CHARACTERIZATION OF THE GENETIC DIVERSITY CONTAINED WITHIN THE GENUS ARACHIS

by

MELANIE LEIGH NEWMAN

Under the Direction of Carol Robacker

ABSTRACT

The genus *Arachis* contains both the cultivated peanut, *A. hypogaea* and its wild relatives. Although many investigations into the genetic diversity and relationships of species within the genus have been conducted, there is still a need to better characterize the genetic diversity in the genus by studying a larger sampling of species with a common set of molecular markers. By understanding the genetic structure of *A. hypogaea* and its wild relatives, breeding programs will be better informed as to how one might introduce wild ancestry into the cultivated peanut for crop improvement. This research has focused on understanding the genetic diversity and its partitioning in the wild species contained within the genus *Arachis* using Amplified Fragment Length Polymorphisms (AFLP) and on the genetic structure and diversity within *A. hypogaea* using Simple Sequence Repeats (SSR). In the wild species, the results indicate that there is a great deal of diversity within the genus that is partitioned mainly within individual sections as opposed to among taxonomic sections. This is important to both plant breeders and germplasm curation alike by highlighting the potential benefits from choosing plants within a particular section as opposed to broadly sampling plants across different taxonomic sections. The AFLP data also support the idea that *A. monticola* is the tetraploid precursor of *A. hypogaea* and *A. batizicoi* is the B genome donor. In the cultivated peanut, genetic diversity is much more limited and the availability of markers to assess genetic diversity is wanting. However, SSR analysis has revealed that the diversity within the species is partitioned better according to geographic distribution of the plants rather than by taxonomic grouping. This information can be used in germplasm collection and by plant breeders by suggesting that more diversity may be gained when selecting plants from diverse geographic regions rather than basing selection on taxonomy alone.

INDEX WORDS: Arachis, Germplasm, Phylogeny, Amplified Fragment Length Polymorphism, Simple Sequence Repeats

CHARACTERIZATION OF GENETIC DIVERSITY CONTAINED WITHIN THE GENUS ARACHIS

by

MELANIE LEIGH NEWMAN

B.S., The State University of West Georgia, 1993

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2003

CHARACTERIZATION OF GENETIC DIVERSITY CONTAINED WITHIN THE GENUS ARACHIS

by

MELANIE LEIGH NEWMAN

Major Professor: Carol Robacker

Committee: Athey Gillaspie
Jerry Johnson
Peggy Ozias-Akins
Roy Pittman

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2003

DEDICATION

I would like to dedicate this dissertation to my two beautiful children, Savannah Leigh and Thomas Shawn, Jr. I hope this accomplishment of mine inspires you to pursue wholeheartedly the things in life that you desire.

ACKNOWLEDGMENTS

I would like to thank Carol Robacker for her support and guidance in work that has been done to complete this degree. She is not only a great advisor but a great friend. I would also like to thank each of my committee members for their support - Graves Gillaspie, Orville Lindstrom, Roy Pittman and Peggy Ozias-Akins. I am especially grateful to Dr. Pittman, who has not only offered his support as a committee member, but as my supervisor and friend. Without his support and understanding through these years, I would not have been able to accomplish earning my degree. I want to thank Rob Dean who has been my coach and helped me with all my data analysis heartaches. I want to thank Tracie Jenkins for her support both academically and personally. I would like to thank my Father for being so proud of me and for his encouragement. I would like to thank my Mother for inspiring me to be the best I could be in life and in my career. All those hours you patiently sat with me helping me with my homework has finally paid off - thank you! And mostly, I would like to thank the Lord for giving me the gifts of intelligence and determination that has enabled me to make this dream of mine come true.

TABLE OF CONTENTS

Page

CHAPTER

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Purpose of Study

The genus *Arachis* is a diverse taxon containing both the cultivated peanut, *A. hypogaea*, and its wild relatives. The genus contains 69 known species (Krapovickas and Gregory, 1994) partitioned into nine sections based on morphological and interspecific hybridization studies (Gregory and Gregory, 1979). In addition, Valls and Simpson (1997) have identified 11 new species which have yet to be fully described. Currently, the only species of significant economic importance is *A. hypogaea* which is an important oilseed crop grown worldwide. However, other species have been utilized. For example, *A. glabrata* and *A. pintoi* are used as forage crops (Maass et al. 1993); *A. repens* is used as an ornamental and a groundcover (Gimenes et al., 2000); and *A. villosulicarpa* is used as a food crop in the Mate Grosso State of Brazil (Galgaro L. et al., 1997).

A. hypogaea is an allotetraploid that is believed to have arisen through the hybridization of two wild diploid *Arachis* species. Much debate has centered on the identification of the progenitor species, but it is generally agreed that the progenitors are most likely species found within the section *Arachis*. Therefore, most, but not all, of the studies characterizing the genetic and morphological features of the different wild *Arachis* species has primarily focused on species within this section. There have been, however, several studies focused on species in the other sections (Singh et al. 1994; Cherry, 1975; and Klozova 1983 a,b; Stalker et al., 1994).

The overall goal of this dissertation is twofold. First, AFLP marker technology will be used to assess the genetic variability within and among seven of the nine sections of the genus *Arachis*. Additionally, the AFLP data will also be used to determine relationships of wild species to cultivated peanut, particularly in section *Arachis*, to shed further light on the origin of the cultivated peanut. Secondly, SSRs will be used to look solely within the species *A. hypogaea*, to determine the amount of genetic variability within the species and to determine the best manner to sample from this species in order to capture most of the genetic variability it has to offer. It will be used to assess the degree of genetic uniformity within accessions - a key interest in maintaining a germplasm system.

Justification

Curation of a germplasm collection is complex and involves balancing regeneration of current material with collection of new material. Currently, there are 7951 accessions of *A. hypogaea* and 566 accessions of other *Arachis* species available for distribution from the United States Germplasm System located in Griffin, Georgia (GRIN, 2003). The cost to regenerate each accession once every 10 years in order to maintain seed viability would far exceed the funds available to the curator, leaving many accessions vulnerable to loss. Having a marker system to evaluate the germplasm would allow for more efficient regeneration. Through marker-assisted identification of genetically redundant accessions, the curator could eliminate unnecessary regeneration of duplicated accessions. Markers allow the curator to assess uniformity within an accession and facilitate decisions on whether or not to split one accessions into multiple accessions. Markers allow a curator to characterize the overall genetic variation within a collection enabling the validation of the existing cultivated core collection that is meant to represent the majority of genetic diversity held in the overall cultivated collection. Core collections could be used by plant breeders to identify traits of interest by increasing the probability of finding different levels of resistance, drought tolerance, etc. in a smaller, genetically diverse sub-population. Hence, by evaluation of the genetic diversity within *Arachis* as put forth in this dissertation, a great benefit would be received by curators of a germplasm system and plant breeders. This work will be the most comprehensive, cohesive evaluation of the genetic diversity within the genus to date. It will include evaluation of the majority of the species and taxonomical sections in *Arachis* with the same genetic marker system.

Literature Review

The genus *Arachis* comprises nine sections based on morphology, geographical distribution and hybridization studies (See Appendix One). Most of the work to support this taxonomic scheme has dealt with analysis of protein profiles. Klozova et al. (1983a,b) studied seed protein patterns using immunochemical techniques for a limited number of species spanning five sections of *Arachis* including *Arachis, Erectoides, Caulorhizae, Extranervosae*, and *Triseminatae*. The differences found in the proteins supported the breakdown of the genus into the different sections. It also revealed a strong relationship between *A. monticola* and *A. hypogaea*, supporting the hypothesis that they are closely related (Gregory and Gregory, 1976) and not biologically distinct species (Smartt, 1979). Both *A. batizicoi* and *A. cardenasii* showed a high degree of similarity to *A. monticola* which supports the idea that these wild diploid species may be possible progenitors of *A. hypogaea*. Using the same techniques, Neucere and Cherry

(1975) investigated 36 species across seven sections and found "extensive hetereogeneities" within most of the studied species. Grosso et al. (2000) used seed protein profiles to investigate relationships of accessions within *A. hypogaea*. The data supported the division of the species into two distinct subspecies, subsp. *fastigiata* and subsp. *Hypogaea* and, to a lesser extent, the further division into varieties. Using seed protein profiles, Singh et al. (1994) studied phylogenetic relationships of 19 accessions representing seven sections of genus *Arachis*. The data supported the seven section classification of *Arachis* and revealed a high level of variation within but not between species within the section *Arachis*. Lu and Pickersgill (1993) used isozymes to estimate genetic diversity and relationships among species within the section *Arachis*. They identified 13 polymorphic genetic loci using eight enzymes which were able to group species according to whether they were annuals or perennials. Two of the 13 loci were able to separate *A. hypogaea* into its two subspecies. Stalker et al. (1994) did an extensive study of the genus as a whole and surveyed 113 accessions representing six sections of *Arachis*. Few polymorphic loci were identified within the cultivated peanut but the diploid species were highly variable. Within-accession variability was also identified, giving concern over germplasm maintenance. The geographic areas that exhibited the greatest interspecific genetic diversity were identified and included Mato Grosso, Brazil for wild species and north, north-central, south and southeast Brazil for *A. hypogaea*.

Although protein\isozyme work has proven insightful, molecular studies at the DNA level to investigate relationships and diversity in the genus *Arachis* have focused mainly on species within section *Arachis* while molecular investigations into other sections have been limited. Genetic variation and phylogenetic relationships of species contained within the sections *Extranervosae, Caulorhizae, Heteranthae, and Triseminatae* have been studied by Galgaro et.al. (1998) using both restriction fragment length polymorphism (RFLP) and randomly amplified polymorphic DNA (RAPD). The study included 75 individuals representing 13 species. The species *A. pietrarellii* and *A villsulocarpa*, both from section *Extranervosae*, were very similar, supporting the hypothesis that these two species are closely related based on previous isozyme work (Galgaro et al., 1997). The accession, V 7786, *Arachis* sp. *Pietrarellii,* also was distinguished from *A. pietrarellii* suggesting that it may be a new species. The species, *A. dardanoi*, which has been classified into section *Heteranthae*, showed more similarity to species of section *Extranervosae* than *A. pusilla*, the other species in *Heteranthae,* suggesting *A. dardani* may be grouped in the incorrect section. This is interesting since the *A. dardani* specimen used in the study was collected from an area in northeastern Brazil where the species of these sections overlap. Galgaro et al. (1998) recommends more work in this area to determine the proper classification. Gimenes et.al. (2000) studied genetic variation and phylogenetic relationships in the section *Caulorrhizae* using RAPD markers. The study population consisted of 52 accessions of *A. pintoi* and 12 accessions of *A. repens*. *A. pintoi* and *A. repens* have been shown to be useful as a forage crop, an ornamental plant, and a ground cover. Recent collections of germplasm for these two species have proven difficult to unequivocally identify based on the defining morphological characters, such as shape and size of leaflets, due to the fact that many intermediate forms of these characters are being seen in the new collections (Valls, 1996). The RAPD data produced in a study by Gimenes et al. (2000) were able to discriminate between the two species, but a high degree of similarity (74% of RFLP bands were shared by the two species) did not justify the separation of the population into two species. These studies have provided useful insights into the relationship between the species within various sections of genus *Arachis* and demonstrate the need for further molecular studies investigating the genetic diversity and relationships that exist within the genus.

In trying to understand the amount and structure of genetic diversity within *A. hypogaea*, the main obstacle has been the lack of polymorphisms detected. Initial studies utilized isozymes to study the genetic variability within *A. hypogaea* but were unsuccessful in finding any variation (Lu and Pickersgill, 1993; Stalker et al., 1994). RFLPs revealed a high degree of diversity within the wild species of section *Arachis*, but no significant polymorphisms were found in *A. hypogaea* (Kochert et al., 1991; Halward et al., 1991; Halward et al., 1993; Paik-Ro, et al., 1992). Hilu and Stalker (1995) were also unable to find any polymorphisms using random amplified polymorphic DNA (RAPD). V. Subramanian and colleagues (2000) found a marginal amount of polymorphism in the cultivated species using RAPDs, with 27 polymorphic bands out of 408 total bands, or 6.6% polymorphism. However, taking into consideration the lack of reproducibility associated with RAPDs, this level of polymorphism is not ideal. More recently, Raina et.al (2001) used RAPDs and inter-simple sequence repeat (ISSR) fingerprints to study genetic structure within *A. hypogaea*. The study found highly polymorphic markers, 42.7% for the RAPD markers and 54.4% for the ISSR markers. A reproducibility study was conducted for the ISSR markers and only robust bands were used in the study. A dendrogram revealed that the markers were capable of separating the different varieties and subspecies of *A. hypogaea*. Only 13 cultivated accessions were included in the sample population with a single individual representing each accession. He and Prakash (1997) looked for polymorphic AFLP markers and found an average of 3.96 polymorphic bands per primer pair. Again, as in the RAPD data from Subramanian et al. (2000), this low level of polymorphism is not ideal and may be due to random error. No reproducibility study was mentioned in the published manuscript. Simple sequence repeats (SSRs) were developed by Hopkins et al. (1999) for *A. hypogaea* and seven highly reproducible, polymorphic markers were identified. These markers have not been used to study the genetic diversity within a large population of *A. hypogaea* accessions as of yet.

AFLP technology has proven to be a useful molecular tool in characterizing biodiversity for numerous plant populations including, but not limited to, sunflower (Quagliaro et al., 2001), rice (Zhu et al., 1998), cotton (Abdalla et al., 2001), rapeseed (Lombard et al., 2000), cassava (Wong et al., 1999), *Arabidopsis* (Breyne et al., 1999), *Lactuca* spp. (Hill et al., 1996), soybean (Maughan et al., 1996; Powell et al., 1996), potato (Milbourne et al., 1997), *Camellia sinensis* (Paul et al., 1997), barley (Schut et al., 1997), *Vitis vinifera* L. (Cervera et al., 1998), maize (Pejic et al., 1998), and wheat (Barrett and Kidwell, 1998). It was first envisioned by Vos et al. (1995) for use as a new technique for DNA fingerprinting. The basic methodology behind AFLPs consists of three main steps. The first step, the digestion-ligation step, involves digestion of genomic DNA with two restriction enzymes to produce a mixture of DNA fragments of varied sizes which have overhangs corresponding to the restriction enzymes used in the digestion. These fragments are then ligated to adaptors, double-stranded oligos approximately 20 base pairs in length which have overhangs that are homologous to the overhangs on the DNA fragments. The second step, the pre-selective reaction, involves amplificatin of these ligated DNA fragments with primers that are homologous to the adaptor sequences and extended by one selective base. This amplification allows for the fragments to be visualized through gel electrophoresis. However, due to the large number of fragments produced, a smear of DNA will be seen. In order to reduce the number of fragments to a sufficient level to allow for visualization of distinct, unique bands, a selective amplification is performed which comprises the third step. The primers in the selective amplification are homologous to the adaptors sequences with an additional 2-3 base pairs added to the 3' end. Thus, only DNA fragments that have bases homologous to this 2-3 base pair extension will be amplified resulting in a reduction in the number of DNA fragments. The degree to which the number of fragments are reduced is a function of the number of bases contained in the primer extension. Usually small genomes (5×10^7 to 6×10^8 base pairs) work best using a two extension primer, whereas larger genomes (5×10^8 to 5×10^9 base pairs) usually require a three extension primer. Selective reactions can be electorophoresed on a polyacrylamide gel and fluorescently detected using a DNA sequencer. The software used in the collection and analysis of the data automatically sizes DNA fragments based on the migration of a known size standard that is loaded with each sample. This method is useful in characterizing genetic diversity because it is simple, automated, requires small amounts of DNA, can produce numerous loci for analysis, and does not require prior knowledge of DNA sequence.

The analysis of AFLP data is straightforward in theory but can become tedious in reality. The data are scored in such a way that the presence of a particular sized band is represented by a “1" and the absence of a band is represented by a “0". Although Vuylsteke et al. (1999) argues that band intensity can be used to distinguish homozygotes from heterozygotes, AFLPs are commonly used as dominant markers.

SSRs, tandemly repeated units of DNA <6 base pairs in length, are considered one of the best molecular markers for studying genetic diversity that we have to date. They are highly reproducible, PCR-based, co-dominant markers that can be adapted to automation (Mitchell et al., 1997). The main drawbacks are the initial expense incurred in marker development and the lack of transferability to other species. Di- and trinucleotides are the most commonly observed repeats with (AT)n being particularly abundant in plants (Ma et al., 1996). The number of repeat units for a particular SSR locus may vary among different individuals within a single species. This high degree of variability can be used in the construction of phylogenetic trees and to build a fingerprint of a particular cultivar.

CHAPTER 2

ANALYSIS OF GENETIC DIVERSITY AND STRUCTURE IN THE CULTIVATED PEANUT, *ARACHIS HYPOGAEA*, USING SIMPLE SEQUENCE REPEATS[1]

---

[1] Newman, M.L., R.E. Dean, M.S. Hopkins and R.N. Pittman. To be submitted to *Euphytica*.

INTRODUCTION

Conservation and evaluation of *Arachis* germplasm is necessary in order to maintain the genetic variation of the genus and ensure the crop's future. The cultivated peanut, *Arachis hypogaea*, displays a wide range of variation in morphology and resistance to various diseases and insects, but demonstration of genetic variability through the use of molecular markers has been limited. Most studies, which have used techniques including Restriction Fragment Length Polymorphisms (RFLP), Randomly Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphisms (AFLP) and isozymes, have been successful at showing variation in the wild species of *Arachis* but not in the cultivated peanut (He and Prakash, 1997; Garcia et al., 1995, 1996; Stalker et al., 1994; Lacks and Stalker, 1993; Paik-Ro et al., 1992; Halward et al., 1991, 1992; Kochert at al., 1991). Six simple sequence repeat (SSR) markers were identified and found to be successful in detecting genetic variation in the cultivated peanut (Hopkins et al., 1999). In addition, a seventh polymorphic SSR has also been isolated (unpublished data). To date, these markers had not been used to study the genetic diversity within a large population of *A. hypogaea* accessions.

Perhaps the best molecular markers for studying genetic diversity that are currently available are SSRs. They are highly reproducible, PCR-based, co-dominant markers that can readily be adapted to automation (Mitchell et al., 1997). The major drawbacks are the initial expense incurred in marker development and the lack of transferability to other species. SSRs are tandemly repeated units of DNA <6 base pairs in length. Di- and trinucleotides are the most commonly used with $(AT)_n$ being particularly abundant in plants (Ma et al., 1996).

In the present study, we investigated the partitioning of and level of genetic variability in a set of 44 accessions of cultivated peanut representing the six recognized botanical varieties. Ten individuals per accession were included to assess within accession variability - an important concern for curation of a germplasm collection.

MATERIALS AND METHODS

Plant Material\DNA Extraction

Plant material was obtained from the USDA Plant Genetic Resources Conservation Unit located in Griffin, Georgia. Accessions chosen for the study were landraces from Latin America. These included a total of 44 accessions; 25 from *Arachis hypogaea* var. *hirsuta*, 11 from var. *peruviana*, and two from each variety of *hypogaea, vulgaris, fastigiata,* and *aequatoriana* (Table 1.1). Ten seeds from each accession were germinated and grown in the dark at 37^0C for 7 to 21 days. The cotyledons were removed and the etiolated seedlings stored at -70^0C. Total genomic DNA was extracted from seedlings according to the CTAB procedure (Saghai-Maroof et al., 1984).

SSR Protocol

Six primer pairs shown previously to amplify polymorphic SSR regions in peanut (Hopkins, et al., 1999) plus another newly identified SSR primer pair (unpublished data) designated as A1-193 were used in the study. Primers labeled at the 5' end of each forward primer with a fluorescent dye, either 6-carboxyfluorescein (6-FAM), tetrachloro-6-carboxyfluorescein (TET) or hexachloro-carboxyfluorescein (HEX) were obtained commercially (Operon and Research Genetics). Two multiplex sets (Table 1.2) plus individual primer sets were used to perform PCR reactions. In cases of unsuccessful amplification of one of the primer pairs in a multiplex set, single- plex PCR were performed. Each 25 uL reaction contained 1X Mg-free PCR Buffer (Perkin-Elmer), 1.5 mM $MgCl_2$, 0.25 mM dNTPs, 1.0 U of Taq polymerase, 5 pmol of each primer and 25 ng of template DNA. Cycling parameters included an initial denaturation of template DNA at 95^0C for 4 min, followed by 25 cycles of 95^0C for 1 min, 55^0C for 2 min, 72^0C for 2 min with an extension time of 72^0C for 10 min during the last cycle. Samples were electorophoresed on 36cm well-to-read, 4.25% (w/v) denaturing acrylamide:bisacrylamide (29:1) gels using an ABI PRISM 377 DNA Sequencer (PE Applied Biosystems) according to manufacturer's protocol. Fragment sizes were determined using Genescan ver. 2.1. Sizes of some SSR fragments could not be unequivocally put into a locus size category, i.e. they did not fall into two/three base pair increments. When this occurred, fragments were sequenced to determine exact fragment size.

Statistical Analyses

Microsat (E. Minch) was utilized to calculate number of alleles per locus and observed heterozygosities. Since *A. hypogaea* is a tetraploid, data was formatted in such a way that each primer set was considered twice assuming two homozygotes, one for the A and one for the B genome. Primer set ah4-026 which contained three alleles per individual was an exception. In this case, one allele at 155 bp was consistent for all individuals and assumed homozygous and the other two alleles present scored as heterozygous. A set of 100 bootstrapped distance matrices based on proportion of alleles shared was constructed using Microsat and input into Phylogeny Inference Package (Phylip) Version 3.5c (Felsenstein, 1993) to construct a phylogenetic tree using the Neighbor Joining option. Analysis of Molecular Variance (AMOVA) (Weir, B.S. and Cockerham, C.C. 1984. Excoffier, L., Smouse, P., and Quattro, J. 1992. Weir, B. S., 1996) was performed using Arlequin ver. 2.000 (Schneider et al., 2000). A geographic distance matrix was constructed using longitude and latitude coordinates obtained from the Germplasm Resources Information Network (GRIN) (USDA, 2003). The following formula was used to calculate the distance (km) between two accessions, where both longitude (lon) and latitude (lat) values are in radians:

$$distance = \cos(\sin(lat1) * \sin(lat2)) + \cos(lat1) * \cos(lat2) * \cos(lon1 - lon2)$$

A Mantel Test was performed using Isolation By Distance (IBD) software (Bohonak, 2002) to determine if there was a significant correlation between the genetic distance matrix and the geographic distance matrix.

RESULTS

The average number of alleles across accessions varied from 1.5 for primer set Ah4-024 to 3.9 for primer set Ah4-026 (Table 1.3). Average heterozygosity (H_O) varied according to SSR primer set and ranged from 0.13 for primer set Ah4-024 to 0.69 for primer set Ah4-026. One of the primer sets, LEC-1, was excluded from the study due to the lack of reproducibility and high number of bands produced. Additionally, primer set Ah6-125 was found to be monomorphic and excluded from the analysis.

The dendogram derived from the SSR data shows that the accessions cluster together better according to geographical location rather than variety (Figure 1.1). Five *hirsuta* accessions clustered together and were are all from the state of Pichincha, Ecuador in a remote area of high elevation. All the *hirsuta* accessions from Mexico grouped together into two major clades. Two accessions of *hirsuta* from Imbabura, Ecuador grouped into a separate clade. Two *peruviana* accessions from Beni, Bolivia grouped together but separate from a *fastigiata* from Beni, Bolivia.

The AMOVA (Table 1.4) revealed that the majority of the genetic variation (66.11%) was partitioned among populations with a considerable amount of variation (28.02%) also being partitioned within populations when groups and populations were defined based on taxonomy. When groups and populations were based on geographic location (Table 1.5), the majority of genetic diversity (61.98%) was partitioned among groups (or geographic location).

A Mantel test using the entire data set revealed no significant correlation between genetic distance and geographic distance (R-value = 0.149). The data set was reduced to include only those accessions that grouped phylogenetically into discrete geographic clades as noted above. Again, there was no significant correlation between geographic distance and genetic distance (R-value = 0.112).

DISCUSSION

In trying to understand the genetic diversity within *A. hypogaea*, the main obstacle has been the lack of polymorphisms detected. Initial studies attempted to utilize isozymes to study the genetic variability within *A. hypogaea* but were unsuccessful in finding any variation (Stalker et al., 1994; Lu and Pickersgill, 1993). Restriction fragment length polymorphisms (RFLPs) have revealed a high degree of diversity within the wild species of section *Arachis*, but no significant polymorphisms were found in *A. hypogaea* (Kochert et al., 1991; Halward et al., 1991; Halward et al., 1993; Paik-Ro, et al., 1992). Hilu and Stalker (1995) were also unable to find any polymorphisms using random amplified polymorphic DNA (RAPD). V. Subramanian and colleagues (2000) found a marginal amount of polymorphism in the cultivated species using RAPDs with 27 polymorphic bands out of 408 total bands or 6.6% polymorphism. However, taking into consideration the lack of reproducibility associated with RAPDs (Karp et al., 1996), this level of polymorphism is not ideal. More recently, Raina et al. (2001) used RAPDs and inter-simple sequence repeats (ISSR) fingerprints to study genetic structure within *A. hypogaea*. The study found highly polymorphic markers, 42.7% for the RAPD markers and 54.4% for the ISSR markers. A reproducibility study was conducted for the ISSR markers and only robust bands were used in the study. A dendogram revealed that the markers were capable of separating the different varieties and subspecies of *A. hypogaea*. Only 13 cultivated accessions were included in the sample population with a single individual representing each accession. He and Prakash (1997) looked for polymorphic AFLP markers and found an average of 3.96 polymorphic bands per primer pair. Again, as in the RAPD data from Subramanian et al. (2000), this low level of polymorphism is not ideal and may be due to random error.

The present study uses highly robust, polymorphic SSR markers to determine genetic diversity and structure in a select group of cultivated peanut accessions available from the USDA National Plant Germplasm System. The amount of genetic variability detected using the SSR markers in the study was high considering the small number of SSR markers available. Most of the markers had a high polymorphic information content as evident from the high values obtained for observed heterozygosity for markers Ah4-026 and Ah4-020. This data would suggest that SSR marker technology may be the marker system of choice for use in the cultivated peanut due to the high degree of polymorphism found when compared to that of other marker systems.

It is a difficult task to balance the need for acquiring new genetic resources with the high cost and demand of maintaining these resources. As more accessions are entered into the germplasm collection from exploration efforts, this creates an increased demand on the germplasm system to process, document, regenerate, and store these new accessions. Future collection trips should be done in such a manner that maximizes the amount of genetic diversity gained with a reduced number of samples collected. The dendogram produced using the SSR data demonstrates that the genetic diversity within the cultivated peanut is partitioned according to geographic location rather than among subspecies taxa. Although the Mantel test did not find any significant correlation between geographic distance and genetic distance, this may be due to the fact that the distance calculated was based on actual distance and does not take landforms and other barriers such as streets into consideration. Taking into further consideration the results from the AMOVA, it can be seen that more genetic diversity may be gained when collection trips focus on gathering new accessions from diverse geographic locations rather than focusing collection efforts solely on obtaining different botanical varieties or morphological characteristics associated with taxonomy. By prioritizing collection efforts, the overall operation of the germplasm system is made more efficient.

Within accession variability is another important consideration in germplasm maintenance. New germplasm accessions will often by obtained from a local market which may contain a mixture of genotypically disimilar individuals. Accessions with a high degree of variability will require different regeneration processes than ones with a highly uniform genotypic profile. AMOVA results indicate that there is a significant amount of within accession variability due to the large partitioning of variability within accessions, defined as populations. Regeneration of highly uniform accessions will require a much smaller sample size than genotypically heterogeneous accessions in order to assure preservation of the genetic content contained in the accessions. This will have a significant impact on regeneration plans each year in that larger sample size requirements will necessitate increased field space, labor demands and cold room storage space. By knowing beforehand the genetic variability contained in individual accessions and the steps required to capture this variability, regeneration will become a more efficient process. In addition, within accession variability may indicate a need to split a single accession into multiple accession.

This study highlights some of the valuable information gained through molecular genetic studies that can assist in the overall operation of a germplasm system. The germplasm collection for *A. hypogaea* maintained in Griffin, Georgia is a valuable collection but a large collection. To ensure preservation of this valuable genetic resource for future generations, it will be required that the collection of new material and regeneration of current holdings be made as efficient as possible. Although significant variation was found in the sample population using the SSR markers currently available, there is still a need to develop more SSR markers to discriminate between more closely related accessions and to fully evaluate the amount of genetic diversity in *A. hypogaea*.

REFERENCES

Bohonak, A.J. 2002. IBD (Isolation By Distance): a program for analyses of isolation by distance. J of Heredity. In press.

Felsenstein, J. 1989. PHYLIP - Phylogeny Inference Package (Version 3.2). Cladistics 5: 164-166.

Garcio G.M., Stalker H.T. and Kochert G. 1995. Introgression analysis of an interspecific hybrid population in peanuts (*Arachis hypogaea* L.) using RFLP and RAPD markers. Genome. 38:166-176.

Halward T.M., Stalker H.T., LaRue E.A., Kochert G. 1991. Genetic variation detectable with molecular markers among adapted germ-plasm resources of cultivated peanut and related wild species. Genome, 34:1013-1020.

Halward T.M., Stalker H.T., LaRue E.A., Kochert G. 1992. Use of single-primer DNA amplifications in genetic studies of peanut (*Arachis hypogaea* L.). Plant Mol. Biol. 18:315-325.

He G. and Prakash C.S. 1997. Identification of polymorphic DNA markers in cultivated peanut (*Arachis hypogaea* L.). Euphytica. 97:143-149.

Hilu K.W. and Stalker H.T. 1995. Genetic relationships between peanut an wild species of *Arachis* sect. *Arachis* (Fabaceae): evidence from RAPDs. Pl. Syst. Evol. 198:167-178.

Hopkins M.S., Casa A.M., Wang T., Mitchell S.E., Dean R.E., Kochert G.D., Kresovich S. 1999. Discovery and characterization of polymorphic simple sequence repeats (SSRs) in peanut. Crop Science. 39(4):1243-1247.

Karpa A., Seberg O., Buiatti M. 1996. Molecular Techniques in the Assessment of Botanical Diversity. Annals of Botany. 8:143-149.

Kochert G., Halward T., Branch W.D., Simpson C.E. 1991. RFLP variability in peanut *(Arachis hypogaea* L.) cultivars and wild species. Theor. Appl. Genet. 5 65-570.

Lu J., and Pickersgill B. 1993. Isozyme variation and species relationships in peanut and its wild relatives (*Arachis* L.: *Leguminosae*). Theor. Appl. Genet. 85:550-560.

Ma ZQ, Roder M, Sorells ME. 1996. Frequencies and sequence characteristics of di-, tri-, and tetra-nucleotide microsatellites in wheat. Genome. 39:123-130.

Minch, E. Microsat. http://hpgl.stanford.edu/projects/microsat/microsat.html

Mitchell S.E., Kresovich S., Jester C.A., Hernandez C.J., Szewc-McFadden. 1997. Application of multiple PCR and fluorescence-based, semi-automated, allele sizing technology for genotyping plant genetic resources. Crop Sci. 37:617-624.

Paik-Ro O.G., Smith R.L., Knauft D.A. 1992. Restriction fragment length polymorphism evaluation of six peanut species within the *Arachis* section. Theor. Appl. Genet. 84:201-208.

Raina S.N., Rani V., Kojima T., Ogihara Y., Singh K.P., and Devarumath R.M. 2001. RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity, varietal identification, and phylogenetic relationships in peanut (*Arachis hypogaea*) cultivars and wild species. Genome. 44:763-772.

Saghai-Maroof M.A., Soliman K.M., Jorgensen R.A.,Allard R.W. 1984. Ribosomal DNA spacer-length polymorphism in barley: Mendelian inhertance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA. 81:8014.

Schneider S., Roessli D., Excoffier L. 2000. Arlequin ver. 2.000. A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Stalker HT, Phillips T.D., Murphy J.P. and Jones T.M. 1994. Variation of isozyme patterns among *Arachis* species. Theor Appl Genet. 87:746-755.

Subramanian V., Gurtu S., Nageswara Rao R.C., and Nigam S.N. 2000. Identification of DNA polymorphism in cultivated groundnut using random amplified polymorphic DNA (RAPD) assay. Genome. 43:656-660.

USDA, ARS, National Genetic Resources Program. Germplasm Resources Information Network - (GRIN). 2003. [Online Database] National Germplasm Resources Laboratory, Beltsville, Maryland. Available: http://www.ars-grin.gov.

Weir, B.S. 1996. Genetic Data Analysis II: Methods for Discrete Population Genetic Data. Sinauer Assoc., Inc., Sunderland, MA, USA.

Weir, B.S. and Cockerham,C.C. 1984. Estimating F-statistics for the analysis of population structure. Evolution. 38:1358-1370.

Table 1.1. List of *Arachis hypogaea* accessions included in the study.

Accession ID	**Variety**	**Country of Origin**
Grif 12529	*aequatoriana*	Ecuador
Grif 12536	*aequatoriana*	Ecuador
PI 493722	*fastigiata*	Argentina
PI 497285	*fastigiata*	Bolivia
Grif 12577	*peruviana*	Ecuador
Grif 13809	*peruviana*	Ecuador
PI 502015	*peruviana*	Peru
PI 502043	*peruviana*	Peru
PI 502045	*peruviana*	Peru
PI 502053	*peruviana*	Peru
PI 502088	*peruviana*	Peru
PI 502096	*peruviana*	Peru
PI 540833	*peruviana*	Bolivia
PI 590455	*peruviana*	Brazil
PI 599611	*peruviana*	Bolivia
PI 493808	*vulgaris*	Argentina
PI 602088	*vulgaris*	India
Grif 12506	*hirsuta*	Ecuador
Grif 12510	*hirsuta*	Ecuador
Grif 12511	*hirsuta*	Ecuador
Grif 12513	*hirsuta*	Ecuador
Grif 13794	*hirsuta*	Ecuador
Grif 13799	*hirsuta*	Ecuador
Grif 14060	*hirsuta*	Guatemala
PI 280688	*hirsuta*	Mexico
PI 501296	*hirsuta*	Peru
PI 501297	*hirsuta*	Peru
PI 576612	*hirsuta*	Mexico
PI 576613	*hirsuta*	Mexico
PI 576614	*hirsuta*	Mexico
PI 576615	*hirsuta*	Mexico
PI 576616	*hirsuta*	Mexico
PI 576617	*hirsuta*	Mexico
PI 576618	*hirsuta*	Mexico
PI 576621	*hirsuta*	Mexico
PI 576633	*hirsuta*	Mexico
PI 576634	*hirsuta*	Mexico
PI 576635	*hirsuta*	Mexico
PI 576636	*hirsuta*	Mexico
PI 576637	*hirsuta*	Mexico
PI 576638	*hirsuta*	Mexico
PI 602356	*hirsuta*	Ecuador

Table 1.1. cont.

Accession ID	**Variety**	**Country of Origin**
PI 497290	*hypogaea*	Bolivia
PI 497443	*hypogaea*	Bolivia

Table 1.2. SSR primers for use in the analysis of *A. hypogaea* with multiplex grouping.

Primer Set	**Motif\ Length (b.p)**	**Primer Sequence**	**Multiplex**
Ah4-004	(GA)	5'-TET-CGATTTCTTTACTGAGTGAG-3'(F)	A
	90	5'-ATTTTTTGCTCCACACA-3'(R)	
Ah4-026	(CT)	5'-FAM-TGGAATCTATTGCTCATCGGCTCTG-3'(F)	A
	180	5'-CTCACCCATCATCATCGTCACATT-3'(R)	
Lec-1	(AT)	5'-TET-CAAGCATCAACAACAACGA-3'(F)	Single
	250	5'-GTCCGACCACATACAAGAGTT-3'(R)	
A1-193	(GA)	5'-TET-CTTGCTGAAGGCAACTCCTACG-3'(F)	A
	445	5'-TCGGTTTGTCTCTTTGGTCACTC-3'(R)	
Ah4-020	(GA)	5'-HEX-ACCAAATAGGAGAGAGGGTTCT-3'(F)	B
	210	5'-CTCTCTTGCTGGTTCTTTATTAACTC-3'(R)	
Ah4-024	(ATA)	5'-TET-TTCTGATTTTAGTAGTCTTCTTTCACT-3'(F)	B
	405	5'-CTCCTTAGCCACGGTTCT’-3'(R)	
Ah6-125	(TTC)	5'-TET-TCGTGTTCCCGATTGCC-3'(F)	Single
	190	5'-GCTTTGAACATGAACATGCC-3'(R)	

Table 1.3. Observed heterozygosity (H_o) and number of alleles (A). Values are listed for each locus in each cultivar with averages across loci and cultivars.

	Ah4-004		**Ah4-026**		**A1-193**		**Ah4-020**		**Ah4-024**	
Cultivar	**H_o**	**A**	**H_o**	**A**	**H_o**	**A**	**H_o**	**A**	**H_o**	**A**
Grif 12529	0.3750	2	0.7083	4	0.0000	1	0.6719	4	0.0000	1
Grif 12536	0.0000	1	0.7311	6	0.0000	1	0.5000	2	0.3200	2
PI 493722	0.0000	1	0.6667	3	0.2188	2	0.5000	2	0.1800	2
PI 497285	0.6420	3	0.6667	3	0.4444	2	0.5000	2	0.1975	2
Grif 12577	0.2188	2	0.6867	4	0.1653	2	0.5429	3	0.0000	1
Grif 13809	0.0000	1	0.6659	3	0.1800	2	0.4986	2	0.1653	2
PI 502015	0.2449	2	0.6667	3	0.0000	1	0.4959	2	0.2449	2
PI 502043	0.1975	2	0.6886	4	0.4938	2	0.5000	2	0.1975	2
PI 502045	0.0000	1	0.6667	3	0.0000	1	0.5000	2	0.0000	1
PI 502053	0.4260	2	0.7044	5	0.1800	2	0.4938	2	0.0000	1
PI 502088	0.4861	2	0.6657	3	0.4444	2	0.4983	2	0.0000	1
PI 502096	0.1800	2	0.7200	4	0.5000	2	0.5000	2	0.0000	1
PI 540833	0.0000	1	0.7188	4	0.0000	1	0.5000	2	0.0000	1
PI 590455	0.0000	1	0.6667	3	0.0000	1	0.5000	2	0.0000	1
PI 599611	0.1975	2	0.6886	4	0.0000	1	0.5000	2	0.0000	1
PI 493808	0.0000	1	0.7022	4	0.4800	2	0.5000	2	0.0000	1
PI 602088	0.0000	1	0.6670	3	0.0000	1	0.5000	2	0.2449	2
Grif 12506	0.3400	3	0.7022	4	0.0000	1	0.5000	2	0.3457	2
Grif 12510	0.0000	1	0.7044	5	0.1800	2	0.5000	2	0.0000	1
Grif 12511	0.0000	1	0.6667	3	0.3200	2	0.4986	2	0.0000	1
Grif 12513	0.0000	1	0.6867	4	0.0000	1	0.5000	2	0.3200	2
Grif 13794	0.2975	2	0.6667	3	0.0000	1	0.4970	2	0.0000	1
Grif 13799	0.1800	2	0.6667	3	0.1975	2	0.5494	3	0.0000	1
Grif 14060	0.3400	3	0.7222	5	0.1800	2	0.4986	2	0.5278	5
PI 280688	0.0000	1	0.6670	3	0.0000	1	0.5000	2	0.3200	2
PI 501296	0.2188	2	0.6939	4	0.0000	1	0.5000	2	0.2188	2
PI 501297	0.0000	1	0.6670	3	0.0000	1	0.5000	2	0.0000	1
PI 576612	0.0000	1	0.7022	4	0.1800	2	0.0000	1	0.0000	1
PI 576613	0.1800	2	0.6867	4	0.0000	1	0.5000	2	0.1800	2
PI 576614	0.0000	1	0.6886	4	0.3457	2	0.2778	2	0.0000	1
PI 576615	0.1975	2	0.6886	4	0.1975	2	0.1975	2	0.1975	2
PI 576616	0.3200	2	0.7600	5	0.0000	1	0.5000	2	0.0000	1
PI 576617	0.0000	1	0.7420	5	0.3200	2	0.5000	2	0.3200	2
PI 576618	0.4898	2	0.7120	4	0.0000	1	0.0000	1	0.0000	1
PI 576621	0.0000	1	0.6667	3	0.1800	2	0.5000	2	0.0000	1
PI 576633	0.2975	2	0.6867	4	0.1800	2	0.4986	2	0.1800	2
PI 576634	0.0000	1	0.7215	4	0.5679	3	0.1800	2	0.0000	1
PI 576635	0.3750	2	0.7188	4	0.2188	2	0.4898	2	0.3750	2
PI 576636	0.0000	1	0.6867	4	0.1800	2	0.5000	2	0.0000	1
PI 576637	0.5400	3	0.7333	5	0.5600	3	0.4082	2	0.4800	2

Table 1.3. cont.

	Ah4-004		**Ah4-026**		**A1-193**		**Ah4-020**		**Ah4-024**	
Cultivar	**H_o**	**A**	**H_o**	**A**	**H_o**	**A**	**H_o**	**A**	**H_o**	**A**
PI 576638	0.3200	2	0.6867	4	0.0000	1	0.5000	2	0.0000	1
PI 602356	0.3200	2	0.6867	4	0.1800	2	0.5450	3	0.3200	2
PI 497290	0.1800	2	0.7051	4	0.4938	3	0.5000	2	0.3457	2
PI 497443	0.3200	2	0.7222	5	0.0000	1	0.4688	2	0.0000	1
Average	**0.1716**	**1.6**	**0.6930**	**3.9**	**0.1686**	**1.6**	**0.4513**	**2.0**	**0.1262**	**1.5**

Table 1.4. AMOVA Analysis According to Taxonomy. AMOVA Analysis (Weir, B.S. and Cockerham, C.C. 1984. Excoffier, L., Smouse, P., and Quattro, J. 1992. Weir, B. S., 1996) was performed using Arlequin software ver. 2.000 (Schneider et al., 2000). The groups are defined as the two subspecies of *A. hypogaea* (subsp. *fastigiata* and subsp. *hypogaea*) and the populations are defined as single accessions containing multiple individuals per accession.

Source of variation	**d.f.**	**Sum of squares**	**Variance components**	**Percentage of variation**
Among groups	1	104.221	0.17832 Va	5.88
Among populations within groups	42	1582.392	2.00595 Vb	66.11
Within populations	764	649.557	0.85021 Vc	28.02
Total		**807**	**2336.170**	**3.03448**

Table 1.5. AMOVA Analysis According to Geography. AMOVA Analysis (Weir, B.S. and Cockerham, C.C. 1984. Excoffier, L., Smouse, P., and Quattro, J. 1992. Weir, B. S., 1996) was performed using Arlequin software ver. 2.000 (Schneider et al., 2000). The groups are defined as geographic areas based on the dendogram and the populations are defined as single accessions containing multiple individuals per accession.

Source of variation	**d.f.**	**Sum of squares**	**Variance components**	**Percentage of variation**
Among groups	4	723.061	2.05619 Va	61.98
Among populations within groups	19	176.700	0.44343 Vb	13.37
Within populations	434	354.846	0.81762 Vc	24.65
Total		**457**	**1254.607**	**3.31723**

Figure 1.1. Dendogram depicting genotypic relationships among various *A. hypogaea* varieties using SSR data.

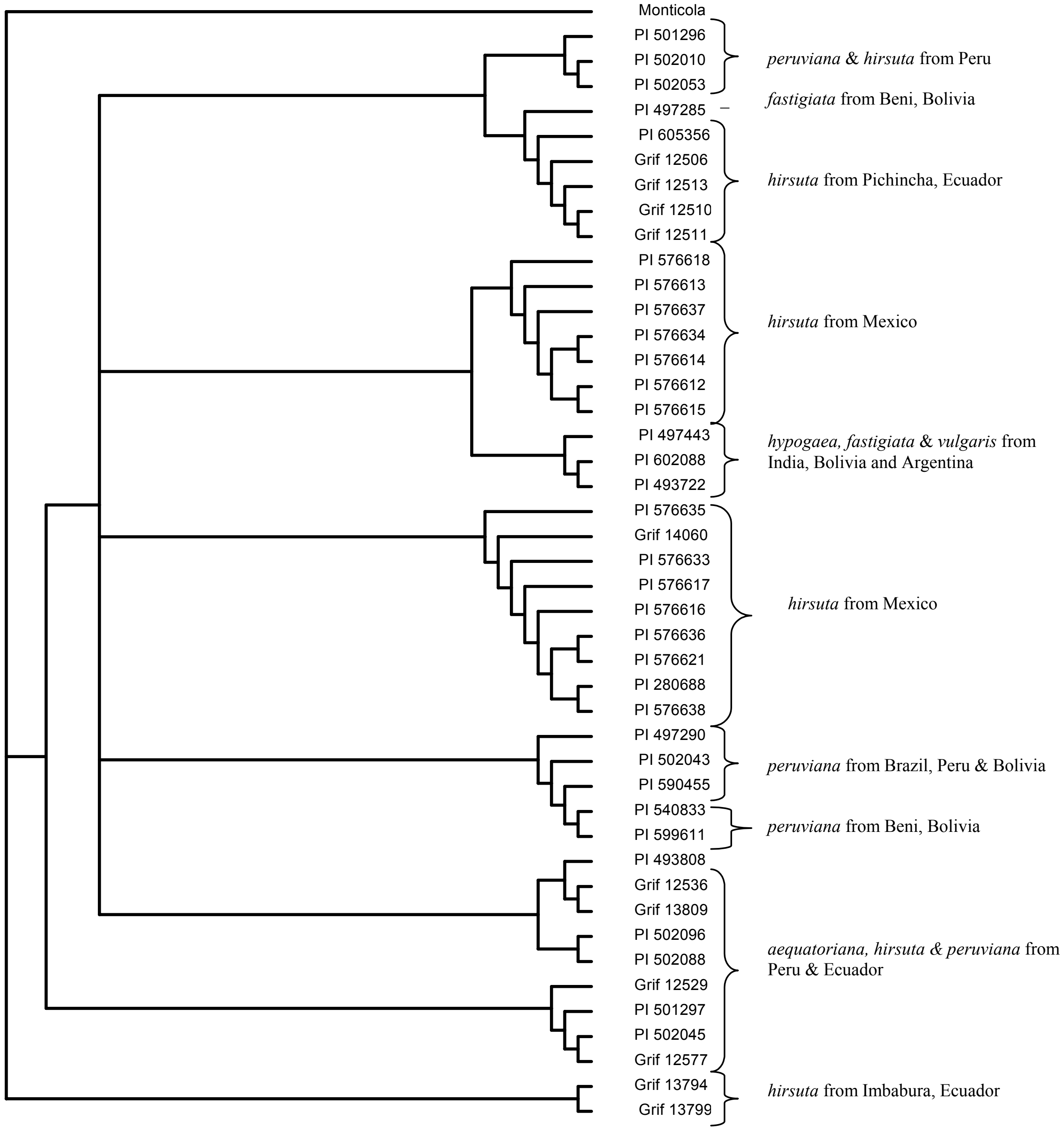

Figure 1.1. Dendogram depicting genotypic relationships among various *A. hypogaea* varieties using SSR data.

CHAPTER 3

PARTITIONING OF GENETIC DIVERSITY IN THE GENUS ARACHIS ASSESSED BY AFLP ANALYSIS

---

[1] Newman, M.L., R.E. Dean and R.N. Pittman. To be submitted to *Euphytica*.

INTRODUCTION

The genus *Arachis* is a diverse taxon containing both the cultivated peanut, *A. hypogaea*, and its wild relatives. The genus contains 69 recognized species (Krapovickas and Gregory, 1994) which has been divided into nine sections based on morphological and interspecific hybridization studies (Gregory and Gregory, 1979). In addition, Valls and Simpson (1997) have identified 11 new species which have yet to be fully described. Currently, the only species of significant economic importance is *A. hypogaea* which is an important oilseed crop grown worldwide. However, other species have been utilized. For example, *A. glabrata* and *A. pintoi* are used as forage crops (Maass et al. 1993); *A. repens* is used as an ornamental and as a groundcover (Gimenes et al., 2000); and *A. villosulicarpa* is used as a food crop in the Mato Grosso State of Brazil (Galgaro L. et al., 1997).

A. hypogaea is an allotetraploid that is believed to have arisen through the hybridization of two wild diploid *Arachis* species (Smartt and Stalker, 1982). Much debate has centered on the identification of the progenitor species, but it is generally agreed that the progenitors are most likely species found within the section *Arachis* (Table 2.1). Therefore, most of the studies characterizing the genetic and morphological features of the different wild *Arachis* species has primarily focused on species within this section. There have been, however, several studies focused on species in the other sections (Singh et al., 1994; Stalker et al.,1994; Klozova et al., 1983 a,b; Cherry, 1975). These studies have provided useful insights into the relationship between the species within various sections of the genus *Arachis* and demonstrate the need for further molecular studies investigating the genetic diversity and relationships that exist within the genus.

AFLP technology has proven to be a useful molecular tool in characterizing biodiversity for numerous plant populations including, but not limited to, sunflower (Quagliaro et al., 2001), rice (Zhu et al., 1998), cotton (Abdalla et al., 2001), rapeseed (Lombard et al., 2000), cassava (Wong et al., 1999), *Arabidopsis* (Breyne et al., 1999), *Lactuca* spp. (Hill et al., 1996), soybean (Maughan et al., 1996; Powell et al., 1996), potato (Milbourne et al., 1997), *Camellia sinensis* (Paul et al., 1997), barley (Schut et al., 1997), *Vitis vinifera* L. (Cervera et al., 1998), maize (Pejic et al., 1998), and wheat (Barrett and Kidwell, 1998). It was first envisioned by Vos et al. (1995) for use as a new technique for DNA fingerprinting. It is an excellent choice for looking at diversity across species because it is not a species specific marker, as is the case many times with simple sequence repeats and other markers. It also requires no prior knowledge of DNA sequence, requires small quantities of DNA and is highly reproducible (Karp et al., 1996).

The overall goal of this study is twofold. First, AFLP marker technology will be used to assess the genetic variability within and between the majority of the sections of the genus *Arachis*. Although this type of investigation into the genus *Arachis* has been done with other technologies such as seed protein profiles and isozymes (Singh et al. 1994; Stalker et al.; 1994 Klozova et al. 1983 a,b; Cherry, 1975), this study will be more comprehensive in that it will survey seven of the nine sections with one marker set. Secondly, the AFLP data will be used to determine relationships of wild species, particularly in section *Arachis,* to cultivated peanut to shed further light on the origin of the cultivated peanut.

METHODS AND MATERIALS

Plant Material

Plant material was collected from wild *Arachis* accessions maintained at the USDA's Plant Genetic Resources Conservation Unit in Griffin, Georgia. Three accessions per species were used in the study when plant material was available (USDA, GRIN, 2003). Due to lack of availability, many species had only one or two accessions representing the species (Table 2.1) and some species in the genus were not represented at all. A small subset of accessions was used to assess within accession variability (PI 262133, PI 405933, PI 468197, PI 468199, Grif 7724, PI 468327, PI 468329, PI 468321, PI 468326, PI 468196). For this subset, six to twelve individuals per accession were utilized. DNA was also isolated from *Stylosanthes* for use as an outgroup in phylogenetic analysis.

DNA Extraction

Newly unfolded leaf material (~200 mg) was extracted using an E.Z.N.A® Plant DNA Miniprep Kit (Omega Bio-Tek**,** Inc.). DNA quality for each extraction was verified using agarose gel electrophoresis. Samples were quantitated using a TKO 100 (Hoefer Scientific Instruments) fluorometer.

Digestion/Ligation Reactions

An AFLP profile was generated for each sample according to a protocol adapted from Zhu et al. (1998) and Herbergs et al. (1999). Briefly, DNA was digested with Mse1 and Pst1 at 37ºC for one hour followed by enzyme inactivation for 20 min at 80ºC in a mixture containing 200ng of DNA, 5U of *Ms*I and 10U of *Pst*I (New England Biolabs), 1X NEB2 Buffer (supplied with enzyme), 0.075 mg Bovine Serum Albumin in a total volume of 50uL. A 10 uL ligation mix containing 11U of T4 DNA Ligase (New England Biolabs), 1X NEB2 Buffer (supplied with enzyme), 6mM ATP and 5 pmol of each Pst1 adaptor (Table 2.2) and 50 pmol of each Mse1 adaptor was then added to the digestion and incubated at 16ºC overnight. Before adding the adaptors to the ligation mixture, each pair of adaptors were mixed together, incubated at 95ºC for 5 minutes, and cooled at room temperature for 10 minutes.

Pre-selective Reactions

Pre-selective reactions contained 1X PCR Buffer (supplied with Taq polymerase), $MgCl_2$, 0.16mM dNTPs, 1U of Taq polymerase, 50 pmol of each pre-selective primer (Table 2.2) and 1.0uL of digestion-ligation reaction in a total volume of 20uL. The reactions were denatured at 95ºC for 5 minutes followed by 30 cycles of 94ºC for 30s, 50ºC for 30s, and 72ºC for 60s.

Selective Reactions

Selective reactions contained 1X PCR Buffer (supplied with Taq polymerase), $MgCl_2$, 0.16mM dNTPs, 1U of Taq polymerase, 50 pmol of each selective primer (Table 2.2) and 0.5 uL of digestion-ligation reaction in a total volume of 20uL. The reactions were denatured at 95ºC for 5 minutes followed by a touchdown PCR of 94ºC for 30s, 65ºC for 30s, and 72ºC for 60s with 0.7ºC decrease in annealing temperature for an additional 20 cycles. This was followed by 24 cycles at 94ºC for 30s, 65ºC for 30s, and 72ºC for 60s. Selective reactions were electrophoresed at 2500V on a 4.25% denaturing polyacrylamide using an ABI 377 Sequencer using the "C" filter according to manufacturers protocol. Reactions were prepared for loading in a mixture containing 1.25 uL of 5:1 (formamide:50mM EDTA\30mg blue dextran), 0.5 uL of reaction, and 0.5 uL of TAMRA 500 Genescan standard (Applied Biosystems) of which 1.0 uL is loaded on the gel. A subset of selective primers was chosen based on degree of polymorphisms, reproducibility, and well-defined peak patterns obtained from data generated in an initial screen of all 20 primer pair combinations using a small subset of peanut individuals. The primer combinations chosen for study were M12/P66, M14/P75 and M14/P78 (Table 2.2).

Data Analysis

AFLP data were processed using GeneScan®, Version 3.1.2 (Applied Biosystems), to ensure accurate size calling of the peaks. Data were imported into Genotyper, Version 2.5 (Applied Biosystems), to transform data into a matrix. Using this matrix, the Phylogeny Inference Package (PHYLIP; Felsenstein, 1989) was used to generate 100 bootstrapped distance matrices using the Seqboot and RestDist programs. A consensus phylogenetic tree was constructed using the Neighbor Joining program (UPGMA option) and Consensus program. Arlequin ver. 2.000 (Schneider et al., 2000) software was used to calculate Nei's Genetic Diversity for each taxonomic section of *Arachis* and perform Analysis of Molecular Variance (AMOVA) (Weir, 1996; Excoffier, et al., 1992; Weir and Cockerham, 1984). A reproducibility study, where reactions were set up in duplicate but on different days, was conducted using a small subset of the samples to determine the error rate associated with the AFLP technique utilized in the study.

RESULTS

Nei's genetic diversity ranged from 0.278 for *Heteranthae* to 0.153 for *Caulorrhizae* (Table 2.3). There was no significant difference in the level of genetic diversity between the different sections. The AMOVA (Table 2.4) revealed that the majority of genetic diversity (78.1%) resides within sections. The AMOVA performed using the within accession study sub-population revealed that most of this variability (93.45%) resides within individual accessions.

The phylogenetic tree derived from the AFLP data was not able to discriminate accessions according to species. It did, however, group accessions according to taxonomic section (Figure 2.1). The grouping indicates that the sections, *Arachis* - annuals and perennials are closely related to each other as are *Erectoides, Rhizomatosae* and *Procumbentes*. The sections *Extranervosae* and *Heteranthae* were more distantly related to the other sections; and *Caulorrhizae* was separate from all other sections. In the reproducibility study, overall cladistic relationships remained constant.

The proportion of shared alleles was determined for all accessions in section *Arachis* compared to the cultivated species, *A. hypogaea*. Fifteen different accessions had 50% or more alleles in common with *A. hypogaea* representing species *A. monticola*, *A. batizicoi*, *A. cardenasii*, *A. cruziania*, *A. williamsii, A. diogoi, A. ipaensis, A. helodes, A. glandulifera* and *A. duranensis*. All three accessions of *A. batizicoi*, the two accessions of *A. monticola* and two of three accessions for *A. cruziania* and *A. diogoi* were included in this group. All other species in this group were represented by only one accession. *A. monticola*, PI 468199, shared the most alleles with *A. hypogaea* (64%), more than any other accession.

DISCUSSION

The level of genetic diversity contained within *Arachis* as measured by Nei's genetic diversity index is what one would expect for a self-pollinating crop such as peanut (Hamrick and Godt ,1987). The AMOVA analysis based on the AFLP profiles reveals that the majority of the genetic diversity is partitioned within species. This has implications for germplasm acquisition. It suggests that more genetic diversity may be obtained by collecting more plants within a few species rather than broadly sampling across several species and different sections.

The genus *Arachis* evolved in the Brazilian Shield and its drainage basins, which include the countries Argentina, Brazil, Bolivia, Paraguay, and Uruguay. It is these drainage basins that separated the genus long ago, isolating them, and leading to genetic drift within the genus. The sections defined in the genus *Arachis* correspond with these major drainage basins. Due to the genetic drift that occurred, the sections would most likely have different genetic make-ups which should lead to their distinction in molecular analysis. This is seen in the grouping of the sections in the phylogenetic analysis (Figure

2.1). Also, this genetic drift would most likely lead to reproductive incompatibilities between species of different sections. This is seen in a study done by Gregory and Gregory (1979). They performed 42,283 within-sectional crosses and 79,843 between-sectional crosses, to determine the interspecific cross-compatibility in the genus. Results of this study are summarized in Figure 2.2. The phylogenetic analysis performed in the current study (Figure 2.1) is in agreement with the cross-compatibility study. The relationships between the different sections, namely the grouping of section *Arachis* series annuals and perennials; *Erectoides, Rhizomatosae* and *Procumbentes*; and *Heteranthae* and *Extranervosae* are similar in both the phylogenetic analysis and the hybridization studies. In the present study, the different accessions grouped by section and did not group by species within sections. The cross-compatibility studies by Gregory and Gregory (1979) show that crosses within sections between different species were more successful than between section crosses. The fact that the species did not group together phylogenetically and the higher degree of cross-compatibility of species within sections suggest a period of increased gene flow occurred at one time in the evolutionary history of these species which may have resulted in less genetic differentiation between species within a section.

A subset of the study population was used to assess within accession genetic diversity. AMOVA analysis indicated that the level of genetic diversity within this subset is mainly within accessions. This agrees with previous findings that found a high percentage of seed to seed variability in isozyme patterns within accessions from section *Arachis* (Stalker et al., 1994). We were unable to include ten or more individuals per accession for the overall study due to the large number of accessions included. We suggest future studies focus on either a few species or a single section with more individuals per accession in order to get a clearer picture of the genetic diversity and its structure.

To investigate possible progenitors of *A. hypogaea*, the accessions that had 50% or more alleles in common with *A. hypogaea* were identified. It is interesting that *A. monticola*, which is believed by some to be either a tetraploid precursor or a subspecies of *A. hypogaea* (pers. comm., Roy Pittman), was found to share the most alleles in common with the cultivated peanut and that all accessions of *A. monticola* were included in the 50% shared alleles group. The fact that all three accessions of *A. batizicoi* were included in the 50% shared alleles group supports previous studies which suggest *A. batizicoi* as the B genome donor of *A. hypogaea* (Table 2.5). Although *A. cruziania* and *A. diogoi* have not been previously indicated as putative progenitors, two of the three accessions included in this study for these species showed 50% or more alleles in common with *A. hypogaea,* suggesting that they may be added to the list of possible progenitors. The fact that only one accession for the remaining species was included in the 50% shared alleles group suggests that their relationship with *A. hypogaea* may not be as strong and may not warrant their consideration as progenitors.

The results of this study taken collectively have implications not only for molecular studies but also studies that investigate traits such as disease or pest resistance. Due to the partitioning of diversity within sections and more narrowly into accessions, it is recommended that increasing the numbers of individuals per accession and the number of accessions per section is more beneficial than broadly sampling across several different sections. This is also applicable to germplasm acquisition efforts. More genetic diversity may be obtained by collecting more individuals in a single species than collecting as many different species as possible. This information is important to both plant breeders and curators of germplasm collections due to the expense in both time and monies involved in screening plants for important traits and collecting and maintaining germplasm. By maximizing the diversity obtained when choosing a study or collection population, the efforts of the investigation are reduced and made more efficient.

REFERENCES

Abdalla A.M., Reddy O.U.K., El-Zik K.M., Pepper A.E. 2001. Genetic diversity and relationships of diploid and tetraploid cottons revealed using AFLP. Theor Appl Genet. 102:222-229.

Barrett B.A., Kidwell K.K. 1998. AFLP-based genetic diversity assessment among wheat cultivars from the Pacific Northwest. Crop Sci. 38:1261-1271.

Breyne P., Rombaut D., Van Gysel A., Van Montagu M., Gerats T. 1999. AFLP analysis of genetic diversity within and between *Arabidopsis thaliana* ecotypes. Mol Gen Genet. 261:627-634.

Cervera M.T., Cabezas J.A., Sancha J.C., Martinez de Toda F., Martinez-Zapatar J.M. 1998. Application of AFLPs to the charactarization of grapevine *Vitis vinifera* L. genetic resources. A case study with accessions from Rioja (Spain). Theor Appl Genet. 97:51-59.

Cherry J.P. 1975. Comparative studies of seed proteins and enzymes of species and collections of *Arachis* by gel electrophoresis. Peanut Science. 2:57-65.

Felsenstein, J. 1989. PHYLIP - Phylogeny Inference Package (Version 3.5). Cladistics 5: 164-166.

Galgaro L., Valls J.A., and Lopes C.R. 1997. Study of the Genetic Variability and Similarity Among and Within *Arachis villosulicarpa*, *Arachis pietrarellii* and *Arachis hypogaea* through Isoenzyme Analysis. Genetic Resources and Crop Evolution. 44(1):9-15.

Gimenes A.M., Lopes C.R. Galgaro M.L., Valls J.F.M., and Kochert G. 2000. Genetic variation and phylogenetic relationships based on RAPD analysis in section *Caulorrhizae*, genus *Arachis* (Leguimosae). Euphytica. 116:187-195.

Gregory, M.P. and W.C. Gregory. 1979. Exotic germplasm of *Arachis L*. interspecific hybrids. J. Heredity 71:185-193.

Herbergs J., Siwek M., Crooijmans R.P.N.A., Van der Poel J.J., Groenen M.A.M. 1999. Multicolour fluorescent detection and mapping of AFLP markers in chicken (*Gallus domesticus*). Animal Genetics. 30:274-285.

Hill M., Witsenboer H., Zabeau M., Vos P., Kesseli R., Michelmore R. 1996. PCR-based fingerprinting using AFLPs as a tool for studying genetic relationships in *Lactuna* spp. Theor Appl Genet. 93:1202-1210.

Karp A., Seberg O., Buiatti M. 1996. Molecular Techniques in the Assessment of Botanical Diversity. Annals of Botany. 78(2):143-149.

Klozova E.J. Svachulova J., Smartt J., Hadag E., Turkova V., Hadaco-Va V. 1983a. The comparison of seed protein patterns withing the genus *Arachis* by polyacrylamide gel electrophoresis. Biol Plant (Praha). 25:266-273.

Klozova E., Turkova V., Smartt J., Pitterova K., Svachulova J. 1983b. Immunochemical Characterization of Seed Proteins of Some Species of the Genus *Arachis* L. Biologia Plantarum. 25(3):201-208.

Krapovickas A. and Gregory W.C. 1994. Taxonomia del genero *Arachis* (*Leguminosae*). Bonplandia. 8:1-186.

Lombard V., Baril C.P., Dubreuil P., Blouet F., Zhang D. 2000. Genetic Relationships and Fingerprinting of Rapeseed Cultivars by AFLP: Consequences for Varietal Registration. Crop Sci. 40:1417-1425.

Maass B. L., Tores A.M., Ocampo C.H. 1993. Morphological and Isozyme Characterization of *Arachis pintoi* Krap e Greg Nom Num Germplasm. Euphytica. 70(1-2):43-52.

Maughan P.J., Saghai Maroof M.A., Buss G.R., Huestis G.M. 1996 Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near isogenic line analysis. Theor Appl Genet. 93: 392-401.

Milbourne D., Meyer R., Bradshaw J.E., Baird E., Bonar N., Provan J., Powell W., Waugh R. 1997. Comparison of PCR-based marker systems for the analysis of genetic relationships in cultivated potato. Mol Breed. 3:127-136.

Paul S., Wachira F.N., Powell W., Waugh R. 1997. Diversity and genetic differentiation among populations of Indian and Kenyan tea (*Camellia sinensis* (L.) O. Kuntze) revealed by AFLP markers. Theor Appl Genet. 94:255-263.

Pejic I., Ajmone-Marsan P., Morgante M., Kozumplick V., Castigliioni P., Taramino G., Motto M. 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. Theor Appl Genet. 97:1248-1255.

Powell W., Morgante M., Andre C., Hanafey M., Vogel J., Tingey S., Rafalski A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Molecular Breeding. 2:225-238.

Quagliaro G., Vischi M., Tyrka M., and Olivieri A.M. 2001. Identification of Wild and Cultivated Sunflower for Breeding Purposes by AFLP Markers. J Heredity. 92:38-42.

Schneider S., Roessli D., Excoffier L. 2000. Arlequin ver. 2.000. A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Schut J.W., Qi X., Stam P. 1997. Association between relationship measures based on AFLP-markers, pedigree data and morphological traits in barley. Theor Appl Genet. 95:1161-1168.

Singh A.K., Gurtu S., Jambunathan R. 1994. Phylogenetic Relationships in the Genus *Arachis* Based on Seed Protein Profiles. Euphytica. 74(3):219-225.

Smartt J. and Stalker T. 1982. Speciation and Cytogenetics in *Arachis*. In Pattee H.E. and Young C.T. (Eds.) “Peanut Science and Technology”. American Peanut Research and Education Society, Inc., Yoakum, Texas.

Stalker , H.T., T.D. Phyllips, J.P. Murphy, and T.M. Jones. 1994. Diversity of isozyme patterns in Arachis species. Theor. Appl. Genet. 87:746-755.

USDA, ARS, National Genetic Resources Program. Germplasm Resources Information Network - (GRIN). 2003. [Online Database] National Germplasm Resources Laboratory, Beltsville, Maryland. Available: http://www.ars-grin.gov.

Valls, J.F.M. and Simpson C.E. 1997. Novas expecies de *Arachis* (*Leguminosae*). In: Veiga R.F. de A., Bovi M.L.A., Betti J.A. and Volton R.B.X. (Eds.). “Programas e resumos del I Simposio Latinoamericano de Recursos Geneticos Vegetais.” Campinas.

Vos, P., R. Hogers, M. Beeker. M. Reijans, T. Vandelee, M. Hornes, A. Frijiters, J. Pot,

J. Peleman, M. Kuiper & M. Zabeau, 1995. AFLP - a new technique for DNA fingerprinting. Nucl Acids Res 23:4407-4414.

Weir, B.S. 1996. Genetic Data Analysis II: Methods for Discrete Population Genetic Data. Sinauer Assoc., Inc., Sunderland, MA, USA.

Weir, B.S. and Cockerham,C.C. 1984. Estimating F-statistics for the analysis of population structure. Evolution. 38:1358-1370.

Wong H., Yeoh H., Lim S. 1999. Customization of AFLP analysis for cassava varietal identification. Phytochemistry. 50:919-924.

Zhu J, Gale M.D., Quarrie S. 1998. AFLP markers for the study of rice biodiversity. Theor Appl Genet. 96:602-611.

Table 2.1. List of *Arachis* accessions. Identification numbers, Grif and PI, correspond to those in the GRIN database (http://www.ars-grin.gov/npgs/).

***Arachis* - Annuals**	
batizicoi	Grif 14249, PI 468326, PI 468329
benensis	PI 475879, PI 540870
cruziania	Grif 14257, PI 476003
dardanoi	PI 591364
decora	Grif 7722
duranensis	PI 262133, PI 468197, PI 468201, PI 219823
glandulifera	PI 468341, PI 468343
hoehni	Grif 7682, Grif 7683
ipaensis	PI 468321, PI 468322
magna	PI 468338
monticola	PI 405933
stenosperma	Grif 7482, PI 338280, PI 497578
valida	Grif 7689, PI 468154
williamsii	Grif 14229
***Arachis* - perennials**	
cardenasii	PI 262141, PI 476011
correntina	PI 262808
diogoi	PI 468344, PI 468345, PI 604829
helodes	PI 476042, PI 476044
kempff-mercadoi	PI 468333, PI 468334, PI 475875
simpsonii	Grif 13710, Grif 13728
villosa	Grif 7724, PI 330651
Caulorrhizae	
pintoi	Grif 14999, Grif 14997, PI 604797, PI 604818
repens	Grif 7520, PI 338277
Erectoides	
archeri	Grif 11998, PI 468188
benthamii	PI 468162, PI 604834
cryptopotmaica	Grif 12007, PI 468165
hermannii	PI 604846, PI 604847
major	Grif 11996, Grif 7702
oteri	Grif 11994
paraguariensis	Grif 11974, Grif 11976
paraguariensis var. *capibarensis*	PI 468176
paraguariensis var. *paraguariensis*	PI 468155, PI 468181, PI 468362
stenophylla	PI 468170

Table 2.1. cont.

Extranervosae	
burchelli	Grif 12026, Grif 7598
macedoi	PI 604851
prostrata	PI 497538, PI 338449
villosulocarpa	PI 604839, PI 604840
Heteranthae	
dardonoi	Grif 7753
pusilla	PI 476136
Procumbentes	
appressipila	PI 468149, PI 468153
chiquitana	PI 476004
kretschmeri	Grif 7709, PI 468151
matiensis	PI 476105, PI 476121
rigonii	PI 262142
subcoriacea	PI 468169, PI 604826
Rhizomatosae	
glabrata var. *glabrata*	PI 118457, PI 163452, PI 234318
glabrata var. Hagenbeckii	PI 262832, PI 262833
pseodovillosa	PI 262847, PI 468173

Table 2.2. AFLP Protocol. Nucleotide sequences for adaptors, pre-selective primers and selective primers used in the AFLP protocol using Mse1 and Pst1 restriction enzymes are listed.

Primer\Adaptor	**Sequence**
Mse1 adaptor 1	5' - GACGATGAGTCCTGAG - 3'
Mse1 adaptor 2	3' -TACTCAGGACTCAT - 5'
Pst1 adaptor 1	5' -CTCGTGACTGCGTACATGCA - 3'
Pst1 adaptor 2	3' -CATCTGACGCATGT- 5'
Mse1 Pre-selective primer	5' -GATGAGTCCTGAGTAA- 3'
Pst1 Pre-selective primer	5' -AGACTGCGTACATGCAG - 3'

	Extension	**Label**
Pst Selective Primers		
P66	GAT	FAM
P67	GCA	TET
P75	GTA	HEX
P77	GTG	TET
P78	GGT	FAM
Mse Selective Primers		
M12	AC	
M13	AG	
M14	AT	
M15	AA	

Table 2.3. Genetic Diversity Indexes. Nei's genetic diversity indexes generated using AFLP profile data according to taxonomic section of *Arachis* are listed.

Section	**Nei's Genetic Diversity**
Arachis-annuals	0.260916 +/- 0.127837
Arachis-perennials	0.220006 +/- 0.111875
Caulorrhizae	0.152872 +/- 0.085743
Erectoides	0.230296 +/- 0.115582
Extranervosae	0.227321 +/- 0.131579
Heteranthae	0.277639 +/- 0.207311
Procumbentes	0.214972 +/- 0.115496
Rhizomatosae	0.199501 +/- 0.111796

Table 2.4. AMOVA Analysis. Using the AFLP data AMOVA analysis (Weir, B.S. and Cockerham, C.C. 1984. Excoffier, L., Smouse, P., and Quattro, J. 1992. Weir, B. S., 1996) was performed using Arlequin software ver. 2.000 (Schneider et al., 2000). The groups are defined as the four major clades observed in the phylogenetic analysis of *Arachis* (Clade 1 - *Arachis* annuals and perennials; Clade 2 - *Rhizomatosae, Erectoides* and *Procumbentes*; Clade 3 - *Heteranthae* and *Extranervosae*; Clade 4 - *Caulorrhizae*). The populations are defined as the eight taxonomic sections of *Arachis* used in the study.

Source of variation	**d.f.**	**Sum of squares**	**Variance components**	**Percentage of variation**
Among groups	3	2395.045	26.46843 Va	14.92
Among populations within groups	4	1100.505	12.288778 Vb	6.93
Within populations	86	11925.726	128.67123 Vc	78.16
Total	**93**	**15421.277**	**177.42744**	

Table 2.5. Summary of research results regarding possible progenitors of the cultivated peanut, *Arachis hypogaea*.

Reference	**A Genome**	**B Genome**	**Evidence**
Seetharam et al., 1973	*A. duranensis*	-	Cytogenetics
Varisai Muhammad, 1973	*A. villosa*	-	Cytogenetics
Krapovickas, Fernandez, Seeligman, 1974	-	*A. batizocoi*	Cytogenetics
Gregory and Gregory, 1976	*A. duranensis*	*A. cardenasii*	Cytogenetics
Smartt, Gregory, and Gregory, 1978a,b	*A. cardenasii*	*A. batizocoi*	Cytogenetics
Gregory, Krapovickas, and Gregory, 1980	*A. cardenasii*	*A. batizocoi*	Cytogenetics
Singh and Moss, 1982	*A. cardenasii*	*A. batizocoi*	Cytogenetics
Smartt and Stalker, 1982	*A. cardenasii*	*A. batizocoi*	Cytogenetics
Klozova et al., 1983	-	*A. batizocoi*	Seed storage proteins
Krishna and Mitra, 1988	*A. duranensis*	*A. batizocoi*	Seed storage proteins
	A. cardenasii	-	
Singh, 1988	*A. duranensis*	*A. batizocoi*	Cytogenetics, morphology
Kochert et al., 1991	*A. duranensis*	*A. ipaensis*	RFLP
Singh et al., 1991	*A. duranensis*	*A. batizocoi*	Seed storage proteins
Lu and Pickersgill, 1993	*A. spegazzinii*	*A. glandulifera*	Isozymes
Paik-Ro, Smith and Knauft, 1992	*A. duranensis*	Not *A. batizocoi*	RFLP
Bianchi-Hall et al., 1993	-	Not *A. batizocoi*	Seed storage proteins
Lanham et al., 1994	*A. spegazzinii*	*A. batizocoi*	Seed storage proteins
Stalker et al., 1994	-	Not *A. batizocoi*	Isozymes
Hilu and Stalker, 1995	-	Not *A. batizocoi*	RAPD
Kochert et al., 1996	*A. duranensis*	*A. ipaensis*	RFLP, cytogenetics
Raina and Mukai, 1999	*A. villosa*	*A. ipaensis*	in situ hybridization
Raina et al., 2001	*A. villosa*	*A. ipaensis*	RAPD, ISSR

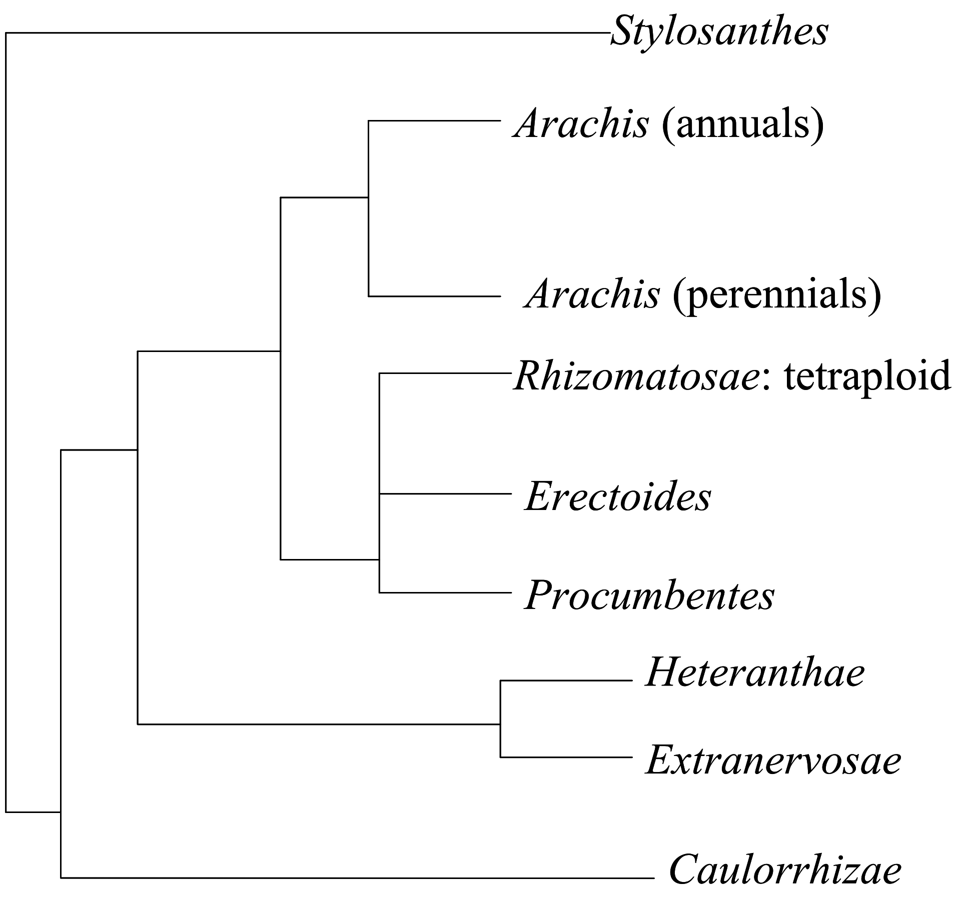

Figure 2.1. Schematic representation of dendogram obtained from AFLP profiles. Clades have been collapsed to show the clustering of accessions according to section and series within section (*Arachis* - annuals and perennials).

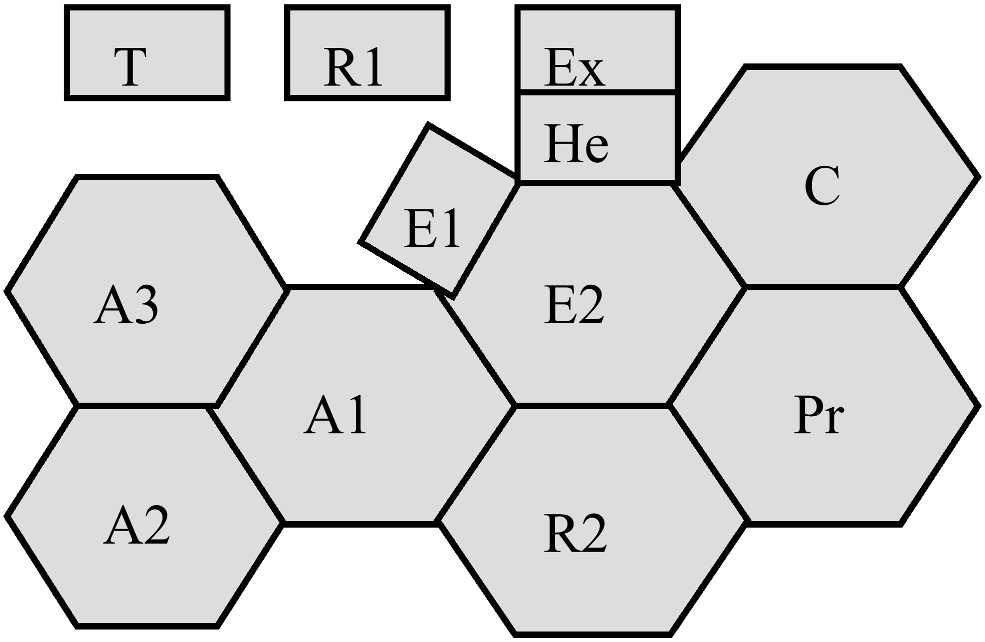

Figure 2.2. Schematic summary of interspecific hybridization studies. (Gregory and Gregory, 1979)

The shapes represent the taxonomic sections of the genus *Arachis* as follows: A1, *Arachis* annuals; A2, *Arachis* perennials; A3, *Arachis hypogaea*; C, *Caulorrhizae*; E1, *Trierectoides*; E2, *Erectoides*; EX, *Extranervosae*; He, *Heteranthae*; Pr, *Procumbentes*; R1, *Rhizomatosae*:series *prorhizomatosae*; R2, *Rhizomatosae* - tetraploid; T, *Triseminatae*. Where a face of each shape touches another indicates that species from this section will produce hybrids from species in the other section.

CHAPTER 4

CONCLUSIONS

A. hypogaea is an important oilseed crop grown worldwide. There is constant work being performed to improve the crop whether by increasing the nutrient content, looking for sources of resistance to disease and pests, increasing drought tolerance, etc. One of the means of crop improvement is through introduction of traits from the wild species in *Arachis.* Wild *Arachis* species can be found throughout South America, but an important source of wild material in the United States is the USDA Plant Germplasm System. Approximately 8000 accessions of *A. hypogaea* and over 500 accessions of wild *Arachis* species are available for distribution from the United States Germplasm System located in Griffin, Georgia (USDA, 2003). Preservation of this valuable resource is crucial to ensure the continued improvement of the peanut.

Two main tasks involved in ensuring a high quality germplasm collection include acquiring new plant material and maintaining current material. As wild species are being lost due to destruction of native habitat, it is necessary to make collections of new material before it is lost forever. This must be balanced with the expense in maintaining and characterizing newly collected material. Understanding the amount of genetic diversity in the genus *Arachis* and how it is partitioned is key in accomplishing these two tasks and is the main focus of this dissertation.

To understand the genetic diversity and its partitioning in the species, *A. hypogaea*, SSR technology was utilized. Phylogenetic analysis demonstrated that the accessions grouped more naturally into clades based on geographic location than according to taxonomic subspecies and varietal classification. Additionally, AMOVA analysis constructed according to taxonomy showed that the majority of the genetic diversity is partitioned among accessions (66.11%) rather than within accessions (28.02%) or among subspecies (5.88%). Also, AMOVA analysis according to geographic location showed that the majority of the genetic diversity is found among geographic locations (61.98%) rather than among species (13.37%) or within species (24.65%). Taken together, these results would indicate that it may be better to sample across diverse geographic locations when collecting new plant material rather than basing selection on taxonomy alone.

To understand the genetic diversity and its partitioning in the wild species of *Arachis*, AFLP technology was utilized. According to the data generated, the amount of genetic diversity within the genus is in accordance of what one would expect from a self-pollinating crop (Hamrick and Godt, 1987). The values for Nei's genetic diversity ranged from 0.278 for section *Heteranthae* to 0.153 for section *Caulorrhizae* with no significant differences in the values for the different sections. Secondly, the genetic diversity is partitioned such that the majority of the diversity can be found within sections (78.16%). Further evaluation revealed that the majority of the genetic diversity within a species is within accessions (93.45%) rather than among accessions (6.57%). Lastly, the dendogram revealed that the relationships that exist among the sections using the AFLP data agree with previous interspecific hybridization studies (Gregory and Gregory, 1979). Taken collectively, these data can offer guidance to crop curation by directing collection efforts as well as regeneration efforts. In order to collect and preserve more genetic diversity with a reduced collection or regeneration size, a curator may benefit from focusing on more individuals per species within a single section rather than broadly sampling across different taxonomic sections. Additionally, the data offered insight into the origin of the cultivated peanut. It was found that *A. monticola,* PI 468199*,* shared the most alleles (64%) with *A. hypogaea*. Also, both accessions of *A. monticola* and all three accessions of *A. batizicoi* used in the study had greater than 50% shared alleles with *A. hypogaea*. This gives credence to the idea that *A. monticola* is either a tetraploid precursor or simply a subspecies of *A. hypogaea* and that *A. batizicoi* may be the B genome donor of *A. hypogaea*. The species, *A. diogoi* and *A. cruziania*, also had two out of three accessions showing 50% or more shared alleles with *A hypogaea* warranting their inclusion in the list of possible progenitors.

Both of these studies offer direction for germplasm curation by suggesting more efficient ways for the collection of new material and the regeneration of current material. It can also be used by plant breeders in their efforts to identify more variability in traits of interest. By sampling within accessions and within species of wild material and across landraces of *A. hypogaea* from diverse geographic origins, a breeder may be more likely to find variability due to the increase in genetic diversity held within such a defined study population. The studies also highlight the need to develop more SSR markers for genetic studies of *A. hypogaea*. The seven markers used in the study are insufficient for genotyping purposes and identifying duplicate accessions in the germplasm system. Additionally, it points out the need to include more individuals per accession when studying wild species due to the partitioning of genetic diversity within species.

REFERENCES

Abdalla A.M., Reddy O.U.K., El-Zik K.M., Pepper A.E. 2001. Genetic diversity and relationships of diploid and tetraploid cottons revealed using AFLP. *Theor Appl Genet*. 102:222-229.

Barrett B.A., Kidwell K.K. 1998. AFLP-based genetic diversity assessment among wheat c ultivars from the Pacific Northwest. *Crop Sci*. 38:1261-1271.

Bianchi-Hall C.M., Keys R.D., Stalker H.T. and Murphy J.P. 1993. Diversity of seed storage protein patterns in wild peanut (*Arachis*, Fabaceae) species. *Plant Systematics and Evolution* 186:1-15.

Black, M.C. and Smith, D.H. 1987. Spotted wilt and rust reactions in south Texas among selected peanut genotypes. *Proc. Amer. Peanut Res. Educ. Soc.* 19:31 (abstr).

Breyne P., Rombaut D., Van Gysel A., Van Montagu M., Gerats T. 1999. AFLP analysis of genetic diversity within and between *Arabidopsis thaliana* ecotypes. *Mol Gen Genet.* 261:627-634.

Brown S., Todd J., Culbreath A., Pappu H., Baldwin J., and Beasley J. 1999. Tomato spotted wilt of peanut: Identifying and avoiding high-risk situations. *Univ. of Georgia Ext. Bull.* 1165.

Cervera M.T., Cabezas J.A., Sancha J.C., Martinez de Toda F., Martinez-Zapatar J.M. 1998. Application of AFLPs to the charactarization of grapevine *Vitis vinifera* L. genetic resources. A case study with accessions from Rioja (Spain). *Theor Appl Genet*. 97:51-59.

Cherry J.P. 1975. Comparative studies of seed proteins and enzymes of species and collections of *Arachis* by gel electrophoresis. *Peanut Science*. 2:57-65.

Culbreath A.K., J.W. Todd, J.W. Demski, and J.R. Chamberlin. 1992. Disease progress of tomato spotted wilt in peanut cultivars Florunner and Southern Runner. *Phytopathology* 82:766-771.

Culbreath A.K., J.W. Todd, D.W. Gorbet, F.M. Shokes, and H.R. Pappu. 1997. Field response of new peanut cultivar UF 91108 to tomato spotted wilt virus. *Plant Des.* 81:1410-1415.

Culbreath A.K., J.W. Todd, D.W. Gorbet, S.L. Brown, J.A. Baldwin, H.R. Pappu, C.C. Holbrook, and F.M. Shokes. 1999. Response of Early, Medium, and Late Maturing

Peanut Breeding Lines to Field Epidemics of Tomato Spotted Wilt. *Peanut Science* 26:100-106.

Galgaro L., Lopes C.R., Gimenes M., Valls J.F.M., Kochert G. 1998. Genetic variation between several species of sections Extranervosae, Caulorrhizae, Heteranthae, and Triseminatae (genus *Arachis*) estimated by DNA polymorphism. *Genome.* 41:445-454.

Galgaro L., Valls J.A., and Lopes C.R. 1997. Study of the Genetic Variability and Similarity Among and Within *Arachis villosulicarpa*, *Arachis pietrarellii* and *Arachis hypogaea* through Isoenzyme Analysis. *Genetic Resources and Crop Evolution*. 44(1):9-15.

Garcio G.M., Stalker H.T. and Kochert G. 1995. Introgression analysis of an interspecific hybrid population in peanuts (*Arachis hypogaea* L.) using RFLP and RAPD markers. *Genome.* 38:166-176.

Gimenes A.M., Lopes C.R. Galgaro M.L., Valls J.F.M., and Kochert G. 2000. Genetic variation and phylogenetic relationships based on RAPD analysis in section *Caulorrhizae*, genus *Arachis* (Leguimosae). *Euphytica.* 116:187-195.

Gregory W.C., Gregory M.P. 1976. Groundnuts: *Arachis hypogaea - Leguminosae - Papilionatae*. In: Simmonds N.W. (Ed.): Evolution of Crop Plants. pp.151-154. Longman Group Ltd. London, New York.

Gregory, M.P. and W.C. Gregory. 1979. Exotic germplasm of *Arachis* L. interspecific hybrids. J. Hered. 70:185-193.

Gregory W. C., Krapovickas A., and Gregory M.P. 1980. Structure, variation, evolution, and classification in *Arachis. In* R.J. Summerfield and A.H.Bunting [eds.], *Advances in legume science*. pp. 469-481. Royal Botanic Gardens, Kew.

Grosso N.R., Krapovickas A., Pietrarelli J.R. and Guzman C.A. 2000. Chemical composition of some wild peanut species (Arachis L.) seeds.. J Agric Food Chem. 48(3):806-809.

Halward T.M., Stalker H.T., Kochert G. Development of an RFLP linkage map in diploid peanut species. *Theor. Appl. Genet*. 87:379-384.

Halward T.M., Stalker H.T., LaRue E.A., Kochert G. 1991. Genetic variation detectabale with molecular markers among adapted germ-plasm resources of cultivated peanut and related wild species. *Genome.* 34:1013-1020.

Halward T.M., Stalker H.T., LaRue E.A., Kochert G. 1992. Use of single-primer DNA amplifications in genetic studies of peanut (*Arachis hypogaea* L.). *Plant Mol. Biol.* 18:315-325.

Halward T.M., Stalker H.T., LaRue E.A., Kochert G. 1991. Genetic variation detectabale with molecular markers among adapted germ-plasm resources of cultivated peanut and related wild species. *Genome.* 34:1013-1020.

Halward T.M., Stalker H.T., LaRue E.A., Kochert G. 1992. Use of single-primer DNA amplifications in genetic studies of peanut (*Arachis hypogaea* L.). *Plant Mol. Biol*. 18:315-325.

He G. and Prakash C.S. 1997. Identification of polymorphic DNA markers in cultivated peanut (*Arachis hypogaea* L.). *Euphytica.* 97:143-149.

Herbergs J., Siwek M., Crooijmans R.P.N.A., Van der Poel J.J., Groenen M.A.M. 1999. Multicolour fluorescent detection and mapping of AFLP markers in chicken (*Gallus domesticus*). *Animal Genetics*. 30:274-285.

Hill M., Witsenboer H., Zabeau M., Vos P., Kesseli R., Michelmore R. 1996. PCR-based fingerprinting using AFLPs as a tool for studying genetic relationships in *Lactuna* spp. *Theor Appl Genet.* 93:1202-1210.

Hilu K.W. and Stalker H.T 1995. Genetic relationships between peanut an wild species of *Arachis* sect. *Arachis* (Fabaceae): evidence from RAPDs. Pl. Syst. Evol. 198:167-178.

Hopkins M.S., Casa A.M., Wang T., Mitchell S.E., Dean R.E., Kochert G.D., Kresovich S. 1999. Discovery and characterization of polymorphic simple sequence repeats (SSRs) in peanut. *Crop Science.* 39(4):1243-1247.

Klozova E.J. Svachulova J., Smartt J., Hadag E., Turkova V., Hadaco-Va V. 1983a. The comparison of seed protein patterns withing the genus *Arachis* by polyacrylamide gel electrophoresis. *Biol Plant* (Praha). 25:266-273.

Klozova E., Turkova V., Smartt J., Pitterova K., Svachulova J. 1983b. Immunochemical Characterization of Seed Proteins of Some Species of the Genus *Arachis* L. *Biologia Plantarum*. 25(3):201-208.

Kochert G., Halward T., Branch W.D., Simpson C.E. 1991. RFLP variability in peanut *(Arachis hypogaea* L.) cultivars and wild species. *Theor. Appl. Genet*. 5 65-570.

Kochert G., Stalker T.H., Gimenes M., Galgaro L., Lopes C.R., Moore K. 1996. RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). *Am. J. Botany.* 83(10):1282-1291.

Krapovickas A., Fernandez A., and Seeligman P. 1974. Recuperacion de la fertilidaden un hibrido interspecifico de *Arachis* (Leguminosae). *Bonplandia* 3:129-142.

Krapovickas A. and Gregory W.C. 1994. Taxonomica del genero *Arachis* (*Leguminosae*). *Bonplandia*. 8:1-186.

Krishna T.G. and Mitra R. 1988. The probable genome donors to *Arachis hypogaea* L. based on arachin seed storage protein. *Euphytica*. 37:47-52.

Lanham P.G. and Fennell S. 1992. Detection of polymorphic loci in *Arachis* germplasm using random amplified polymorphic DNAs. *Genome.* 35:885-889.

Lanham P.G., Forster B.P., McNicol P., Moss J.P. and Powell W. 1994. Seed storage protein variation in *Arachis* species. *Genome.* 37. 487-496.

Lavia G.I. 2000. Chromosome studies in wild *Arachis* (Leguminosae). *Caryologia.* 53:277-281.

Lombard V., Baril C.P., Dubreuil P., Blouet F., Zhang D. 2000. Genetic Relationships and Fingerprinting of Rapeseed Cultivars by AFLP: Consequences for Varietal Registration. *Crop Sci.* 40:1417-1425.

Lu J., and Pickersgill B. 1993. Isozyme variation and species relationships in peanut and its wild relatives (*Arachis* L.: *Leguminosae*). Theor. Appl. Genet. 85:550-560.

Ma ZQ, Roder M, Sorells ME. 1996. Frequencies and sequence characteristics of di-, tri-, and tetra-nucleotide microsatellites in wheat. *Genome*. 39:123-130.

Maass B.L. and Ocampo C.H. 1995. Isozyme Polymorphism Provides Fingerprints for Germplasm of *Arachis glabrata* Bentham. *Genetic Resources and Crop Evolution.* 42(1):77-82.

Maass B. L., Tores A.M., Ocampo C.H. 1993. Morphological and Isozyme Characterization of *Arachis pintoi* Krap e Greg Nom Num Germplasm. *Euphytica.* 70(1-2):43-52.

Maughan P.J., Saghai Maroof M.A., Buss G.R., Huestis G.M. 1996 Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near isogenic line analysis. *Theor Appl Genet*. 93: 392-401.

Milbourne D., Meyer R., Bradshaw J.E., Baird E., Bonar N., Provan J., Powell W., Waugh R. 1997. Comparison of PCR-based marker systems for the analysis of genetic relationships in cultivated potato. *Mol Breed.* 3:127-136.

Mitchell S.E., Kresovich S., Jester C.A., Hernandez C.J., Szewc-McFadden. 1997. Application of multiple PCR and fluorescence-based, semi-automated, allele sizing technology for genotyping plant genetic resources. *Crop Sci*. 37:617-624.

Paik-Ro O.G., Smith R.L., Knauft D.A. 1992. Restriction fragment length polymorphism evaluation of six peanut species within the *Arachis* section. *Theor. Appl. Genet*. 84:201-208.

Paul S., Wachira F.N. Powell W., Waugh R. 1997. Diversity and genetic differentiation among populations of Indian and Kenyan tea (*Camellia sinensis* (L.) O. Kuntze) revealed by AFLP markers. *Theor Appl Genet*. 94:255-263.

Pejic I., Ajmone-Marsan P., Morgante M., Kozumplick V., Castigliioni P., Taramino G., Motto M. 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. *Theor Appl Genet*. 97:1248-1255.

Powell W., Morgante M., Andre C., Hanafey M., Vogel J., Tingey S., Rafalski A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding*. 2:225-238.

Quagliaro G., Vischi M., Tyrka M., and Olivieri A.M. 2001. Identification of Wild and Cultivated Sunflower for Breeding Purposes by AFLP Markers. *J Heredity.* 92:38-42.

Raina S.N. and Mukai Y. 1999. Genomic in situ hybridization in *Arachis* (Fabaceae) identifies the diploid wild progenitors of cultivated (A. hypogaea) and related wild (A. monticola) peanut species. *Pl. Syst. Evol.* 214:251-262.

Raina S.N., Rani V., Kojima T., Ogihara Y., Singh K.P., and Devarumath R.M. 2001. RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity, varietal identification, and phylogenetic relationships in peanut (*Arachis hypogaea*) cultivars and wild species. *Genome.* 44:763-772.

Schut J.W., Qi X., Stam P. 1997. Association between relationship measures based on AFLP-markers, pedigree data and morphological traits in barley. *Theor Appl Genet.* 95:1161-1168.

Seetharam A.K., Nayar M.D., Sreekantaradhya R., and Achar D.K.T. 1973. Cytological studies on the interspecific hybrid of *Arachis hypogaea* X *Arachis duranensis*. *Cytologia.* 38:277-280.

Sharma S.B., Ansari M.A., Varaprasad K.S., Singh A.K. and Reddy L.J. 1999. Resistance to *Meloidogyne javaica* in wild *Arachis* species. *Genetic Resources and Crop Evolution.* 46:557-568.

Shelton, K.L., A.K. Culbreath, H.R. Pappu, and J.W. Todd. 1998. Consideration of spatial patterns in field characterization of peanut genotype response to TSWV. *Phytopathology*. 88:S127 (abstr.).

Singh A.K. 1988. Putative genome donors of *Arachis hypogaea* (Fabaceae) evidence from crosses with synthetic amphidiploids. *Plant Systematics and Evolution.* 160:143-151.

Singh A.K., Gurtu S., Jambunathan R. 1994. Phylogenetic Relationships in the Genus *Arachis* Based on Seed Protein Profiles. *Euphytica.* 74(3):219-225.

Singh A.K., Sivaramakrishnan S., Mengesha M.H. and Ramaiah C.D. 1991. Phylogenetic relations in section *Arachis* based on seed protein profile. *Theor Appl Gen.* 82:593-597.

Singh A.K., and Moss J.P. 1982. Utilization of wild relatives in genetic improvement of *Arachis hypogaea* L. Part 2: Chromosome complements of species in section *Arachis*. *Theor Appl Gen*. 61:305-314.

Smartt J., Gregory W.C., Gregory M.P. 1978a. The genomes of *Arachis* hypogaea 1. Cytogenetic studies of putative genome donors. *Euphytica.* 27:665-675.

Smartt J., Gregory W.C., Gregory M.P. 1978b. The genomes of *Arachis* hypogaea 2.The implications in interspecific breeding. *Euphytica.* 27:677-680.

Smartt J., and Stalker H.T. 1982. Speciation and cytogenetics in *Arachis*. In H.E. Pattee and C.T. Young [eds.], *Peanut science and technology*, pp. 21-49, American Peanut Research and Education Society, Yoakum, TX.

Stalker H.T. 1981. Hybrids in the genus *Arachis* between sections *Erectoides* and *Arachis*. *Crop Science*. 21:359-362.

Stalker HT. 1985. Cytotaxonomy of *Arachis*. In: Moss JP (ed) *Proc Int Workshop on Cytogenetics of Arachis*. International Crops Research Institute for the Semi-Arid Tropics, Patancheru, A.P., India, pp 64-79.

Stalker H.T., Dhesi J.S., Kochert G. Genetic diversity within the species *Arachis* duranensis Krapov. & W.C. Gregory, a possible progenitor of cultivated peanut. *Genome.* 38:1201-1212.

Stalker HT, Phillips T.D., Murphy J.P. and Jones T.M. 1994. Variation of isozyme patterns among *Arachis* species. *Theor Appl Genet.* 87:746-755.

Stalker HT and Simpson CE. 1995. Germplasm Resources in *Arachis*. Chapter in *Advances in Peanut Science*. Edited Pattee HE and Stalker HT. American Peanut and Research and Education Society, Inc. Stillwater, OK, USA.

Stalker HT, Young CT, and Jones TM. 1989. A survey of the fatty acids of peanut species. *Oleagineux.* 44:419-424.

Subramanian V., Gurtu S., Nageswara Rao R.C., and Nigam S.N. 2000. Identification of DNA polymorphism in cultivated groundnut using random amplified polymorphic DNA (RAPD) assay. *Genome.* 43:656-660.

Todd, J.W. Culbreath A.K., Chamberlin J.R., Beshear R.J., and Mullinix B.G. 1995. Colonization and population dynamics of thrips in peanuts in the southern United States, pp. 453-460. In B.L. Parker, M. Skinner, an T. Lewis (eds) *Thrips Biology and Management.* Plenum Press, New York.

USDA, ARS, National Genetic Resources Program. Germplasm Resources Information Network - (GRIN). 2003. [Online Database] National Germplasm Resources Laboratory, Beltsville, Maryland. Available: http://www.ars-grin.gov.

Valls, J.F.M. 1996. Variability in the genus *Arachis* and potential forage uses. In: T.L. Springer & R.N. Pittman (EDS.), *Identifying Germplasm for successful Forage Legume-Grass Interactions*, 1994, Seattle, Proceedings of a Symposium of the Crop Science Society of America, Washington: USDA-ARS. pp. 15-27.

Valls, J.F.M. and Simpson C.E. 1997. Novas expecies de *Arachis* (*Leguminosae*). In: Veiga R.F. de A., Bovi M.L.A., Betti J.A. and Volton R.B.X. (EDS.). *"Programas e resumos del I Simposio Latinoamericano de Recursos Geneticos Vegetais."* Campinas.

Varisai M.S. 1973. Cytogenetical investigations in the genus *Arachis* L. II. Triploid hybrids and their derivatives. *Madras Agricultural Journal*. 60:1414-1427.

Vos P., Hogers R., Bleeker M., et al. 1995. AFLPs: a new technique for DNA fingerprinting. *Nucleic Acids Research.* 23:4407-4414.

Wong H., Yeoh H., Lim S. 1999. Customization of AFLP analysis for cassava varietal identification. *Phytochemistry.* 50:919-924.

Zhu J, Gale M.D., Quarrie S. 1998. AFLP markers for the study of rice biodiversity. *Theor Appl Genet*. 96:602-611.

APPENDIX

Morphological features of the different sections of *Arachis* (Stalker and Simpson, 1995).

Section *Arachis*
- plant tap-rooted with vertical pegs
- flowers without red veins on back of standard

  Series *Annuae*
  - flowers medium to small
  - standard 14mm wide X 12 mm high
  - short-lived, usually annual
  - 2n = 2x = 20

  Series *Perennes*
  - Flowers medium to large
  - standard 14 mm wide x 12 mm high
  - perennial
  - 2n = 2x = 20

  Series *Amphiploides*
  - Flowers small to large
  - standard 10-21 mm wide x 8-14 mm high
  - short-lived
  - 2n = 4x = 40

Section *Erectoides*
- plants tap-rooted or with tuberiform hypocotyl
- plants erect or prostrate
- pegs horizontal or nearly so
- flowers medium to large 16-24 mm x 12-20 mm
- 2n = 2x = 20

  Series *Trifoliolatae*
  - hypocotyl tuberiform
  - leaves trifoliolate

  Series *Tetrafoliatae*
  - plants erect or prostrate
  - hypocotyls not tuberiform
  - leaves tetrafoliolate
  - standard orange

Section *Procumbentes*
- plants prostrate
- standard yellow
- 2n = 2x = 20

Section *Caulorhizae*
- plants with hollow stems
  rooting at nodes
- pegs vertical
- standard yellow
- 2n = 2x = 20

.

Section *Rhizomatosae*
- plants rhizomatous
- solid stems
- flowers large

Series *Prorhizomatosae*
- plants delicate
- flowers large
- red veins on both faces of standard
- $2n = 2x = 20$

Series *Eurhizomatosae*
- plants usually robust
- flowers large
- without red veins on back of standard
- $2n = 4x = 40$

Section *Extranervosae*
- plants with thickened lomentiform tuberoid roots
- pegs vertical, sometimes producing adventious roots
- flowers small to medium with red veins on back
- $2n = 2x = 20$

Section *Heteranthae*
- plants tap-rooted
- pegs vertical
- flowers very small
- standard with red veins on front and back

Section *Triseminatae*
- plants tap-rooted
- pegs horizontal
- flowers small
- purple mark inside orange standard
- fruits often three-segmented
- $2n = 2x = 20$