity of plant mitochondrial genomes were revealed in these intraspecific comparisons.

A similar low level of intraspecific mtDNA diversity occurs in the common bean, Phaseolus vulgaris L. (Chapter 1). Nonetheless, five RFLPs allowed the 23 bean lines studied to be differentiated into the two major gene pools of beans. These are recognized as the small-seeded Mesoamerican and the large-seeded Andean gene pools. In addition to the differences in their seed sizes and centers of origin, the two gene pools are distinguished by differences in growth habits, environmental adaptation, disease resistance, zymograms and phaseolin types (reviewed in Gepts 1988b). In order to understand how and when the mitochondrial genomes of the two gene pools became distinct, we decided to examine the mtDNA of extant wild bean populations from both centers of diversity.

As shown in Figure 2.1, wild forms of P. vulgaris still exist in an almost continuous band extending from the state of Chihuahua, Mexico, in the north to the province of San Luis, Argentina, in South America (Debouck and Tohme 1989). These wild types cross freely with cultivated beans and are classified as botanical varieties of P. vulgaris (Baudet 1977, Delgado-Salinas 1985). Morphological and phenological variability exists among the populations in the different geographical areas for such characters as seed size, hypocotyl texture, bracteole size and shape, days to, and duration of flowering (Vanderborght 1983, Debouck and Tohme 1989). In addition, genetic variation has been observed for physiological characters (Lynch et al. 1989) and biochemical attributes (Gepts et al. 1986, Koenig and Gepts 1989, Koenig et al. 1990).

Although the variability appears to be gradual along the range of distribution, the Mexican and the Argentinian forms are the extremes and are quite distinct. Vanderborght (1983) reported an average seed size of 6.2 and 13.9 g/100 seeds for the two forms, respectively. Seed size (Gentry 1969, Evans 1973, Baudet 1977, Gepts et al. 1986) and phaseolin type (Gepts and Bliss 1986, Gepts et al. 1986) were found to vary among wild beans and among domesticated beans in a parallel fashion along the geographic range. This has been interpreted to mean that multiple domestication events have occurred (Gepts 1988a).

The main objectives of this study were to examine the mtDNA variability in a sample of wild beans by restriction and hybridization analyses, in order to compare the mtDNA diversity of wild and cultivated lines. It was also intended as a preliminary study to identify new RFLPs that can be used to examine a larger number of wild bean accessions for a better understanding of the origin, evolution and domestication of the common bean.

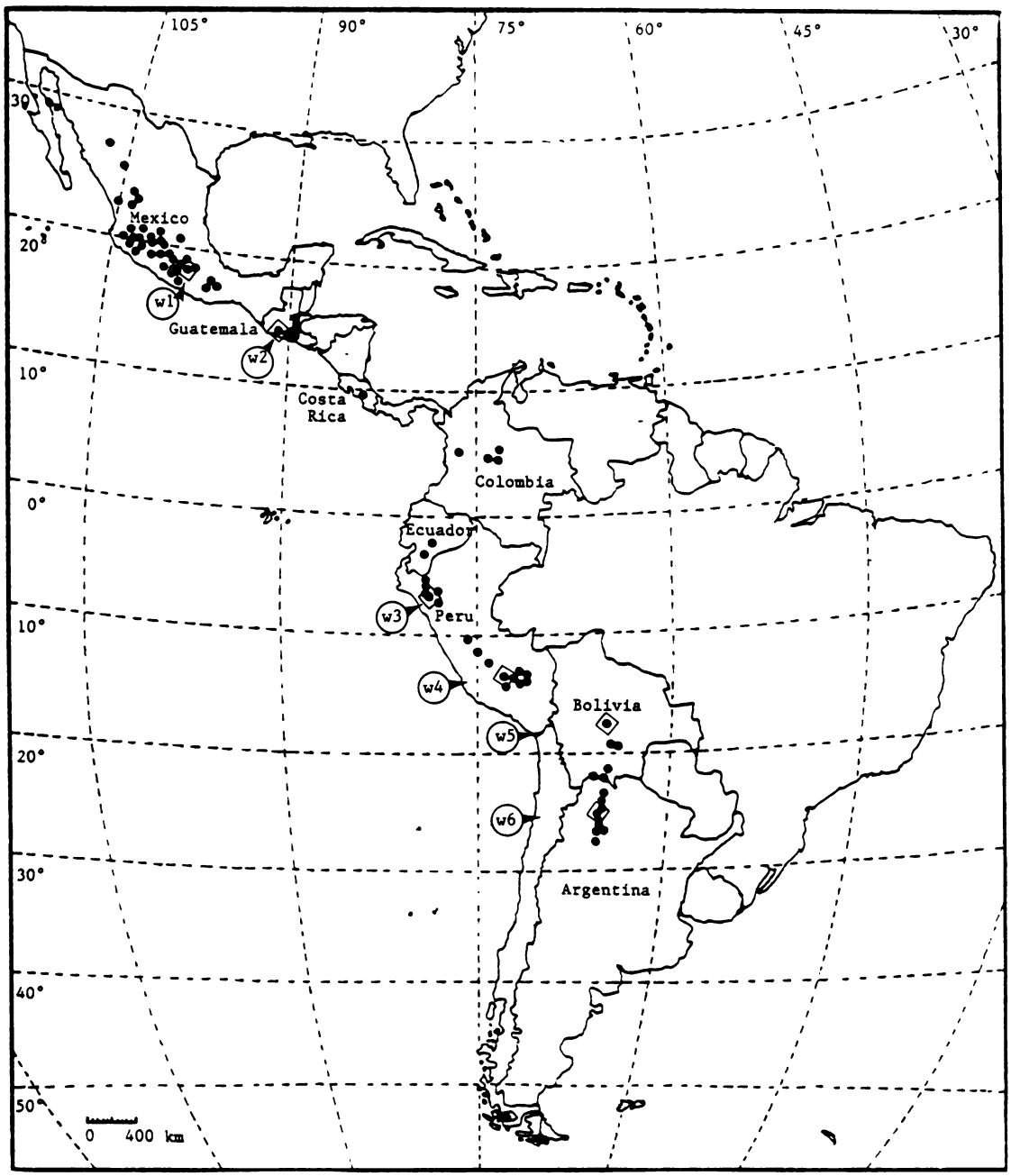

Figure 2.1. Known distribution of wild Phaseolus vulgaris in North and South America (Torro et al. 1990), and collection site of the six accessions (w1 to w6) used in this study.

MATERIALS AND METHODS

Bean lines

Six wild bean accessions sampled across the range of distribution of wild Phaseolus vulgaris were used in this study (Table 2.1). Their exact collection locations are indicated in Figure 2.1. Seed samples were obtained from Drs. D. Debouck and J. Tohme at the Centro Internacional de Agricultura Tropical (CIAT), Colombia. Based on seed availability and amounts, Malawian bean lines 4-11 and 4-22 were used as the representatives of the cultivated Mesoamerican and Andean controls, respectively (Chapter 1).

Seeds of the wild bean accessions were scarified prior to germination in vermiculite trays in darkness.

MtDNA RFLP analysis

Seedlings that had been grown in the dark for 10 to 15 days were used to extract mtDNA. The procedures and conditions for mtDNA isolation, digestion and electrophoresis, Southern blotting, nick translation and hybridization were described in Chapter 1. Nine restriction endonucleases were utilized to digest the mtDNAs: AsnI, HindIII, PstI (Boehringer Mannheim Biochemicals), BamHI, DraI, EcoRI, EcoRV, SalI (Bethesda Research Lab, Inc.) and XhoI (New England Biolabs).

Southern blots were consecutively hybridized to five cosmid clones and strip-washed after each hybridization according to Gatti et al. (1984). The clones were provided by Dr. C. D. Chase (University of Florida, Gainesville). They contained random inserts from the bean mitochondrial genome ranging in size from 29 to 38 kb (Chapter 1).

Table 2.1. Origin and some characteristics of the wild bean accessions used in this study.

Line	CIAT Number	Collection Number	Country, State	Seed Size gms/100 seeds	Phase-olin
w1	G12878	Gentry 22535	Mexico, Guerrero	8.5	M1
w2	G19906	DGD 1610	Guatemala, Sacatepequez	5.7	S,Sd
w3	G21245	DGD 1962	Peru, Cajamarca	8.8	I
w4	G23455	DGD 2581	Peru, Cuzco	9.0	C,H
w5	G23442	DGD 2484	Bolivia, Cochabamba	8.7	To
w6	G19892	DGD 629	Argentina, Salta	7.0	T

a Phaseolin types were determined at the Centro Internacional de Agricultura Tropical (CIAT).
Types 'S', 'T', 'C' were first described in Brown et al. (1981); 'M1' in Romero-Andreas and Bliss (1985); 'H' in Gepts et al. (1986); 'I' and 'Sd' in Koenig et al. (1990) and 'To'in Vargas et al. (1990).

RESULTS

Mitochondrial DNA from six wild P. vulgaris accessions and two cultivated lines representing the Mesoamerican and the Andean types, were digested with nine restriction endonucleases and the Southern blots were hybridized to five cosmid clones. Seven of the endonucleases were used previously to probe mtDNA diversity among cultivated bean lines but only two of these, EcoRI and DraI, revealed RFLPs (Chapter 1). These two enzymes differ from the other five in that they contain a minimum of four adjacent AT pairs in their recognition sequences, GAATTC and TTTAAA, respectively. Based on this observation, two other enzymes having this trait, EcoRV (GATATC) and AsnI (ATTAAT), were included in our analysis of wild bean mtDNA diversity.

The restriction fragment profiles for all the enzymes were complex, and relatively uniform across the eight lines tested, as expected for an intraspecific comparison. However, a few more polymorphic bands were observed in this analysis as compared with mtDNA isolated from cultivated material (Chapter 1). Differences in ethidium bromide-stained gels were observed repeatedly in the case of DraI (Figure 2.2A) but were not as clear with the other enzymes and thus are not considered in this analysis. Hybridization experiments provided more conclusive results and the RFLPs thus detected are listed in Table 2.2. Six enzymes out of nine revealed 20 polymorphisms although DraI, EcoRI, EcoRV and AsnI produced most of the variation (Figures 2.2 to 2.5).

Figure 2.2 shows the six RFLPs detected with DraI. The cultivated Mesoamerican line showed an extra 12.7 kb fragment (RFLP # 1), while line w2 showed two extra fragments, a 9.0 kb (RFLP # 2) and a 6.2 kb also seen with clone C8 (RFLP # 6). Line w3 and the cultivated Andean line had, respectively, a unique band of 4.9 kb (RFLP # 3) and 3.7 kb (RFLP # 4). RFLP # 5 was detected with probes C3 and C6, and is most probably due to a small length mutation on the order of 80

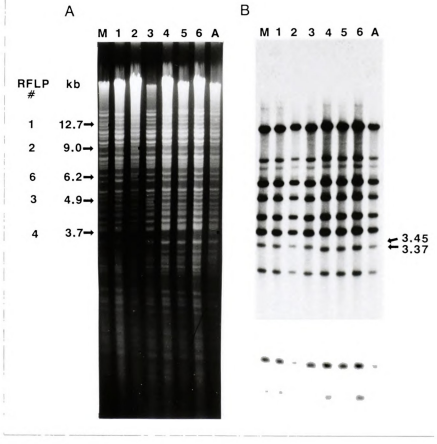

Figure 2.2. DraI mtDNA digestion patterns of six wild and two cultivated bean accessions. M,A = Mesoamerican and Andean controls, respectively. Numbers 1 to 6 above the gel lanes indicate the wild bean accessions as in Table 2.1. (A) Ethidium bromide stained gel, arrows indicating polymorphic fragments, RFLP numbers and sizes are given at the left of the gel. (B) Southern blot hybridized with clone C3 showing RFLP # 5, with sizes indicated at right.

Table 2.2. RFLPs detected in wild and cultivated bean mtDNA.

RFLP Number	Enzyme	Revealed by	Definition of Pattern # 1	Definition of Pattern # 2	Number of Polymorphic Fragments
1	Dra I	Gel	12.7kb	no 12.7kb	1
2	Dra I	Gel	no 9.0kb	9.0kb	1
3	Dra I	Gel	no 4.9kb	4.9kb	1
4	Dra I	Gel	no 3.7kb	3.7kb	1
5	Dra I	C3/C6 *	3.45kb	3.37kb	2
6	Dra I	C8 *	no 6.2kb	6.2kb	1
7	EcoR I	C2	18.6 + 4.3kb	22.3kb	3
8	EcoR I	C2	1.6kb	1.2 + 0.7kb	3
9	EcoR I	C4	6.2kb	11.2kb	2
10	EcoR V	C3	18.3 + 12.7kb	no 18.3, 12.7kb **	2
11	EcoR V	C3	no 6.6kb	6.6kb	1
12	EcoR V	C8/C4	7.2 + 0.7kb #	6.7 + 1.1kb #	4
13	Asn I	C2	4.9kb	11.8kb	2
14	Asn I	C3	24.9 + 17.3kb	21.4kb ##	3
15	Asn I	C3/C6	3.4 kb	6.4 kb	2
16	Asn I	C6	10.4 + 4.5kb	no 10.4, 4.5kb	2
17	Asn I	C8/C4	10.7 + 1.5kb	12.7kb	3
18	Asn I	C8	17.3kb	23.6kb	2
19	Hind III	C4	2.3 + 1.5kb	3.8kb	3
20	Xho I	C6	8.8kb	no 8.8kb	1

a The RFLP pattern numbers were assigned so that the cultivated Mesoamerican control always displayed pattern # 1.

* Also detected in the stained gel.

** Line w6 showed a 29.1 and a 23.9kb fragment that are possible results of partial digestion.

\# Only fragments 7.2 and 6.7kb hybridized to C8, whereas all all four fragments hybridized to C4.

A 27.9kb fragment seen in line w3 has not been included because it was probably due to partial digestion.

basepairs since the cultivated Mesoamerican line and lines w1, w2 and w3 had a 3.45 kb band while the cultivated Andean line and lines w4, w5 and w6 had a 3.37 kb fragment.

Figure 2.3 shows the two RFLPs seen in the Southern blot of the EcoRI digest hybridized to clone C2. In both cases, the three polymorphic bands may result from a basepair mutation in an EcoRI recognition site. As indicated in the figure for both RFLP # 7 and 8, the total of two fragments present in pattern # 1 is approximately equal to the size of the third fragment present in pattern # 2. While RFLP # 7 differentiates the cultivated Andean line from all others, RFLP # 8 differentiates line w3 from all others.

The three EcoRV RFLPs are shown in Figure 2.4. RFLP # 10 is seen with clone C3 as the absence of two large fragments (18.3 and 12.7 kb) in lines w1 and w6, while RFLP # 11 is an extra 6.6 kb band present only in line w1 (Figure 2.4A). The hybridization of the EcoRV blot to C8 revealed RFLP # 12 that differentiates line w1 (6.7 kb) from all the other lines (7.2 kb) (Figure 2.4B). Clone C4 hybridizes to the same two polymorphic bands and two others, 1.1 kb in w1 and 0.7 kb in all other lines (blot not shown). These four polymorphic fragments defining RFLP # 12 could be due to an inversion because 6.7 + 1.1 kb are approximately equal to 7.2 + 0.7 kb.

The six RFLPs detected in the AsnI digest are shown in Figure 2.5. Hybridization to clone C2 shows the absence of an 11.8 kb band and the presence of a 4.9 kb band in the cultivated Mesoamerican and w2 lines (RFLP # 13, Figure 2.5A). RFLP # 14 (clone C3) differentiates line w3 from all others by the absence of two bands (24.9 and 17.3 kb) and the presence of a 21.4 kb fragment (Figure 2.5B). On the other hand, RFLP # 15, seen with both clones C3 and C6, separates line w1 from all others by the presence of a 6.4 kb band instead of a 3.4 kb fragment (Figure 2.5B,C). RFLPs # 16, 17 and 18 again distinguish line w3 from all other lines. For RFLP # 16, which is detected with clone C6 (Figure 2.5C), w3 is missing two fragments (10.4 and

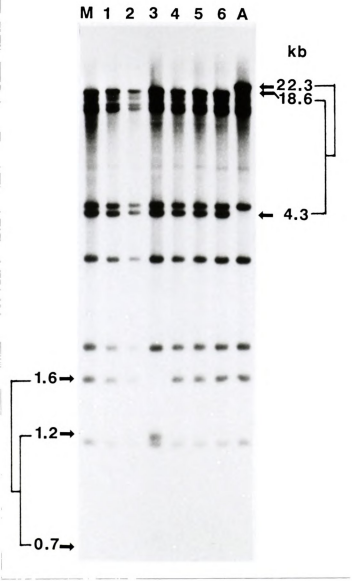

Figure 2.3. Hybridization of EcoRI digested mtDNAs from cultivated and wild beans to clone C2. M, A, numbers 1 to 6 as in Figure 2.2. Arrows on the right show RFLP # 7 and on the left RFLP # 8 (Table 2.2).

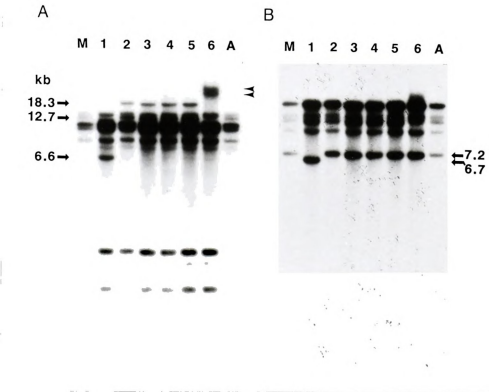

Figure 2.4. Hybridization of EcoRV digested mtDNAs from cultivated and wild beans to clones C3 and C8. M, A, numbers 1 to 6 as in Figure 2.2. (A) Probe C3, the upper two arrows on the left point to the fragments missing in accessions 1 and 6 (RFLP # 10), the third arrow indicates RFLP # 11. The arrows on the right point to mtDNA fragments that probably resulted from partial digestion in line 6. (B) Probe C8, polymorphic bands are part of RFLP # 12 (see Table 2.2).

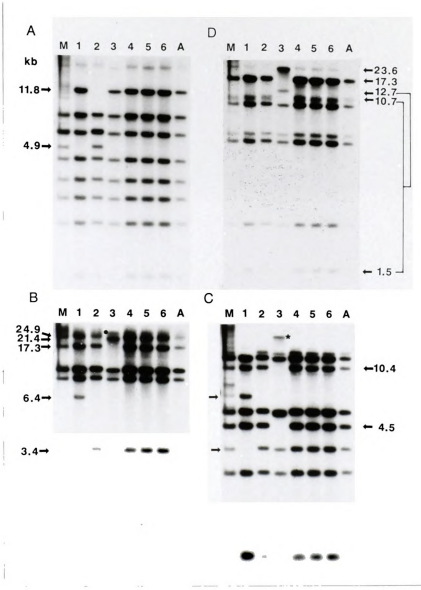

Figure 2.5. Hybridization of AsnI digested mtDNAs from cultivated and wild beans to clones C2, C3, C6 and C8. M, A, numbers 1 to 6 as in Figure 2.2. (A) Probe C2, RFLP # 13. (B) Probe C3, the upper three arrows point to RFLP # 14, the lower two to RFLP # 15 also seen with clone C6. The dot shows a probable partially digested band in line 3. (C) Probe C6, arrows on the left show RFLP # 15 and on the right RFLP # 16. The "*" shows a fragment remaining from the previous hybridization with clone C8. (D) Probe C8, the upper two arrows point to RFLP # 18 and the lower three to RFLP # 17 shared with clone C4.

4.5 kb). For RFLP # 17 seen with clone C8 (Figure 2.5D), w3 lacks a 10.7 and a 1.5 kb fragment but has a 12.7 kb band which seem to indicate a basepair mutation event. Finally, for RFLP # 18 (C8) it shows a 23.6 kb but lacks a 17.3 kb band (Figure 2.5D).

Although our interpretation of some RFLPs as restriction mutations or inversions has already been mentioned, the nature of the mutations giving rise to some RFLPs was ambiguous and difficult to decifer. The absence of large fragments in some lines (e.g. RFLPs # 6, 10, 11, 20) or a large difference in fragment sizes in the two patterns (e.g. RFLPs # 9, 13, 14, 15, 18) are not likely to be due to large length mutations, since these differences should have then been detected with other enzymes using the same probe. If, as in other plant species, the bean mitochondrial genome has a set of small repeat elements, the deletion of one of those may prevent hybridization of a fragment to a clone that has sequences similar to the repeat rather than to the fragment itself. Alternatively, these RFLPs may be due to basepair mutations or inversions, which are not readily observed because the probe(s) may not extend into the other polymorphic fragments.

The results of all the hybridizations are summarized in Table 2.3. The scores (pattern #1 or #2) of the eight bean lines with respect to the polymorphic enzyme x probe combinations are reorganized in Table 2.4 in order to show some of the patterns of variation. The first six RFLPs (# 1, 9, 4, 7, 5 and 13) differentiated the gene pools of the Mesoamerican and Andean cultivated beans. The other 14 RFLPs distinguish one or two of the wild bean lines from all of the others. This is a good indication of the higher variability of mtDNA in the wild material.

Accession w3 was quite distinct from all others, as exhibited by seven unique polymorphisms, RFLPs # 3, 8, 14, 16, 17, 18 and 20. It is followed by the Mexican accession w1, with three unique polymorphisms (RFLPs # 11, 12 and 15), then by w2 from Guatemala with two unique RFLPs (# 2, 6) and finally w6 from Argentina with a single unique RFLP, # 19. Accessions w4 and w5 were the most uniform.

Table 2.3. Summary of results when mtDNA from cultivated and wild bean lines were digested with nine restriction endonucleases and probed with five different mtDNA clones (C_2, C_3, C_4, C_6, C_8).

Gel or Probe	Enzyme	RFLP Number[a]	Bean Lines[b]							
			cM	w1	w2	w3	w4	w5	w6	cA
Gel	Dra I	1	1	2	2	2	2	2	2	2
	Dra I	2	1	1	2	1	1	1	1	1
	Dra I	3	1	1	1	2	1	1	1	1
	Dra I	4	1	1	1	1	1	1	1	2
C_2	Asn I	13	1	2	1	2	2	2	2	2
	BamH I		1	1	1	1	1	1	1	1
	Dra I		1	1	1	1	1	1	1	1
	EcoR I	7	1	1	1	1	1	1	1	2
	EcoR I	8	1	1	1	2	1	1	1	1
	EcoR V		1	1	1	1	1	1	1	1
	Hind III		1	1	1	1	1	1	1	1
	Pst I		1	1	1	1	1	1	1	1
	Sal I		1	1	1	1	1	1	1	1
	Xho I		1	1	1	1	1	1	1	1
C_3	Asn I	14	1	1	1	2	1	1	1	1
	Asn I	15	1	2	1	1	1	1	1	1
	BamH I		1	1	1	-	1	1	1	1
	Dra I	5	1	1	1	1	2	2	2	2
	EcoR I		1	1	1	1	1	1	1	1
	EcoR V	10	1	2	1	1	1	1	2	1
	EcoR V	11	1	2	1	1	1	1	1	1
	Hind III		1	1	1	1	1	1	1	1
	Pst I		1	1	1	1	1	1	1	1
	Sal I		1	1	1	-	1	1	1	1
	Xho I		1	1	1	1	1	1	1	1
C_4	Asn I	17	1	1	1	2	1	1	1	1
	BamH I		1	1	1	1	1	1	1	1
	Dra I		1	1	1	1	1	1	1	1
	EcoR I	9	1	2	2	2	2	2	2	2
	EcoR V	12	1	2	1	1	1	1	1	1
	Hind III	19	1	1	1	1	1	1	2	1
	Pst I		1	1	1	1	1	1	1	1
	Sal I		1	1	1	1	1	1	1	1
	Xho I		1	1	1	1	1	1	1	1
C_6	Asn I	14	1	2	1	1	1	1	1	1
	Asn I	16	1	1	1	2	1	1	1	1
	BamH I		1	1	1	-	1	1	1	1
	Dra I	5	1	1	1	1	2	2	2	2
	EcoR I		1	1	1	1	1	1	1	1
	EcoR V		1	1	1	1	1	1	1	1
	Hind III		1	1	1	-	1	1	1	1
	Pst I		1	-	1	-	-	-	1	1
	Sal I		1	1	1	1	1	1	1	1
	Xho I	20	1	1	1	2	1	1	1	1
C_8	Asn I	17	1	1	1	2	1	1	1	1
	Asn I	18	1	1	1	2	1	1	1	1
	BamH I		1	1	1	-	1	1	1	1
	Dra I	6	1	1	2	1	1	1	1	1
	EcoR I		1	1	1	1	1	1	1	1
	EcoR V	12	1	2	1	1	1	1	1	1
	Hind III		1	1	1	-	1	1	1	1
	Pst I		1	-	1	-	-	-	1	1
	Sal I		1	1	1	1	1	1	1	1
	Xho I		1	1	1	1	1	1	1	1

a The RFLP numbers as defined in Table 2.2.

b cM, cA = cultivated Mesoamerican and Andean beans, respectively;
w1 to w6 as in Table 2.1;
Numbers indicate the occurrence of pattern #1 or #2 for each probe x enzyme combination as defined in Table 2.2.

\- Not determined.

Table 2.4. Reorganization of the 20 mtDNA RFLPs detected in the cultivated and wild bean lines.

RFLP Number[a]	Enzyme	Gel or Probe	Bean Lines[b]							
			cM	w1	w2	w3	w4	w5	w6	cA
1	Dra I	Gel	1	2	2	2	2	2	2	2
9	EcoR I	C4	1	2	2	2	2	2	2	2
4	Dra I	Gel	1	1	1	1	1	1	1	2
7	EcoR I	C2	1	1	1	1	1	1	1	2
5	Dra I	C3/C6	1	1	1	1	2	2	2	2
13	Asn I	C2	1	2	1	2	2	2	2	2
10	EcoR V	C3	1	2	1	1	1	1	2	1
2	Dra I	Gel	1	1	2	1	1	1	1	1
3	Dra I	Gel	1	1	1	2	1	1	1	1
6	Dra I	C8	1	1	2	1	1	1	1	1
8	EcoR I	C2	1	1	1	2	1	1	1	1
11	EcoR V	C3	1	2	1	1	1	1	1	1
12	EcoR V	C8	1	2	1	1	1	1	1	1
14	Asn I	C3	1	1	1	2	1	1	1	1
15	Asn I	C3/C6	1	2	1	1	1	1	1	1
16	Asn I	C6	1	1	1	2	1	1	1	1
17	Asn I	C8/C4	1	1	1	2	1	1	1	1
18	Asn I	C8	1	1	1	2	1	1	1	1
19	Hind III	C4	1	1	1	1	1	1	2	1
20	Xho I	C6	1	1	1	2	1	1	1	1

a The RFLP numbers as defined in Table 2.2

b cM, cA = cultivated Mesoamerican and Andean beans, respectively;
w1 to w6 as in Table 2.1;
Numbers indicate the occurrence of pattern #1 or #2 for each probe x enzyme combination as defined in Table 2.2

DISCUSSION

Timing and direction of mtDNA mutations distinguishing the two gene pools of cultivated beans

The examination of 23 cultivated bean lines for intraspecific mtDNA variability revealed very uniform restriction patterns (Chapter 1). However, five RFLPs were identified that divided the bean lines into two groups corresponding to the Mesoamerican and Andean gene pools of P. vulgaris. By examining the mtDNA of extant wild bean populations from across the centers of bean diversity in this study, we sought to understand how and when the cytoplasms of the two gene pools diverged. The first five RFLPs in Table 2.4 are the ones that differentiated the mtDNA of the cultivated bean gene pools. In RFLPs # 1 and 9, all six wild beans are similar to the cultivated Andean beans, while in RFLPs # 4 and 7 they are similar to the cultivated Mesoamerican beans, demonstrating that the mutations occurred in the Mesoamerican gene pool giving rise to the first two RFLPs, and in the Andean gene pool resulting in the other two RFLPs. Although these results indicate that the mutations occurred after domestication, they must have happened early enough such that all examined cultivated beans of an individual gene pool were the same. This also argues against multiple domestication events along the range of the wild beans.

As to the fifth polymorphism (Table 2.4, RFLP # 5), it obviously occurred before domestication since the wild beans themselves exhibit both patterns. Surprisingly, however, accession w3 from northern Peru showed the same pattern (#1) as the Mesoamerican wild (w1, w2) and cultivated beans rather than pattern 2 seen in the other Andean wild (w4,5,6) and cultivated beans. Consistent with this observation, Koenig and Gepts (1989) found that allozyme allelic frequencies in accessions from Colombia and northern Peru are more similar to those of Mexican accessions than to those of populations from southern Peru and Argentina. Based on a survey of

phaseolin types in wild and cultivated beans, Colombia was initially suggested as a minor domestication center of the common bean and a geographical meeting place for the Mesoamerican and Andean cultivated forms (Gepts and Bliss 1986). Koenig and Gepts (1989) later defined the Colombia-Peru region as a geographical transition zone. Analysis of a larger number of wild bean accessions from that region for mtDNA variability may help clarify the validity and/or significance of that "transition zone". On the other hand, Debouck (1986) recognized three centers of diversification of the genus Phaseolus in general, and of P. vulgaris in particular: Mesoamerica, the north Andes (western Venezuela to northern Peru), and the south Andes.

A sixth polymorphism (RFLP # 13), detected with AsnI and clone C2 differentiated the cultivated Mesoamerican and Andean beans used in this study, and was also obvious in the wild beans, indicating that the mutation happened before domestication. In this case, however, only line w2 from Guatemala showed the same pattern as the Mesoamerican control. Moreover, the distribution of the polymorphism within domesticated beans is unknown, since this enzyme was not utilized in the previous study of bean mtDNA diversity (Chapter 1).

More mtDNA variability in the wild

It is obvious from Table 2.4 that the wild beans contain quite a few more mtDNA polymorphisms than do the cultivated forms. This has been shown for chloroplast DNA (cpDNA) in barley (Clegg et al. 1984, Neale et al. 1988), although a study of both cpDNA and mtDNA diversity in wild and cultivated barley produced contradictory results (Holwerda et al. 1986). In P. vulgaris, wider genetic diversity in wild populations, illustrated by phaseolin studies (Gepts and Bliss 1986, Gepts et al. 1986) led Debouck and Tohme (1989) to postulate a founder effect of domestication. The reduction in genetic diversity in the crop has important implications to bean breeders who are in continuous search for useful genes to incorporate into their improved cultivars. Wild relatives have served as sources of various important

characters in breeding programs, the most common being disease and pest resistance (Harlan 1976). The work on bruchid resistance in beans is a good example of the expected consequences of the founder effect of domestication, and of the usefulness of wild germplasm. After an unsuccessful search among thousands of cultivated bean genotypes for bruchid resistance (Schoonhoven and Cardona 1982), good resistance levels were found in a small number of wild bean populations from Mexico (Schoonhoven et al. 1983).

Our results confirm previous studies of phaseolin (Gepts et al. 1986, Koenig et al. 1990) and allozyme (Koenig and Gepts 1989) diversity showing that wild bean accessions from Mesoamerica and Colombia contain higher genetic variability than those from the southern Andes. However, it is still unclear if this indicates a Mesoamerican origin of wild P. vulgaris or whether the higher variation in Mesoamerica is due to conditions eliciting higher levels of genetic diversity.

The significantly higher mtDNA variability of accession w3 is quite peculiar. Koenig and Gepts (1989) have shown with allozyme data that this same accession (DGD 1962) from northern Peru is genetically distinct from 82 other wild bean accessions that clustered into two major groups. Until more accessions from that region are examined, it will not be clear whether this higher variability is characteristic of the area in general, or is unique to that particular accession.

The similarity of mtDNA from the cultivated Andean beans to mtDNA of the south Andean wild accessions, especially w4 and w5, is remarkable. These data suggest that the south Peru - Bolivia area may be the site of initial domestication of the Andean beans. The results also favor at most a subspecific classification of the Andean wild bean as suggested by Burkhart and Brücher (1953), P. vulgaris ssp. aborigineus Burk., or Baudet (1977), P. vulgaris var. aborigineus (Burk.) Baudet, in contrast to placing it as a separate species, P. aborigineus Burk., as was initially done (Burkhart 1952) and as Brücher (1988) currently favors.

On the other hand, the mtDNA patterns of the Mesoamerican cultivated beans are most similar to those of the wild bean accession from Guatemala (Table 2.4), which, in a similar fashion, may point to a specific area of bean domestication in Mesoamerica. A primary center of P. vulgaris origin in the western Mexico-Guatemala area has been suggested (Miranda-Colín 1967), although Gentry (1969) proposed multiple origins of cultivated beans in Mesoamerica based on botanical, genetic and archaeological evidence. Because several RFLPs distinguish the cultivated Mesoamerican beans from the wild accessions, analysis of more plant samples, both wild and cultivated, is essential before similar conclusions about a domestication center can be made.

Finally, it was surprising to see that the mtDNA of the bean accessions from the extremes of the range of distribution of wild P. vulgaris, w1 and w6, shared one polymorphism that differentiated them from all other wild beans (Table 2.4, RFLP # 10). It is not clear, though, whether this RFLP arose twice independently or if it is due to common descent.

In summary, this study allowed the identification of four restriction endonucleases that can detect significantly more mtDNA RFLPs in wild bean germplasm than can other enzymes. The reduced number of enzymes will be useful for a more extensive survey of mtDNA polymorphisms in wild populations especially when seed amounts are limiting. Many questions concerning bean origin and domestication have been raised by these preliminary studies: Will we able to delineate specific areas where domestication occurred? Will a clinal pattern of mtDNA variability be more obvious? Can the wild beans be clustered in meaningful groups based on mtDNA data, and will one group be identified as ancestral? Answers to these questions may be obtained through an extensive survey of wild forms from across the range of distribution of wild P. vulgaris.

LIST OF REFERENCES

Baudet JC (1977) Origine et classification des espèces cultivées du genre Phaseolus. Bull Soc Roy Bot Belg 110:65-76

Beckmann JS, Soller M (1986) Restriction fragment length polymorphisms in plant genetic improvement. Oxford Surveys Plant Mol Cell Biol 3:196-250

Berglund-Brücher O, Brücher H (1976) The south American wild bean (Phaseolus aborigineus Burk.) as ancestor of the common bean. Econ Bot 30:257-272

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

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

Burkhart A (1952) Las leguminosas Argentinas silvestres y cultivadas. 2nd ed, ACME, Buenos Aires, 569 pages

Burkhart A, Brücher H (1953) Phaseolus aborigineus Burkhart, die mutmassliche andine Stammform der Kulturbohne. Der Zuchter 23:65-72

Clegg MT, Brown AHD, Whitfeld PR (1984) Chloroplast DNA diversity in wild and cultivated barley: implications for genetic conservation. Genet Res, Camb 43:339-343

Debouck DG (1986) Primary diversification of Phaseolus in the Americas: three centers? Plant Genet Resources Newsl 67:2-8

Debouck DG, Tohme J (1989) Implications for bean breeders of studies on the origins of common beans, Phaseolus vulgaris L. In: Beebe S (ed) Current topics in breeding of common bean. Working Document No. 47. Bean Program, CIAT; Cali, Colombia pp 3-42

Delgado-Salinas AO (1985) Systematics of the genus Phaseolus (Leguminosae) in North and Central America. PhD Dissertation University of Texas, Austin.

Evans AM (1973) Genetic improvement of Phaseolus vulgaris. In: Milner M (ed) Nutritional improvement of food legumes by breeding. Protein Advisory Group, UN, New York, pp 107-115

Gatti RA, Concannon P, Salser W (1984) Multiple use of Southern blots. Biotechniques 2:

Gentry HS (1969) Origin of the common bean, Phaseolus vulgaris. Econ Bot 23:55-69

Gepts P (1988a) Phaseolin as an evolutionary marker. In: Gepts P (ed) Genetic resources of Phaseolus beans. Kluwer Acad Publ Dordrecht, pp 215-241

Gepts P (1988b) A middle American and an Andean common bean gene pool. In: Gepts P (ed) Genetic resources of Phaseolus beans. Kluwer Acad Publ, Dordrecht, pp 375-390

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

Grabau EA, Davies WH, Gengenbach BG (1989) Restriction fragment length polymorphism in a subclass of the 'Mandarin' soybean cytoplasm. Crop Sci 29:1554-1559

Harlan JR (1976) Genetic resources in wild relatives of crops. Crop Sci 16:329-333

Holwerda BC, Jana S, Crosby WL (1986) Chloroplast and mitochondrial DNA variation in Hordeum vulgare and Hordeum spontaneum. Genetics 114:1271-1291

Ichikawa H, Tanno-Suenaga L, Imamura J (1989) Mitochondrial genome diversity among cultivars of Daucus carota (ssp sativus) and their wild relatives. Theor Appl Genet 77:39-43

Kaplan L (1981) What is the origin of the common bean? Econ Bot 35:240-254

Kemble RJ, Gunn RE, Flavell RB (1983) Mitochondrial DNA variation in races of maize indigenous to Mexico. Theor Appl Genet 65:129-144

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

Koenig RL, Singh SP, Gepts P (1990) Novel phaseolin types in wild and cultivated common bean (Phaseolus vulgaris, Fabaceae). Econ Bot 44:50-60

Lynch J, Gonzalez A, Tohme J (1989) Characters related to leaf photosynthesis in wild bean. Agron Abst 81:116

Miranda-Colín S (1967) Origen de Phaseolus vulgaris L. (frijól comun). Agrociencia 1:99-109

Neale DB, Saghai-Maroof MA, Allard RW, Zhang Q, Jorgensen RA (1988) Chloroplast DNA diversity in populations of wild and cultivated barley. Genetics 120:1105-1110

Palmer JD, Shields CR, Cohen DB, Orton TJ (1983) Chloroplast DNA evolution and the origin of amphidiploid Brassica species. Theor Appl Genet 65:181-189

Rines HW, Gengenbach BG, Boylan KL, Storey KK (1988) Mitochondrial DNA diversity in oat cultivars and species. Crop Sci 28:171-176

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

Schoonhoven A von, Cardona C (1982) Low levels of resistance to the Mexican bean weevil in dry beans. J Econ Entomol 75:567-569

Schoonhoven A von, Cardona C, Valor J (1983) Resistance to the bean weevil and the Mexican bean weevil (Coleoptera: Bruchidae) in non cultivated common bean accessions. J Econ Entomol 76:1255-1259

Sisson VA, Brim CA, Levings CS III (1978) Characterization of cytoplasmic diversity in soybeans by restriction endonuclease activity. Crop Sci 18:991-996

Timothy DH, Levings CS III, Pring DR, Conde MF, Kermicle JL (1979) Organelle DNA variation and systematic relationships in the genus Zea:Teosinte. Proc Natl Acad Sci USA 76:4220-4224

Torro O, Tohme J, Debouck DG (1990) Wild bean Phaseolus vulgaris description and distribution. IBPGR and CIAT, Cali, Colombia

Vanderborght T (1983) Evaluation of P. vulgaris wild types and weedy forms. Plant Genet Resources Newsl 54:18-25

Vargas J, Tohme J, Debouck DG (1990) Common bean domestication in the southern Andes. Ann Rep Bean Improvement Coop 33:104-105

Weissinger AK, Timothy DH, Levings CS III, Goodman MM (1983) Patterns of mitochondrial DNA variation in indigenous maize races of Latin America. Genetics 104:365-379

CHAPTER THREE

MITOCHONDRIAL GENOME SIZE VARIATION AND RESTRICTION FRAGMENT LENGTH POLYMORPHISMS IN THREE PHASEOLUS SPECIES.

ABSTRACT. Restriction patterns of mitochondrial DNA (mtDNA) from three Phaseolus species were examined to estimate their relative genome sizes and to determine the level of interspecific variability and relatedness. Three restriction endonucleases that produced relatively simple profiles were identified and used to determine the genome size of the three species. Taking into account fragment stoichiometries, the average estimates across enzymes were 456, 324, and 400kb, respectively, for P.vulgaris , P. coccineus , and P. acutifolius. Restriction fragment length polymorphisms (RFLPs) differentiated the species when the mtDNAs were digested with seven endonucleases and hybridized with five cosmid clones covering about 200kb of mtDNA sequences. Proportions of shared restriction fragments between every two species were computed as F-values and demonstrated that P. vulgaris and P. coccineus are more related to each other than either is to P. acutifolius and that this latter has a similar degree of relationship to the other two species.

INTRODUCTION

The genus Phaseolus L. originated in the American continent and comprises about thirty species (Maréchal et al. 1978) of which approximately 90% occur in Mexico or adjacent areas of the USA and Central America (Smartt 1985). Four species have been most prominent as cultivated food crops: P. vulgaris L. (common bean, dry bean, snap bean), P. coccineus L. (runner or scarlet runner bean), P. acutifolius Gray (tepary bean), and P. lunatus L. (lima bean, sieva bean). All four species are diploid (2n=2x=22) and have both domesticated and wild representatives. In the case of both P. vulgaris and P. lunatus, small-seeded and large-seeded forms are thought to have been domesticated independently in Mesoamerica and Andean South America (Kaplan 1965, 1981). The other two species appear to have been domesticated in Mesoamerica (Kaplan 1965) although Pratt and Nabhan (1988) suggest Aridoamerica (northwestern Mexico and the southwestern United States) as the area of domestication of tepary bean. Despite some common geographical origins, the four species have evolved in different ecological zones. The scarlet runner bean is a cool temperate crop grown at high altitudes while the common bean, a warm temperate crop, grows at intermediate levels and the lima bean is mostly cultivated in humid lowland areas. The tepary has developed into a specialized desert annual (Smartt 1985).

Several studies on morphological and pollen characteristics (Maréchal et al. 1978), interspecific hybridization (reviewed in Hucl and Scoles 1985 and Mok et al. 1986), seed protein composition (Derbyshire et al. 1976, Sullivan and Freytag 1986) and their immunochemical reactions (Kloz et al. 1966, Kloz and Klozová 1974), and isozyme patterns (Bassiri and Adams 1978) have provided information on the relationships between the four cultivated species. The studies suggested that P. vulgaris and P. coccineus are most closely related, with P. acutifolius more distantly related and P. lunatus the most distant from the other three species. According to Harlan and de Wet's (1971) gene pool nomenclature system, the primary gene pool of P. vulgaris

comprises the genetic resources of the wild and domesticated populations, its secondary gene pool contains P. coccineus (and vice versa), and the tertiary gene pool includes the other species (Smartt 1985).

Comparative analysis of restriction fragment length polymorphisms (RFLPs) of organelle DNAs has enabled the resolution of species relationships in such crops as Brassica (Palmer et al. 1983), Coffea (Berthou et al. 1983), Daucus (DeBonte et al. 1984), Aegilops (Terachi and Tsunewaki 1986), Lycopersicon and Solanum (McClean and Hanson 1986), and Pennisetum (Chowdhury and Smith 1988). The relatively small size of the chloroplast genome (120-180kb) allows differences in restriction patterns to be clearly seen in agarose gels, permits restriction site differences to be mapped, and facilitates deduction of the events that resulted in the RFLPs. Due to the larger sizes and higher rates of rearrangements of plant mitochondrial genomes with respect to chloroplast genomes, measures of mitochondrial DNA (mtDNA) relatedness between plant species have been limited to the estimation of the proportion of shared restriction fragments between any two species. Plant phylogenies have been constructed using either the cladistic approach where chloroplast DNA (cpDNA) restriction sites are considered separate characters (e.g. Palmer et al. 1983), or the phenetic approach where relationships are expressed by a distance measure obtained from the number of shared restriction sites (more feasible with cpDNA) or shared restriction fragments (more common with mtDNA, e.g. Terachi and Tsunewaki 1986).

A study of mtDNA diversity within P. vulgaris indicated a high complexity (large size) of the genome and a very low level of variation (Chapter 1). The infrequent RFLPs allowed the differentiation of cultivated common bean into its two major gene pools. In this study, we have utilized mtDNA RFLPs to determine the extent of mtDNA diversity among Phaseolus species and compared it to the diversity within the species. We have used these data to examine species relationships. For this reason, we have limited our study to the primary, secondary, and tertiary gene pool representatives:

P. vulgaris, P. coccineus, and P. acutifolius. An additional objective of this research was to estimate the mitochondrial genome size of the three species because only rough estimates for P. vulgaris and P. coccineus are available.

MATERIALS AND METHODS

Plant materials

The two major Mesoamerican and Andean gene pools of common bean were found to have distinct mtDNAs as defined by five RFLPs, with no variation detected within a gene pool (Chapter 1). In this study only small-seeded accessions were used to compare P. vulgaris mtDNA to that of the other two species because of the common Mesoamerican origin in contrast to the South American origin of large-seeded common beans. For this purpose, small-seeded Malawian lines 5-2 and 5-18 (Chapter 1) were used as sources of mtDNA because of the availability of seeds. A single accession each of P. coccineus and P. acutifolius was examined. These consisted of local cultivars collected at Fransisco I. Madero, Durango, and San Pedro, Coahuila in Mexico, respectively, and were grown for seed increase at Durango, Durango.

Seeds were planted in vermiculite trays and allowed to germinate in darkness. The tepary bean seeds were first scarified to ensure a faster and more uniform germination.

Genome size estimation

The mtDNAs used for estimating the sizes of the three genomes were isolated as in Chapter 1 but were further purified through a cesium chloride (CsCl) gradient as follows. After overnight precipitation at -20^{o}C, the mtDNA pellet was resuspended in 2-3 ml $T_{10}E_{0.1}$ (10mM Tris, 0.1mM EDTA, pH 8.0) and 1.2 g solid CsCl/ml was added and allowed to dissolve. Bisbenzimide was added to a final concentration of 10-11 ug/ml and more CsCl was included to a final density of 1.67 gms/ml which

corresponded to a refractive index of 1.397 ± 0.0008. The gradients were centrifuged for 18-20 hours at 154,300 x g in a Sorvall TV 865 vertical rotor. The mtDNA band was removed and the dye extracted with NaCl-saturated isopropanol. The mtDNA was then precipitated in 1/20 volume of 5M ammonium acetate and 1/2 volume ice cold isopropanol at -20°C overnight.

Samples of the mtDNAs were digested at 37°C for 5-7 hours with NarI, SalI, or SstII (Bethesda Research Lab, Inc. (BRL)). The resulting fragments were separated by electrophoresis in 0.5-0.7% agarose gels at room temperature for 20-35 hours at constant voltage (30-40 volts), using the TAE buffer system (0.04M Tris, 0.02M sodium acetate, 0.001M EDTA, pH 8.0, and 0.5 ug/ml ethidium bromide). Three sets of molecular size markers were used, lambda DNA digested with HindIII, lambda DNA digested with HindIII and EcoRI, and high molecular weight markers from BRL. The gels were photographed with Polaroid type 55 film under UV light using one or two different exposures. The negatives were then enlarged on 8x10 sheet film (Kodak TMAX 100) and were scanned on a Gilford Response II spectrophotometer using 500nm visible light to read the % transmittance. The areas under the peaks were determined manually by counting the number of squares on graph paper in each peak.

Sizes of individual restriction fragments were estimated following the method of Schaffer and Sederoff (1981) using a computer program written in QuickBasic.

RFLP analysis

Procedures and conditions for mtDNA isolation, digestion, and electrophoresis, Southern blotting, nick translation and hybridization were as described in Chapter one. Seven restriction endonucleases were used in this study: BamHI, DraI, EcoRI, HindIII, PstI, SalI, and XhoI. Southern blots were consecutively hybridized to five cosmid clones provided by Dr. C.D. Chase (University of Florida, Gainesville). The clones contain random inserts from the bean mitochondrial genome ranging in size from 29 to 38 kb (see Chapter 1).

For each enzyme x probe combination the total number of fragments hybridizing per species (Nx, Ny) and the number of fragments shared by each pair of species (Nxy) were recorded. These were then added across clones for each enzyme and adjusted by counting fragments that hybridized to more than one clone only once. Using those figures, indices of relatedness (F-values) for the three pairs of species were computed according to Nei and Li (1979), where Fxy= [2 Nxy]/ [Nx + Ny]. Higher F-values indicate more relation between the species.

RESULTS

Genome size estimation

Estimates of the mitochondrial genome size of the three bean species were obtained from restriction profiles of three enzymes by adding up lengths of individual fragments and accounting for their multiplicities. For these analyses, the enzymes NarI, SalI, and SstII were selected from a group of 11 enzymes because they produced the simplest profiles (Figures 3.1A and 3.2A). Each digest was repeated 3-4 times and those replications were used to determine presence/absence of a band, band sizes, and band stoichiometries (Tables 3.1, 3.2 and 3.3).

Sheet film negatives scanned on a spectrophotometer allowed the identification of individual band peaks (Figures 3.1B and 3.2B). Band multiplicities were determined by considering the areas under peaks in the densitometric scans as well as the band intensities in enlarged prints of the gels. Submolar bands or fragments not present in all replicates were not counted in the totals and are not mentioned in Tables 3.1, 3.2, and 3.3.

Figure 3.1. SalI digestion patterns of mtDNAs from three Phaseolus species. V = P. vulgaris, C = P. coccineus and A = P. acutifolius. (A) Ethidium bromide stained gel, dots identify individual bands. (B) Densitometric scans of photographic negatives. Numbers under the peaks identify the individual restriction fragments, and above the peaks indicate the multiplicity of non-stoichiometric bands.

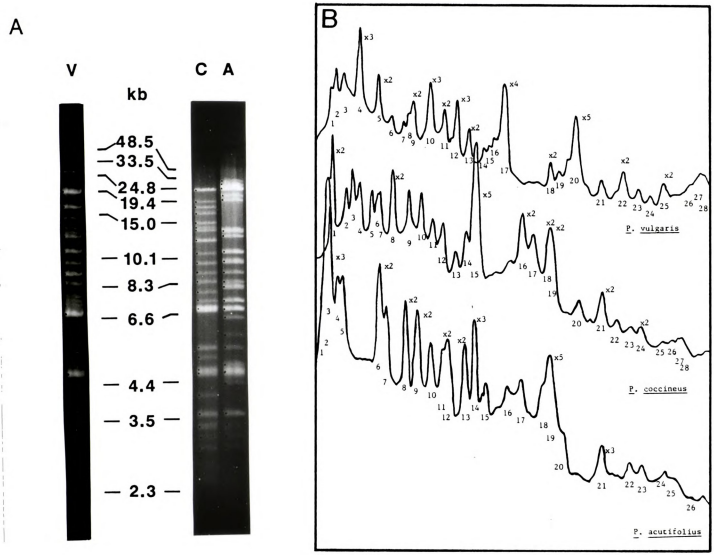
A
V
C A
kb
48.5
33.5
24.8
19.4
15.0
10.1
8.3
6.6
4.4
3.5
2.3
B
P. vulgaris
P. coccineus
P. acutifolius

Figure 3.1.

Figure 3.2. SstII digestion patterns of mtDNAs from three Phaseolus species. V = P. vulgaris, C = P. coccineus and A = P. acutifolius. (A) Ethidium bromide stained gel, dots identify individual bands. (B) Densitometric scans of photographic negatives. Numbers under the peaks identify the individual restriction fragments, and above the peaks indicate the multiplicity of non-stoichiometric bands.

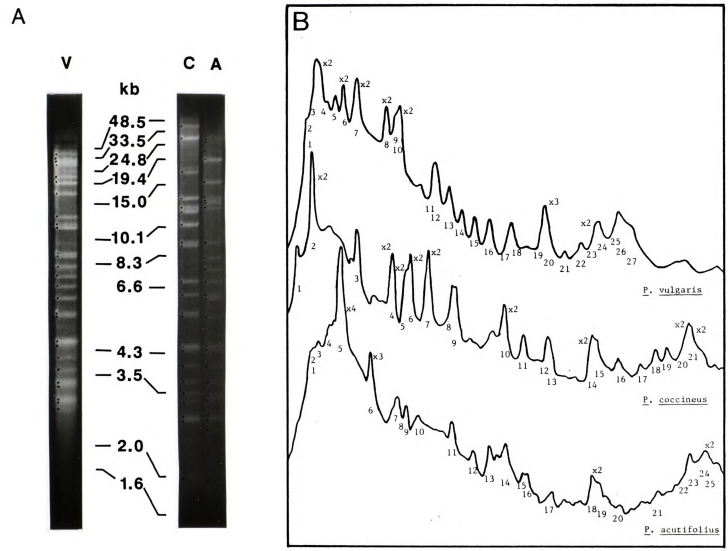
A
V
C A
kb
48.5
33.5
24.8
19.4
15.0
10.1
8.3
6.6
4.3
3.5
2.0
1.6
B
P. vulgaris
P. coccineus
P. acutifolius

Figure 3.2.

Table 3.1. Molecular weights and stoichiometries of NarI, SalI, and SstII restriction fragments of *P. vulgaris* mtDNA.

NarI			SalI			SstII		
Number	Size[a] (kb)	Stoichio-metry	Number	Size[a] (kb)	Stoichio-metry	Number	Size[a] (kb)	Stoichio-metry
1	39.56		1	29.11		1	45.62	
2	31.57		2	27.26		2	39.00	
3	27.02		3	24.88		3	37.77	x2
4	24.24		4	20.73	x3	4	32.96	
5	21.85	x3	5	17.16	x2	5	24.28	
6	19.19		6	15.11		6	22.14	x2
7	16.24		7	13.64		7	18.58	x2
8	14.67	x2	8	13.23		8	13.36	x2
9	14.10		9	12.67	x2	9	12.16	
10	12.54	x2	10	10.97	x3	10	11.70	x2
11	11.93		11	9.97	x2	11	8.92	
12	11.69		12	9.08	x3	12	8.59	
13	9.59	x2	13	8.43	x2	13	7.86	
14	9.28	x2	14	7.68		14	7.19	
15	8.60	x3	15	7.38		15	6.61	
16	7.63		16	7.11		16	6.00	
17	7.45		17	6.67	x4	17	5.28	
18	6.63		18	5.13	x2	18	5.18	
19	5.19	x2	19	4.61		19	4.24	x3
20	4.75		20	4.45	x5	20	4.11	
21	4.33		21	3.88		21	3.74	
22	3.80		22	3.50	x2	22	3.39	
23	3.58	x2	23	3.25		23	3.14	x2
24	2.94		24	3.05		24	3.04	
25	2.83		25	2.87	x2	25	2.78	
26	2.25		26	2.54		26	2.71	
27	2.05		27	2.42		27	2.55	
28	1.70		28	2.36		28	1.91	
29	1.56					29	1.70	
30	1.31					30	1.30	
Total	330.07		Total	279.14		Total	347.81	
	445.82[b]			**458.24[b]**			**462.98[b]**	

a Sizes of individual fragments are average values from 3 or 4 gels.

b Numbers in bold are the sum of fragment sizes, taking into account their relative multiplicities.

Table 3.2. Molecular weights and stoichiometries of NarI, SalI, and SstII restriction fragments of *P. coccineus* mtDNA.

NarI			SalI			SstII		
Number	Size[a] (kb)	Stoichio-metry	Number	Size[a] (kb)	Stoichio-metry	Number	Size[a] (kb)	Stoichio-metry
1	40.07		1	25.80	x2	1	45.55	
2	34.78		2	20.93		2	33.92	x2
3	24.11	x2	3	19.42		3	18.60	
4	22.18	x2	4	17.32		4	13.64	x2
5	18.32		5	15.50		5	12.44	x2
6	16.91		6	14.59		6	11.88	x2
7	14.67		7	14.08		7	10.60	x2
8	9.86	x2	8	12.63	x2	8	9.02	
9	9.47		9	10.94		9	8.73	
10	7.81		10	9.97		10	6.65	x2
11	7.70		11	9.10		11	6.01	
12	6.91		12	8.41		12	5.34	
13	6.78		13	7.69		13	5.21	
14	5.31	x2	14	7.17		14	4.29	x2
15	4.82		15	6.77	x5	15	4.18	
16	4.38		16	5.19	x2	16	3.80	
17	4.27		17	4.88		17	3.43	
18	3.90		18	4.48	x2	18	3.21	
19	3.66		19	4.40	x2	19	3.04	
20	2.98		20	3.89		20	2.79	x2
21	2.87		21	3.51	x2	21	2.73	x2
22	2.20		22	3.26		22	1.96	
23	2.00		23	3.05		23	1.30	
24	1.66		24	2.86	x2			
25	1.50		25	2.62				
26	1.26		26	2.51				
			27	2.39				
			28	2.37				
Total	260.38 **321.84**[b]		Total	245.73 **331.68**[b]		Total	218.32 **317.26**[b]	

a Sizes of individual fragments are average values from 3 or 4 gels.

b Numbers in bold are the sum of fragment sizes, taking into account their relative multiplicities.

Table 3.3. Molecular weights and stoichiometries of NarI, SalI, and SstII restriction fragments of *P. acutifolius* mtDNA.

NarI			SalI			SstII		
Number	Size[a] (kb)	Stoichio-metry	Number	Size[a] (kb)	Stoichio-metry	Number	Size[a] (kb)	Stoichio-metry
1	38.17		1	29.08		1	32.94	
2	32.40	x2	2	27.41		2	30.54	
3	25.11	x3	3	25.13	x3	3	28.45	
4	21.67		4	22.40		4	24.47	
5	18.04		5	21.19		5	22.43	x4
6	17.30		6	14.11	x2	6	16.48	x3
7	14.70	x2	7	13.15		7	13.63	
8	11.53		8	11.12	x2	8	13.03	
9	10.94		9	10.15	x2	9	12.60	
10	9.76		10	9.12		10	12.22	
11	9.22		11	8.47		11	8.91	
12	8.55		12	8.12	x2	12	7.84	
13	7.60	x3	13	7.28	x2	13	7.19	
14	6.67	x2	14	6.83	x3	14	6.58	
15	5.29		15	6.40		15	5.96	
16	4.80	x4	16	5.65		16	5.80	
17	4.34		17	5.17		17	5.15	
18	3.78		18	4.64		18	4.25	x2
19	3.30		19	4.43	x5	19	4.10	
20	2.98		20	4.16		20	3.75	
21	2.87		21	3.50	x3	21	3.39	
22	2.20		22	3.22		22	3.16	
23	2.00		23	3.05		23	3.07	
24	1.82		24	2.83		24	2.80	x2
25	1.66		25	2.51		25	2.73	
26	1.50		26	2.38		26	2.54	
27	1.26					27	1.97	
						28	1.58	
Total	269.46 **403.05**[b]		Total	261.50 **400.92**[b]		Total	287.56 **394.86**[b]	

a Sizes of individual fragments are average values from 3 or 4 gels.

b Numbers in bold are the sum of fragment sizes, taking into account their relative multiplicities.

By averaging the values obtained with the three enzymes, we estimate the mitochondrial genome sizes of P. vulgaris, P. coccineus, and P. acutifolius to be 456kb, 324kb, and 400kb, respectively.

RFLP analysis

In order to determine the level of mitochondrial interspecific variation, seven restriction endonucleases were used to digest the mtDNA of three Phaseolus species. The resulting profiles after electrophoresis were quite complex, but differences between the species were obvious for all en-zymes. Figure 3.3A shows the patterns observed with the enzyme EcoRI. For a more accurate analysis of the variation between species, the gels were blotted and the Southern filters were consecutively hybridized to five cosmid clones. The number of mtDNA fragments hybridizing to each clone varied between species and for each enzyme. Figure 3.3 shows only the case of the EcoRI digestion and the hybridizations with clones C2, C6, and C8. The number of fragments hybridizing per species and those common between two species were totaled across the five clones for each species x enzyme combination and the proportions of shared fragments (F-values) between two species were calculated according to Nei and Li (1979) for all the enzymes (Table 3.4).

F-values ranged from a low 0.23 for DraI to a high of 0.74 for HindIII. For five out of seven of the enzymes, the highest F-value was for the vulgaris-coccineus comparison, indicating that these species are most closely related. When the F-values were averaged across enzymes, the proportion of shared fragments was 0.63, 0.54, and 0.53 for the vulgaris/coccineus, vulgaris/acutifolius, and coccineus /acutifolius comparisons.

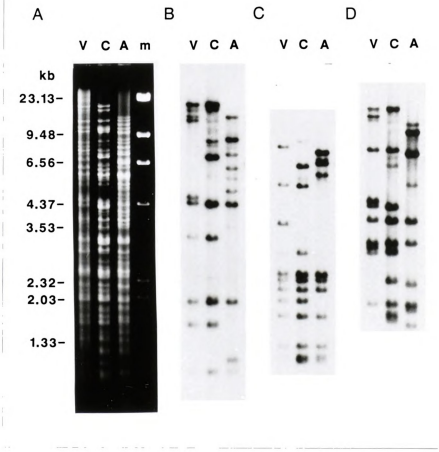

Figure 3.3. EcoRI digestion patterns of mtDNAs from three Phaseolus species. V = P. vulgaris, C = P. coccineus and A = P. acutifolius. (A) Ethidium bromide stained gel. (B,C,D) Southern blots hybridized respectively with clones C2, C6 and C8.

Table 3.4. Proportion of mtDNA fragments shared between three pairs of *Phaseolus* species as obtained from hybridization experiments to five cosmid clones.

	Total fragments[a]			Common fragments[b]			F-values[c]		
Enzyme	Nv	Nc	Na	Nvc	Nva	Nca	Fvc	Fva	Fca
BamHI	36	33	36	24	20	19	0.70	0.56	0.55
DraI	37	29	32	12	8	9	0.36	0.23	0.30
EcoRI	51	49	48	36	26	24	0.72	0.53	0.50
HindIII	49	41	45	33	35	30	0.73	0.74	0.70
PstI	28	28	28	18	17	14	0.64	0.61	0.50
SalI	21	20	18	10	9	9	0.49	0.46	0.47
XhoI	33	31	28	20	18	19	0.63	0.59	0.64
Totals	255	231	235	153	133	124	0.63	0.54	0.53

a Total number of restriction fragments hybridizing to the five clones for each species, v = *P. vulgaris*, c = *P. coccineus*, and a = *P. acutifolius*.

b Total number of fragments across the five clones shared by two species.

c Proportion of shared fragments between two species, $F_{xy} = \frac{2\ N_{xy}}{N_x + N_y}$ (Nei and Li 1979).

DISCUSSION

Genome size estimation

Plant mitochondrial genomes vary remarkably in size (200 to 2400 kb), and are significantly larger than those from animals (15-18kb) and fungi (18-78kb) (Levings 1983, Newton 1988). The variation in size can be quite high within the same family of plants as Ward et al. (1981) showed for the cucurbits (320 - 2400 kb), or minimal (208 - 242 kb) as in the Brassica/Raphanus mitochondrial genomes (Lebacq and Vedel 1981, Palmer and Herbon 1988).

The most accurate method for estimating the mitochondrial genome size is by restriction mapping, which has been used for maize (Lonsdale et al. 1984), for Brassica (Palmer and Shields 1984, Palmer and Herbon 1986), and for spinach (Stern and Palmer 1986). However, the large sizes of mitochondrial genomes have limited both its use, as well as that of electron microscopy in assessing lengths of mtDNA molecules. Two other methods have been utilized more extensively, namely reassociation kinetics and summation of fragment sizes from restriction profiles. Although both methods have resulted in figures that agree, the estimates from restriction maps are usually 15-30% larger. The restriction profile method was used in the present study to determine the genome size of three Phaseolus species. In order to address the problems of the complex banding patterns and non-stoichiometry of some of the fragments associated with this method, three endonucleases that produce relatively simple profiles were chosen, and from these, we have estimated the copy number of individual fragments. To optimize the resolution of the larger fragments obtained with the less frequent cutters, low percentage agarose gels were run at low voltages for long times. In addition, CsCl-purification of mtDNAs helped minimize the background in the gels, three to four gels of each enzyme were run to increase the confidence of the estimates, and 30 molecular weight markers ranging in size from 48.5 to 0.56 kb were present in all gels.

For each species, the estimates obtained with the different enzymes were close (less than 5% difference) to each other (Tables 3.1, 3.2 and 3.3). The average respective values of 456kb, 324kb, and 400kb for common bean, runner bean, and tepary bean fall within or close to the range of 290-430 kb for other leguminous crops (Ward et al. 1981, Bendich 1982, Stern and Palmer 1984). A smaller genome size for P. coccineus compared to P. vulgaris has previously been reported (Bannerot and Charbonnier 1988). Using SalI digestion profiles, Bannerot and Charbonnier (1988) had estimated the mitochondrial genome size of P. vulgaris and P. coccineus to be about 320kb and 240-260kb, respectively. Their estimates did not take into account band multiplicitites, and they recognized that the use of several enzymes will improve the estimation of the genome size. The totals obtained with our SalI digests, disregarding individual band stoichiometries, was 40kb less for P. vulgaris but agree well for P. coccineus when compared to these previous estimates (Tables 3.1 and 3.2). The difference in the common bean mtDNA estimate is probably due to variation in resolution in certain parts of their gels or ours.

Phaseolus species relationships

Considerable variation of mtDNA restriction profiles and hybridization patterns differentiated the three Phaseolus species, in contrast to the high homogeneity within P. vulgaris (Chapter 1). Due to this high level of mtDNA variation and because no restriction maps of the mitochondrial genomes of beans are as yet available, it was not possible to determine the nature of the mutations giving rise to the observed polymorphisms. However, the variability in genome size between the three species indicates that insertion/deletion events must have taken place during the evolution of the Phaseolus genus. The variation in the proportion of shared fragments (Table 3.4, F-values) across the different endonucleases used, in particular the lower values for SalI and DraI, seem to indicate the additional role of nucleotide substitution in the evolution of those genomes. Nevertheless, we cannot rule out the possible past occurrence of

rearrangements, especially in light of the recent study in Brassica where Palmer and Herbon (1988) have clearly demonstrated the significant role of mtDNA rearrangements in the evolution of the mitochondrial genome in that genus.

Due to the probable rearrangements and length mutations, we are not using the F-values as measures of nucleotide divergence (Nei and Li 1979) but as indices of relatedness between every pair of species. For the same reason, the data from the different endonucleases are not pooled since the same length mutation or rearrangement will be detected with several, if not all of the enzymes used. F-values have also been used as indices of relatedness in studies of species relationships in Daucus (DeBonte et al. 1984, Ichikawa et al. 1989), Aegilops (Terachi and Tsunewaki 1986) and Pennisetum (Chowdhury and Smith 1988). These F-values agreed with the conventional classification methods but also provided unique insights into the phylogenetic relationships.

In the case of Phaseolus, the indices of mtDNA relatedness comply with the species relationships as defined by morphological, genetic, and biochemical data. The higher Fvc values as compared to Fva or Fca (Table 3.4) clearly demonstrate that P.vulgaris and P. coccineus are more related to each other than either is to P. acutifolius. A vulgaris-coccineus species complex that includes the wild vulgaris forms and the three other coccineus subspecies polyanthus, obvallatus, and formosus has been recognized by Marèchal et al. (1978) after an extensive study of morphological and pollen characteristics.

The close genetic relationship of the common and the scarlet runner bean was suggested to result from either common descent or through species introgression (Bassiri and Adams 1978). Because of the maternal inheritance of mitochondria, introgression between those species could perhaps be detected as either the presence of a vulgaris-like chondriome (mitochondrial genome) in wild coccineus populations or a coccineus-like chondriome in wild vulgaris populations. Since most successful

hybridizations between the two species occur when P. vulgaris is the female parent (Hucl and Scoles 1985), the first case is more probable than the second. It would be interesting to look at the mtDNA diversity within P. coccineus, especially in wild populations suspected or known to grow within the natural range of P. vulgaris wild populations.

P. acutifolius is often referred to as closely related to the vulgaris-coccineus complex rather than to either spe-cies. Our results indicate the same kind of relationship. The Fva and Fca values are not consistently larger or smaller when compared across the endonucleases used (Table 3.4). Their averages (Fva=0.54 and Fca=0.53) are not significantly different, indicating the same level of mtDNA divergence between P. vulgaris and P. acutifolius, as between P. coccineus and P. acutifolius.

A different approach such as the examination of cpDNA variation and mapping of its various mutations, or the comparison of DNA sequence data will probably provide a finer resolution of such relationships.

LIST OF REFERENCES

Bannerot H, Charbonnier L (1988) Induction de stérilités males cytoplasmiques par alloplasmie dans le genre Phaseolus: étude génétique et moléculaire. In: Berville A (ed.) Variabilité génétique cytoplasmique et stérilité male cytoplasmique. INRA, Paris, pp 181-198

Bassiri A, Adams MW (1978) An electrophoretic survey ofseedling isozymes in several Phaseolus species. Euphytica 27:447-459

Bendich AJ (1982) Plant mitochondrial DNA: the last frontier. In: Slonimsky P, Borst P, Attardi G (eds) Mitochondrial genes. Cold Spring Harbor Lab, New York, pp 477-481

Berthou F, Mathieu C, Vedel F (1983) Chloroplast and mitochondrial DNA variation as indicator of phylogenetic relationships in the genus Coffea L. Theor Appl Genet 65:77-84

Chowdhury MKU, Smith RL (1988) Mitochondrial DNA variation in pearl millet and related species. Theor Appl Genet 76:25-32

DeBonte LR, Matthews BF, Wilson KG (1984) Variation of plastid and mitochondrial DNAs in the genus Daucus. Amer J Bot 71:932-940

Derbyshire E, Yarwood JN, Neat E, Boulter D (1976) Seed proteins of Phaseolus and Vigna. New Phytol 76:283-288

Harlan JR, deWet JMJ (1971) Towards a rational classification of cultivated plants. Taxon 20:509-517

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

Kaplan L (1965) Archaeology and domestication in American Phaseolus (Beans). Econ Bot 19:358-368

Kaplan L (1981) What is the origin of the common bean? Econ Bot 35:240-254

Kloz J, Klozovà E (1974) The protein euphaseolin in Phaseolinae - a chemotaxonomical study. Biol Pl 16:290-300

Kloz J, Klozovà E, Turkovà V (1966) Chemotaxonomy and genesis of protein characters with special reference to the genus Phaseolus. Preslia 38:229-236

Lebacq P, Vedel F (1981) SalI restriction enzyme analysis of chloroplast and mitochondrial DNAs in the genus Brassica. Plant Sci Lett 23:1-9

Levings CS III (1983) The plant mitochondrial genome and its mutants. Cell 32:659-661

Lonsdale DM, Hodge TP, Fauron CMR (1984) The physical map and organization of the mitochondrial genome from the fertile cytoplasm of maize. Nucleic Acids Res 12:9249-9261

Maréchal R, Mascherpa J-M, Stainier F (1978) Etude taxonomique d'un groupe complexe d'espèces des genres Phaseolus et Vigna (Papilionaceae) sur la base de données morphologiques et polliniques, traitées par l'analyse informatique. Boissiera 28:1-273

McClean PE, Hanson MR (1986) Mitochondrial DNA sequence divergence among Lycopersicon and related Solanum species. Genetics 112:649-667

Mok DWS, Mok MC, Rabako-Arihanta A, Shii CT (1986) Phaseolus: wide hybridization through embryoculture. In: Bajaj YPS (ed) Biotechnology in agriculture and forestry. Crops I. Springer Verlag, Berlin, pp 309-318

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269-5273

Newton KJ (1988) Plant mitochondrial genomes: organization, expression and variation. Ann Rev Plant Phys Plant Mol Biol 39:503-532

Palmer JD, Herbon LA (1986) Tripartite mitochondrial genomes of Brassica and Raphanus: reversal of repeat configurations by inversion. Nucleic Acids Res 14:9755-9765

Palmer JD, Herbon LA (1988) Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. J Mol Evol 28:87-97

Palmer JD, Shields CR (1984) Tripartite structure of the Brassica campestris mitochondrial genome. Nature 307:437-440

Palmer JD, Shields CR, Cohen DB, Orton TJ (1983) Chloroplast DNA evolution and the origin of amphidiploid Brassica species. Theor Appl Genet 65:181-189

Pratt RC, Nabhan GP (1988) Evolution and diversity of Phaseolus acutifolius genetic resources. In: Gepts P (ed) Genetic resources of Phaseolus beans. Kluwer Acad Publ, Dordrecht, pp 409-440

Schaffer HE, Sederoff RR (1981) Improved estimation of DNA fragment lengths from agarose gels. Anal Biochem 115:113-122

Smartt J (1985) Evolution of grain legumes. IV. Pulses in the genus Phaseolus. Expl Agric 21:193-207

Stern DB, Palmer JD (1984) Recombination sequences in plant mitochondrial genomes: diversity and homologies to known mitochondrial genes. Nucleic Acids Res 12:6141-6157

Stern DB, Palmer JD (1986) Tripartite mitochondrial genome of spinach: physical structure, mitochondrial gene mapping, and locations of transposed chloroplast DNA sequences. Nucleic Acids Res 14:5651-5665

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

Terachi T, Tsunewaki K (1986) The molecular basis of genetic diversity among cytoplasms of Triticum and Aegilops 5. Mitochondrial genome diversity among Aegilops species having identical chloroplast genomes. Theor Appl Genet 73:175-181

Ward BL, Anderson RS, Bendich AJ (1981) The mitochondrial genome is large and variable in a family of plants (Cucurbitaceae). Cell 25:793-803

SUMMARY AND CONCLUSIONS

A project of the Bean/Cowpea Collaborative Research Support Program (CRSP) was initiated in 1981 between Michigan State University in the United States and Bunda College in Malawi, East Africa, to study the genetic diversity of beans from Malawi. Landraces were collected from 15 farm sites in Malawi, and 25 lines from each landrace were examined for morphoagronomic diversity (Martin 1984). Allozyme variability was later investigated in the same 375 lines (Sprecher 1988). While the first study revealed extensive variation in quantitative traits, the second showed very little variability at six isozyme loci and a clustering of the lines into two major groups which corresponded to the small-seeded Mesoamerican and the large-seeded Andean gene pools of Phaseolus vulgaris. The objectives of the first part of this dissertation were to examine the diversity of a sample of the Malawian lines at the mitochondrial DNA (mtDNA) level, and compare and contrast it to the morphoagronomic and allozyme variation. Twenty lines were selected to include both "related" as well as "distant" lines; three pure-bred cultivars were included as controls. Despite the use of eight restriction endonucleases to digest the mtDNA, and cosmid clones covering approximately 200 kb of the bean mitochondrial genome, the results showed a high level of mtDNA homogeneity among the 23 bean lines examined. However five restriction fragment length polymorphisms (RFLPs) were detected that differentiated again the Malawian lines into the small-seeded and the large-seeded gene pools. This provides the first evidence that the cytoplasms of the two major gene pools of P. vulgaris are distinct. The RFLPs thus identified may be useful in the future as markers to determine the maternal parent in plants believed to be derived from inter-gene pool

crosses. They could even be used to follow the inheritance of mitochondria in beans. Although mitochondria have been thought to be strictly maternally inherited in higher plants as they are in animals, much of the evidence in higher plants is based on the inheritance of cytoplasmic male sterility, a trait associated with mtDNA (Neale and Sederoff 1988). Some paternal transmission of mtDNA has been observed in barley-rye intergeneric crosses (Soliman et al. 1987), and Neale et al. (1989) recently reported strict paternal inheritance of mtDNA in redwood.

In the second part of the dissertation, the mtDNA diversity was examined in extant wild bean populations in order to understand how and when the mitochondrial genomes of the two gene pools of cultivated beans became distinct. The mtDNA of six wild bean accessions from Central and South America were digested with nine endonucleases and analyzed by Southern hybridization. Observation of 20 RFLPs clearly pointed to a higher amount of mtDNA variability in wild than in cultivated beans. The distribution of mtDNA RFLPs among the wild beans indicated that in four out of the five inter-gene pool RFLPs, the mutations arose soon after domestication, two in the Mesoamerican gene pool and two in the Andean. The mtDNA of wild bean accessions from the southern Andes were found to be very similar to that of the cultivated Andean beans, supporting the classification of the Andean wild bean as a botanical form (Baudet 1977, Delgado-Salinas 1985) rather than as a separate species (Brücher 1988). The highest similarity of the mtDNA patterns of the wild bean accessions from Guatemala and from south Peru and Bolivia to those of the cultivated Mesoamerican and Andean beans, respectively, may point to specific areas where wild beans were first domesticated. The observed stability of the mtDNA restriction patterns makes it an attractive system to study bean origin and evolution. The analysis of a larger number of wild bean accessions is essential before definite answers about bean origin and domestication can be presented. The identification of four endonucleases that detected significantly more mtDNA RFLPs than did other enzymes will facilitate

this task.

In the third part of this dissertation, mtDNA restriction patterns from three Phaseolus species were examined in order to estimate their relative genome sizes and determine the level of interspecific variability. Using three endonucleases that produced relatively simple restriction profiles, and taking into account fragment stoichiometries, the mitochondrial genome size estimates for P. vulgaris, P. coccineus, and P. acutifolius were 456, 324, and 400 kb, respectively. The high level of mtDNA variation between the species contrasted with the homogeneity within P. vulgaris. In this context, it would be interesting to examine the level of intraspecific mtDNA variation within P. coccineus and/or P. acutifolius. Indices of relatedness between species were computed as the proportions of shared restriction fragments between every two species and demonstrated that P. vulgaris and P. coccineus are more related to each other than either is to P. acutifolius, which was equally distant from the other two species. The close genetic relationship of the common bean and the scarlet runner bean may have resulted from species introgression (Bassiri and Adams 1978). Examination of mtDNA in wild P. vulgaris and P. coccineus sympatric populations could provide evidence for or against the occurrence of introgression. The relationships between the three species studied here as well as between other Phaseolus species may be resolved better with the use of a less complex system than the mitochondrial genome. The chloroplast genome with its relatively small genome size and low rate of rearrangements would be appropriate. Alternatively, comparisons of sequence data from specific regions of the nuclear or the organellar genomes would allow a finer resolution of phylogenetic relationships.

LIST OF REFERENCES

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

Baudet JC (1977) Origine et classification des especes cultivees du genre Phaseolus. Bull Soc Roy Bot Belg 110:65-76

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

Delgado-Salinas AO (1985) Systematics of the genus Phaseolus (Leguminosae) in North and Central America. PhD Dissertation, University of Texas, Austin

Martin GB (1984) Genetic diversity of bean landraces in Northern Malawi. MS Thesis, Michigan State University, East Lansing, MI. 161 pages

Neale DB, Sederoff RR (1988) Inheritance and evolution of conifer genomes. In: Hanover JW, Keathley DE (eds) Genetic manipulations of woody plants. Plenum Press, New York, pp 251-264

Neale DB, Marshall KA, Sederoff RR (1989) Chloroplast and mitochondrial DNA are paternally inherited in Sequoia sempervirens D. Don Endl. Proc Natl Acad Sci USA 86: 9347-9349

Soliman K, Fedak G, Allard RW (1987) Inheritance of organelle DNA in barley and Hordeum x Secale intergeneric hybrids. Genome 29:867-872

Sprecher SL (1988) Allozyme differentiation between gene pools in common bean (Phaseolus vulgaris L.), with special reference to Malawian germplasm. PhD Dissertation, Michigan State University, East Lansing, MI. 207 pages