Theor Appl Genet (1993) 85:550–560

© Springer-Verlag 1993

Isozyme variation and species relationships in peanut and its wild relatives (*Arachis* L. – Leguminosae)

J. Lu* and B. Pickersgill

Department of Agricultural Botany, School of Plant Sciences, University of Reading, Reading, RG6 2AS, UK

Received March 5, 1992; Accepted October 7, 1992
Communicated by A.L. Kahler

Summary. *Arachis hypogaea* (peanut or groundnut) is an AABB allotetraploid whose precise ancestry is not yet clear. Its closest diploid relatives are the annual and perennial wild species included with it in the section *Arachis*. Variation in these species for 11 different enzymes was studied by starch-gel electrophoresis. Differences attributed to at least 13 genetic loci were found among eight enzymes, while three enzymes appeared uniform throughout the section. Values for Nei's genetic distance were calculated for all pairs of species and were used to estimate relationships. All diploid species, apart from two whose validity had previously been questioned, could be distinguished by their overall zymotypes, but few contained unique alleles. When species were grouped by their mean genetic distances, they formed two clusters, which agreed reasonably well with the division of the section into annual versus perennial species. The single B-genome species was an outlier within the annual group. *A. hypogaea* showed fixed heterozygosity at four loci (in ssp. *hypogaea*) or six loci (in ssp. *fastigiata*), which agrees with previous conclusions that the peanut is an allotetraploid. None of the diploids included in this survey could be conclusively identified as donors of either the A or the B genome to the tetraploids. The two subspecies of *A. hypogaea* differed consistently in two of the thirteen putative loci studied. This may call into question the simple hypothesis that *A. hypogaea* originated from just two diploid species.

Key words: *Arachis* – Allotetraploid – Isozyme – Genetic distance – Fixed heterozygosity

* *Present address:* Center for Viticulture, Florida A and M University, Tallahassee, FL 32307, USA
Correspondence to: B. Pickersgill

Introduction

The genus *Arachis* comprises seven sections and many species, most of which have not yet been formally described (Gregory et al. 1973; Gregory and Gregory 1979). *Arachis hypogaea*, the peanut or groundnut, belongs to the section *Arachis* and is the principal cultivated species in the genus. It is a tetraploid which can be crossed with other species in the section *Arachis* but not with species in other sections (Gregory and Gregory 1979). This implies that its diploid progenitors are members of the section *Arachis* (Gregory and Gregory 1976).

There are currently thought to be more than 20 species in the section *Arachis* (Stalker 1990). All are diploid except *A. hypogaea* and *A. monticola*, which is the wild counterpart of *A. hypogaea* (Krapovickas 1969, 1973). Hybrids between most of these diploids have reduced pollen fertility and produce few seeds. Hybrids involving *A. batizocoi* and *A. glandulifera* are completely sterile and have few chiasmata at meiosis (Smartt et al. 1978; Singh and Moss 1984; Stalker 1991). Stalker (1991) therefore concluded that there are three different genomes in the section *Arachis*: the A genome, present in the majority of species; the B genome, known only in *A. batizocoi*; and the D genome, known only in *A. glandulifera*. The B- and D-genome species are both annuals, like the tetraploids. The A-genome species include both annuals and perennials. Both *A. monticola* and *A. hypogaea* are believed to be AABB allo- or segmental allo-tetraploids (Smartt et al. 1978; Singh and Moss 1982, 1984; Singh 1988; Stalker 1991). *A. batizocoi*, as the only B-genome species so far known, is generally considered to be the source of the B genome. Species suggested as the A-genome donor include the perennials *A. villosa* (Varisai Muhammad 1973) and *A. cardenasii* (Smartt et al. 1978; Krishna and Mitra 1988) and the annuals *A. duranensis* (Singh 1988; Krish-

na and Mitra 1988; Kochert et al. 1991, Singh et al. 1991; *A. spegazzinii* and *A. ipaensis* (Kochert et al. 1991).

A. hypogaea contains two subspecies which differ markedly in their branching patterns and associated physiological and agronomic traits (Bunting 1955; Bunting and Elston 1980) and which also differ slightly in karyotype (Stalker and Dalmacio 1986). Hybrids between the two subspecies are not fully fertile (Gregory et al. 1980) and the genomes of the two differ slightly in their pairing affinities with genomes of the diploid species (Singh and Moss 1984). These data suggested to Singh and Moss (1984) and Singh (1988) that the two subspecies could have originated independently, from crosses between *A. batizocoi* and two different A-genome diploids. However, recent studies of variation in genomic DNA based on restriction fragment length polymorphisms (RFLPs) and random amplified polymorphic DNAs (RAPDs) show very low variability in the tetraploids and have been interpreted as indicating a single recent origin of both *A. hypogaea* and *A. monticola* (Halward et al 1991).

Isozyme studies have proved useful in investigations of species relationships and in resolving the origin and nature of polyploids in many genera. Cherry (1975) conducted the first studies in *Arachis* species, while Grieshammer and Wynne (1990a, b) surveyed isozyme variation within *A. hypogaea*. We report here our preliminary data on isozyme variation in the section *Arachis*, and on the implications of these data for understanding the origin of the tetraploids and the relationships among the wild diploids. The latter are of interest as genetic resources for improvement of the tetraploids.

Materials and methods

Plant materials

Seeds of 37 accessions representing 12 different species in the section *Arachis* (Table 1) were obtained from *Arachis* germplasm collections at North Carolina State University, USA, and ICRISAT, India. No details were available on the numbers of wild plants originally sampled or the number and size of the generations grown subsequently *ex situ*. However, it is unlikely that the available accessions reflect fully either the genetic diversity or the genetic structure of the wild populations. Collecting *Arachis* species from the wild is extremely difficult (Valls, personal communication) and more satisfactory material is not likely to become available in the foreseeable future.

F_1 hybrids from seven combinations of these accessions (Table 1) were also included in the study. Only one of these F_1s (*A. spegazzinii* GKP 10038 × *A. duranensis* GKP 30065) was sufficiently fertile to produce an F_2 population of 26 plants.

Isozyme analysis

Five plants were studied for each accession. Extracts prepared from 20-mg samples of leaflets which were not yet fully expanded were used for all enzymes studied. Pollen extracts from 10–15 flowers per sample were prepared by grinding them in extraction medium to release both cytosolic and organellar enzymes. These extracts were used for studies of malate dehydrogenases, malic enzymes, menadione reductases and phosphoglucoisomerases. Leachates, prepared by soaking pollen from 10–15 flowers in extraction medium for 24h at 0–4°C to release cytosolic enzymes only, were used in addition to ground pollen for malic enzyme.

The extraction medium was that of Roose and Gottlieb (1978). PVP-40, PVPP, and mercaptoethanol were omitted from the medium used for pollen leachate. Mercaptoethanol was also omitted for extracts of ground pollen and for leaf extracts in which activities of aminopeptidases, malic enzyme and peroxidases were to be assessed. Proteins were separated by standard horizontal starch-gel electrophoresis in three gel/electrode buffer systems and gel slices were stained for 11 different enzymes using techniques based on those of Conkle et al. (1982), Soltis et al. (1983), Vallejos (1983) and González de León (1986). A histidine/tris-citrate buffer at pH 7.0 was used for alcohol dehydrogenases (ADH; E.C. 1.1.1.1.), isocitric dehydrogenases (IDH; E.C. 1.1.1.42), malate dehydrogenases (MDH; E.C. 1.1.1.37), menadione reductases (MNR; E.C. 1.6.99.2) and 6-phosphogluconate dehydrogenases (6-PGDH; E.C. 1.1.1.44). Tris-citrate/borate buffer at pH 7.8 was used for malic enzyme (ME; E.C. 1.1.1.39). Tris-EDTA/borate buffer at pH 8.0 (gel) or pH 8.5 (electrode) was used for aminopeptidases (AP; E.C. 3.1.3.2), phosphoglucoisomerases (PGI; E.C. 5.3.1.9), phosphoglucomutases (PGM; E.C. 2.7.5.1), peroxidases (PRX; E.C. 1.11.1.7), and shikimate dehydrogenases (SKDH; E.C. 1.1.1.25).

Bands of enzyme activity on the stained gels were assigned to regions which were numbered sequentially from the anodal end. The numbers of genetic loci and alleles controlling this activity were inferred from the banding patterns observed and from data on quaternary structure and genetic control of these enzymes in other genera (Weeden and Wendel 1989), together with F_1 and F_2 data where available. Loci were numbered to correspond with the region of the gel in which their products were active, thus *Pgi1* controls those PGI bands present in region 1. Putative alleles of the various loci were distinguished by lower-case letters assigned sequentially so that *a* designates the allele coding for the most anodal band in a particular region.

Estimation of relationships between species

Enzymes which did not vary in any of the species studied were assumed to be controlled by a single locus, even when more than one invariant band was present. Our figures may therefore overestimate distances between species. "Allele" frequencies were calculated for the genetically variable loci inferred from the observed banding patterns by pooling data from all accessions of each species. An enzyme which is controlled by a single locus in the diploids should be controlled by duplicate loci in the tetraploids unless one locus has been silenced. Tetraploid genotypes *aaaa* or *aabb* were attributed to two different homozygous loci (with genotypes *aa* and *aa* or *aa* and *bb*) for tetraploid-tetraploid comparisons, which are therefore based on twice as many loci as diploid-diploid comparisons. For diploid-tetraploid comparisons, we pooled allele frequencies at both of the duplicate loci in the tetraploid and treated these as a single locus, comparable to the equivalent locus in the diploids. Thus in a tetraploid monomorphic for genotype *aaaa*, the frequency of allele *a* was 1.0, while in a tetraploid monomorphic for genotype *aabb*, the frequency of allele *a* was considered to be 0.5. Diploid-tetraploid comparisons are therefore based on the same number of "loci" as diploid-diploid comparisons. No genotypes of forms other than homozygotes (e.g., *aaaa*) or fixed heterozygotes (e.g., *aabb*) were identified in the tetraploids.

Table 1. Species accessions and interspecific hybrids used in this study

Species/variety/hybrid	Collector[a] and number	PI no./cultivar	Origin
Diploid annuals			
A. duranensis Krap. et Greg. *nom. nud.*	GKP 30065	468201	Argentina, Salta
	GKP 30072	475847	Bolivia, Tarija
	GKP 30075	468321	Bolivia, Tarija
	K 7988	219833	Argentina, Salta
A. spegazzinii Greg. et Greg. *nom. nud.*	GKP 10038	262133	Argentina, Salta
A. stenosperma Greg. et Greg. *nom. nud.*	HLK 408	338279	Bolivia, Parana
	HLK 409	337308	Bolivia, Parana
	HLK 410	338280	Bolivia, Parana
A. batizocoi Krap. et Greg.	K 9484	298639	Bolivia, Cordillera
	GKBSPSc 30079	468325	Bolivia, Santa Cruz
	GKBSPSc 30081	468328	Bolivia, Santa Cruz
	GKBSPSc 30082	468326	Bolivia, Santa Cruz
	GKBSPSc 30080	468340	Bolivia, Santa Cruz
A. glandulifera Stalker	GKSSc 30091	468336	Bolivia, Santa Cruz
	GKSSc 30099	468342	Bolivia, Santa Cruz
	GKSSc 30100	468343	Bolivia, Santa Cruz
Diploid perennials			
A. cardenasii Krap. et Greg. *nom. nud.*	GKP 10017	262141	Bolivia, Robore
A. chacoense Krap. et Greg. *nom. nud.*	GKP 10602	276235	Paraguay, Puerto Casado
A. correntina (Burk.) Krap. et Greg. *nom. nud.*	K 7830	262137	Argentina, Corrientes
	K 7897	262134	Argentina, Corrientes
	GKP 9530	262808	Argentina, Corrientes
A. villosa Benth.	Bu 22585	298626	Uruguay
A. sp.	GK 30008	468152	Brazil, Mato Grosso
Tetraploid annuals			
A. monticola Krap. et Rig.	K 7264	219824	Argentina, Jujuy
	GKBSPSc 30062	468196	Argentina, Jujuy
A. hypogaea L.			
ssp. *hypogaea* var. *hypogaea* (Virginia)		Florigiant	USA
		Florunner	USA
		Robut 33-1	India
		NC6	USA
		NC7	USA
ssp. *fastigiata* Waldron var. *fastigiata* (Valencia)		New Mexico	USA
		216192-3	Paraguay
		275751	–
var. *vulgaris* Harz. (Spanish)		Argentina	Argentina
		Starr	USA
		244177	Bolivia
		261928	Argentina
Interspecific hybrids			
A. spegazzinii 10038 × *A. duranensis* 30065			
A. duranensis 30065 × *A. stenosperma* 408			
A. spegazzinii 10038 × *A. duranensis* 30075			
A. batizocoi 9484 × *A. cardenasii* 10017			
A. batizocoi 9484 × *A. hypogaea* ssp. *hypogaea* (NC6)			
A. batizocoi 9484 × *A. duranensis* 30065			
A. batizocoi 9484 × *A. glandulifera* 30100			

[a] Abbreviations: Bu, Burkart; GK, Gregory-Krapovickas; GKBSPSc, Gregory-Krapovickas-Banks-Simpson-Pietravelli-Schinini; GKP, Gregory-Krapovickas-Pietravelli; GKSSc, Gregory-Krapovickas-Simpson-Schinini; HL, Hammons-Langford; HLK, Hammons-Langford-Krapovickas; HLKHe, Hammons-Langford-Krapovickas-Hemsy; HLKO, Hammons-Langford-Krapovickas-Ojeda; K, Krapovickas

Differences between species were estimated using Nei's coefficient of genetic distance (Nei 1972). Cluster analysis of the matrix of distance coefficients was carried out by the unweighted pair-group method of Sneath and Sokal (1973) and presented as a dendrogram using the computer program UPGMA developed by Miller (Whitehead Institute, Cambridge, Mass.).

Results

Isozyme banding patterns and their genetic control

Alcohol dehydrogenases (ADH). Diploid and tetraploid species all contained a single ADH band, the position of which was similar in all species. The section *Arachis* is therefore presumed to be monomorphic for a single allele of *Adh.*

Aminopeptidases (AP). Identical banding patterns were obtained with the substrates for both alanine aminopeptidases and leucine aminopeptidases, so the activity detected here is attributed to a single group of generalised aminopeptidases.

Each diploid species had two anodal bands (Fig. 1 A) each in a distinct region of the gel and hence assumed to be controlled by a distinct locus. F_1 hybrids between species which differed in the position of bands in region 1 showed all the bands of both parents but no novel bands. The single available F_2 family segregated into three phenotypes, corresponding to those of the two parents and the F_1. AP in *Arachis* therefore appears to be a monomeric enzyme, controlled by two loci. Variation in the position of bands in each region was attributed to two alleles at *Ap1* and four alleles at *Ap2.*

The tetraploids all had a single band in region 1 and two bands in region 2. They must therefore carry an identical allele at each of the duplicate loci corresponding to *Ap1* (genotype *bbbb* in our notation), but different alleles at the two *Ap2* loci (genotype *bbdd* in our notation). No variation in the number or position of bands in region 2 was found although we examined many different accessions. The wild tetraploid and both subspecies of cultivated tetraploid are therefore considered to be fixed heterozygotes with respect to *Ap2.*

Isocitric dehydrogenases (IDH). A single band, in an apparently uniform position, characterised all species studied. This band is presumed to be controlled by a single locus, *Idh*, with a single allele.

Malate dehydrogenases (MDH). At least three bands were found in leaf extracts of the diploid species. Two were present in all species, did not vary in position, and were attributed to two loci, *Mdh1* and *Mdh2* (Fig. 1 B). Three positions of the third band were detected. F_1 hybrids between species which differed in zymotype for this band showed bands of both parents plus an extra band of intermediate mobility. Region MDH3 thus appears to be a dimeric enzyme controlled by a single locus *Mdh3* with three alleles. The wild and cultivated tetraploids all had three bands in region 3, so were interpreted as fixed heterozygotes for *Mdh3.*

Two additional bands, anodal to the fastest band in region 1, occurred in leaf extracts of two species and pollen extracts of all species. These bands also did not vary and may be post-translational modifications of the products of *Mdh1* and/or *Mdh2.* For the purposes of this analysis, these bands were not attributed to any loci.

Malic enzyme (ME). Leaf extracts showed only a single band of enzyme activity, in region 1, whereas extracts of ground pollen showed bands in two distinct regions of the gel (Fig. 1 c). Pollen leachates usually had bands in region 2 only. This suggests that region 2 contains a cytosolic ME while region 1 contains an organellar ME. The locus coding for the organellar enzyme is designated *Me1* and that coding for the cytosolic enzyme is designated *Me2.* Only one allele was detected at *Me1*, but two were present at *Me2.* Pollen extracts of all plants of the single available accession of *A. cardenasii* and one of the two accessions of *A. correntina* had one band in region 1 but two bands in region 2. These accessions were therefore either uniformly heterozygous for *Me2* or carried a duplicate *Me2* locus. Neither explanation seems very likely, but since perennial peanuts produce very little seed under greenhouse conditions, even if hand-pollinated, no further investigations were possible. Extracts of ground pollen of the tetraploid species had only one band in each region, so in the tetraploids the duplicate loci coding for ME2 must both carry the same allele.

Menadione reductases (MNR). MNR activity occurred in two regions of the gel (Fig. 1 D). In region 2, leaf extracts of most species had a single band in one of two positions. F_1 hybrids between species with different band positions showed the two parental bands and three intermediate bands. Some plants of three of the perennial species also had 5-banded phenotypes in region 2. Extracts from ground pollen of these plants had the fastest and the slowest bands but not the three intermediate bands. These data indicate that bands in region 2 are controlled by a single locus with two alleles, and that the active form of MNR2 is a tetramer. The tetraploid species had a single band in region 2, hence must carry the same allele at both *Mnr2* loci.

Bands in region 1 were much fainter than bands in region 2. Activity in this region was attributed to a different locus, *Mnr1.* Diploid species had a single band in this region, while the tetraploid species and a diploid interspecific hybrid had two bands in region 1. These data suggest that MNR1 is a monomer and that the tetraploids are fixed heterozygotes for *Mnr1.*

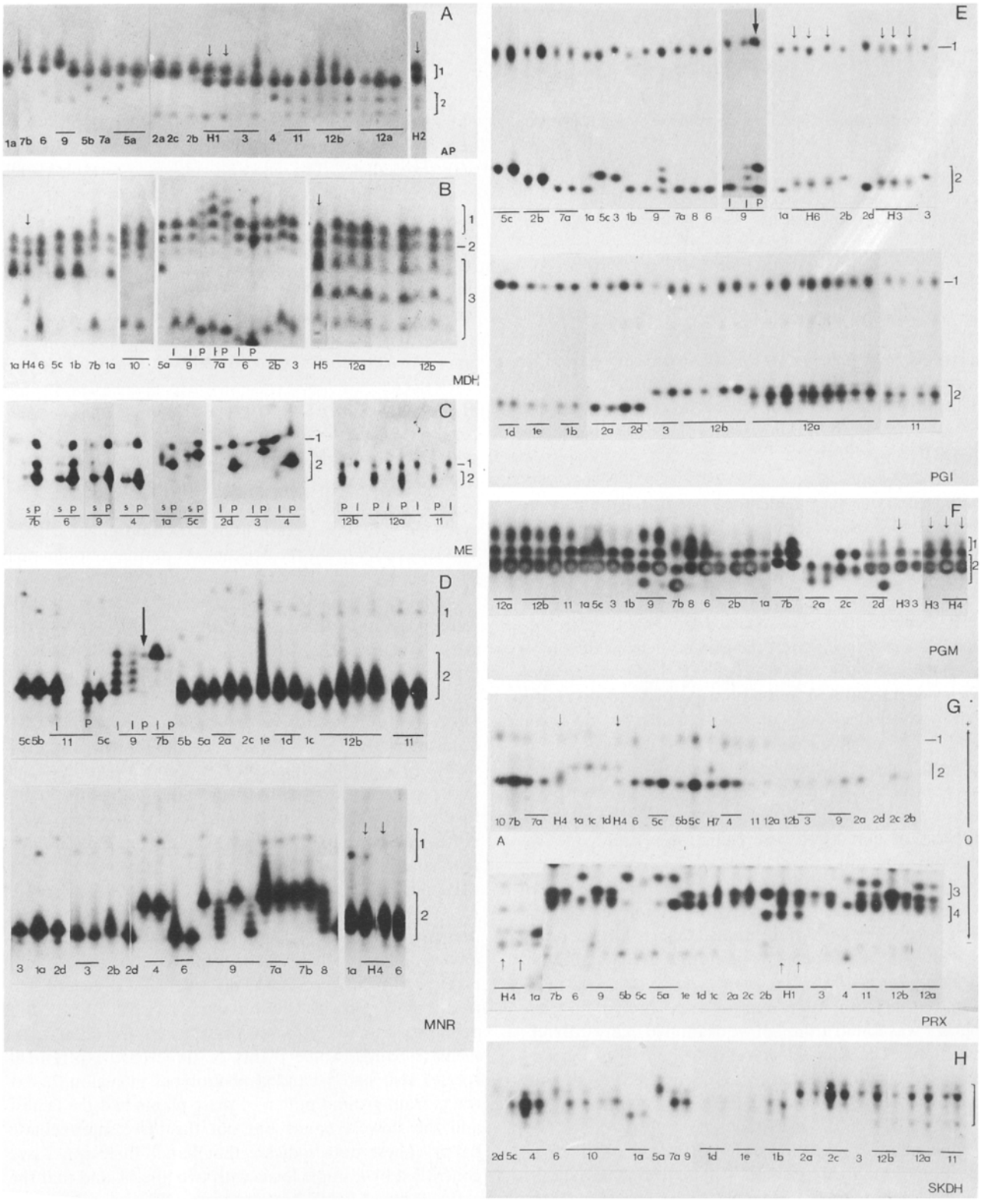

Fig. 1 A–H. Composite photographs of zymograms of eight enzymes in the diploid species, diploid hybrids (*arrow*), and tetraploid species in the section *Arachis*. When a zymogram for one enzyme combines results from different gels, the different gels are separated by a small gap. Identical bands in different gels may migrate to slightly different positions. Each lane represents a single plant. The anode is at the top in all cases and all photographs are of anodal regions, except for peroxidases which include both anodal and cathodal regions. Numbers at the right hand margins of photographs indicate regions of activity controlled by different genetic loci. The symbols assigned to the lanes correspond to those used for species and accessions in Table 2 while the hybrids (*H*) are identified as follows: *H1*, *A. spegazzinii* 10038 (*3*) × *A. duranensis* 30065 (*2b*); *H2*, *A. duranensis* 30065 (*2b*) × *A. stenosperma* 408 (*4*); *H3*, *A. spegazzinii* 10038 (*3*) × *A. duranensis* 30075 (*2d*); *H4*, *A. batizocoi* 9484 (*1a*) × *A. cardenasii* 10017 (*6*); *H5*, *A. batizocoi* 9484 (*1a*) × *A. hypogaea* ssp. *hypogaea* (NC6, *12b*); *H6*, *A. batizocoi* 9484 (*1a*) × *A. duranensis* 30065 (*2b*); *H7*, *A. batizocoi* 9484 (*1a*) × *A. glandulifera* 30100 (*5c*). Where leaf and pollen extracts have been compared for the same accession, these are distinguished as follows: *l* (leaf), *p* (ground pollen), *s*, (pollen leachate). **A**, AP; **B**, MDH; **C**, ME; **D**, MNR; **E**, PGI; **F**, PGM; **G**, PRX; **H**, SKDH

6-phosphogluconate dehydrogenases (6-PGDH). Two invariant bands of 6-PGDH activity were found in all species studied.

Phosphoglucose isomerases (PGI). Two regions of activity were found (Fig. 1 E). Region 1 contained a single band, whose position never varied. In region 2, three band positions were detected in the species examined. Diploid species had a single band, but hybrids between species with different bands had a 3-banded phenotype. Leaf extracts of two plants of one of the perennial species also had three bands in region 2, whereas extracts of ground pollen from these plants had two bands only (intermediate band absent). These results indicate that the active form of PGI2 is a dimer and that activity in region 2 is controlled by a single locus *Pgi2* with three alleles. Activity in region 1 was attributed to a different locus, *Pgi1*, with only one allele.

The two subspecies of the cultivated tetraploid *A. hypogaea* differed consistently in their phenotypes for PGI2. Ssp. *hypogaea* had a single band, suggesting that both *Pgi2* loci were homozygous for the same allele (allele b). Ssp. *fastigiata* had three bands, and so must be a fixed heterozygote (genotype *bbcc*). Both accessions of the wild tetraploid *A. monticola* had the same phenotype as *A. hypogaea* ssp. *fastigiata*.

Phosphoglucomutases (PGM). There were two overlapping regions of activity for PGM (Fig. F). Bands in region 1 were attributed to a single locus, *Pgm1*, with three alleles. A second locus, *Pgm2*, appeared to control a doublet of bands, the fastest of which apparently migrated to the same position on the gel as the slowest band in region 1. Two alleles were detected at *Pgm2*. Phenotypes of F_1 hybrids between species with bands in different positions fitted this interpretation and indicated that the active form of PGM is a monomer. Heterozygous phenotypes were observed in some plants of some wild species.

Peroxidases (PRX). PRX isozymes migrated both anodally and cathodally (Fig. 1 G). There were two anodal regions of activity. Region 1 contained a single invariant band, attributed to *Prx1*. Bands in region 2 occurred in two different positions. Diploid species produced a single band, while hybrids between diploids which differed in band position showed both bands. PRX2 is therefore a monomer, and is apparently controlled by a single locus, *Prx2*, with two alleles. Cathodally, there were also two regions of analysable activity (regions 3 and 4). Each diploid species had one band in each region. Phenotypes of hybrids between appropriate diploids indicated that PRX3 and PRX4 are also monomers. PRX3 is presumed to be controlled by a single locus, *Prx3*, with two alleles. The single available F_2 population segregated for PRX4 only, and contained three phenotypes corresponding to those of the two parents and their F_1 hybrid. The tetraploids have identical genotypes at *Prx1*, *Prx2* and *Prx4* but differ with respect to *Prx3*. Ssp. *hypogaea* contains the same allele (b) at both duplicate loci, whereas ssp. *fastigiata* is a fixed heterozygote (*aabb*). Again, the wild tetraploid *A. monticola* has the same phenotype as ssp. *fastigiata*.

Shikimic dehydrogenases (SKDH). A single region of activity, with bands in three different positions, occurred in gels stained for SKDH (Fig. IH). Diploid species had a single band, while tetraploids had two bands. F_1 hybrids between diploids with different band positions showed the two parental bands only. SKDH therefore appears to be monomeric and to be controlled by a single locus with three alleles. The double-banded phenotype of the tetraploids must then be attributed to fixed heterozygosity.

Genotypic variation within and between species

Table 2 shows the distribution of alleles in accessions and species within the section *Arachis*. Intra-accession polymorphism was very rare and, apart from one example (*Ap1* in *A. chacoense*), consisted entirely of occasional heterozygous plants in an otherwise homozygous accession. These plants were usually heterozygous at a single locus only. Apart from the fixed heterozygosity in the tetraploids, only one example of heterozygosity occurred in an annual species (in *A. duranensis* GKP 30075). All other examples involved the diploid perennials, *A. cardenasii*, *A. correntina*, *A. villosa* and *A. chacoense*.

Polymorphism between different accessions of the same species occurred in all diploids of which more than one accession was available. *Mnr1*, *Pgi2* and *Pgm1* were the loci most likely to vary within species. Some rare alleles occurred in only one accession of one species (e.g., $Pgi2^a$ in a single accession of *A. glandulifera*; $Skdh^c$ in a single accession of *A. batizocoi*). Certain other alleles were confined to a single species, but were present in all accessions of that species (e.g., $Ap2^a$ and $Prx2^a$ in *A. batizocoi*).

The tetraploid species contained no unique alleles. They showed fixed heterozygosity at four (*A. hypogaea* ssp. *hypogaea*) or six (*A. hypogaea* ssp. *fastigiata* and *A. monticola*) of the duplicate loci studied here. No differences were observed within either subspecies, i.e., the Spanish (ssp. *fastigiata* var. *vulgaris*) and Valencia (ssp. *fastigiata* var. *fastigiata*) accessions used in our analyses gave identical banding patterns. Grieshammer and Wynne (1990 b) also found no consistent differences between Spanish and Valencia peanuts in GOT (not studied by us), PGI and IDH. They found variation within Virginia peanuts in IDH (we detected no variation in this enzyme) and attributed this to recent broadening of the genetic

Table 2. Isozyme genotypes in species and accessions of the section *Arachis*

Isozyme loci	*Adh*	*Ap1*	*Ap2*	*Idh*	*Mdh1*	*Mdh2*	*Mdh3*	*Me1*	*Me2*	*Mnr1*	*Mnr2*	*6-Pgdh*	*Pgi1*	*Pgi2*	*Pgm1*	*Pgm2*	*Prx1*	*Prx2*	*Prx3*	*Prx4*	*Skdh*
Diploid annuals																					
A. batizocoi																					
K 9484	+	bb	aa	+	+	+	aa	+	aa	cc	bb	+	+	cc	cc	aa	+	aa	bb	aa	cc
GKP 30081	+	bb	aa	+	+	+	aa	+	aa	dd	bb	+	+	cc	cc	aa	+	aa	bb	aa	bb
GKP 30082	+	bb	aa	+	+	+	aa	+	aa	dd	bb	+	+	cc	cc	aa	+	aa	bb	aa	bb
GKP 30079	+	bb	aa	+	+	+	aa	+	aa	cc	bb	+	+	cc	bb	aa	+	aa	bb	aa	bb
GK 30080	+	bb	aa	+	+	+	aa	+	aa	cc	bb	+	+	cc	bb	aa	+	aa	bb	aa	bb
A. duranensis																					
K 7988	+	aa	dd	+	+	+	bb	+	aa	bb	bb	+	+	cc	–	bb	+	bb	bb	aa	aa
GKP 30065	+	aa	dd	+	+	+	bb	+	aa	bb	bb	+	+	bb	cc	aa	+	bb	bb	bb	aa
GKP 30072	+	aa	dd	+	+	+	bb	+	aa	bb	bb	+	+	cc	bb	aa	+	bb	bb	aa	aa
GKP 30075	+	aa	dd	+	+	+	bb	+	bb	bb	bb	+	+	cc	bb	aa(ab)	+	bb	bb	aa	aa
A. spegazzinii																					
GKP 10038	+	bb	dd	+	+	+	bb	+	aa	bb	bb	+	+	bb	cc	aa	+	bb	bb	aa	aa
A. stenosperma (all accessions)	+	bb	bb	+	+	+	bb	+	bb	aa	aa	+	+	cc	cc	aa	+	bb	aa	aa	bb
A. glandulifera																					
GKSSc 30091	+	bb	bb	+	+	+	aa	+	aa	ee	bb	+	+	bb	aa	aa	+	bb	aa	aa	bb
GKSSc 30099	+	bb	bb	+	+	+	aa	+	aa	ee	bb	+	+	bb	aa	aa	+	bb	aa	aa	bb
GKSSc 30100	+	bb	bb	+	+	+	aa	+	aa	bb	bb	+	+	aa	bb	aa	+	bb	aa	aa	bb
Diploid perennials																					
A. cardenasii																					
GKP 10017	+	bb	bb	+	+	+	cc	+	ab	bb	bb	+	+	cc	bb	aa	+	bb	aa	aa	aa
A. correntina																					
K 7830	+	bb	cc	+	+	+	bb	+	aa	bb	aa(ab)	+	+	cc	bb	aa	+	bb	bb	aa	bb
GKP 9530	+	bb	cc	+	+	+	bb	+	ab	bb	aa	+	+	cc	bb	aa(ab)	+	bb	bb	aa	bb
A. villosa Bu22588	+	bb	cc	+	+	+	bb	+	aa	bb	ab	+	+	cc	bb	aa	+	bb	bb	aa	bb
A. chacoense																					
GKP10602	+	bb(aa)	cc	+	+	+	bb	+	bb	aa	aa(ab)	+	+	cc(bc)	aa	aa(ab)	+	bb	bb	aa	bb
A. sp. GK 30008	+	bb	bb	+	+	+	bb	+	aa	aa	aa	+	+	cc	aa	aa	+	bb	bb	aa	bb
Tetraploids																					
A. monticola (both accessions)	+	bbbb	bbdd	+	+	+	aabb	+	aaaa	ccdd	bbbb	+	+	bbcc	aaaa	aaaa	+	bbbb	aabb	aaaa	aacc
A. hypogaea																					
ssp. *fastigiata* (all accessions)	+	bbbb	bbdd	+	+	+	aabb	+	aaaa	ccdd	bbbb	+	+	bbcc	aaaa	aaaa	+	bbbb	aabb	aaaa	aacc
ssp. *hypogaea* (all accessions)	+	bbbb	bbdd	+	+	+	aabb	+	aaaa	ccdd	bbbb	+	+	bbbb	aaaa	aaaa	+	bbbb	bbbb	aaaa	aacc

Table 3. Genetic distances among the species of the section *Arachis*

Species	*dur.*	*speg.*	*sten.*	*gland.*	*card.*	*corr.*	*vill.*	*chac.*	*sp.30008*	*mont.*	*fast.*	*hyp.*
Diploids												
batizocoi	0.37	0.33	0.42	0.28	0.37	0.30	0.25	0.37	0.32	0.22	0.22	0.25
duranensis		0.12	0.45	0.04	0.25	0.23	0.20	0.33	0.36	0.23	0.23	0.24
spegazzinii			0.41	0.30	0.29	0.28	0.23	0.36	0.34	0.16	0.16	0.13
stenosperma				0.30	0.29	0.26	0.29	0.18	0.15	0.33	0.33	0.43
glandulifera					0.22	0.33	0.28	0.35	0.26	0.13	0.13	0.17
cardenasii						0.28	0.25	0.41	0.36	0.22	0.22	0.31
correntina							0.01	0.15	0.17	0.30	0.30	0.33
villosa								0.17	0.17	0.25	0.25	0.27
chacoense									0.12	0.26	0.26	0.27
sp.30008										0.20	0.20	0.22
Tetraploids												
monticola											0.00	0.03
ssp. *fastigiata*												0.03
ssp. *hypogaea*												

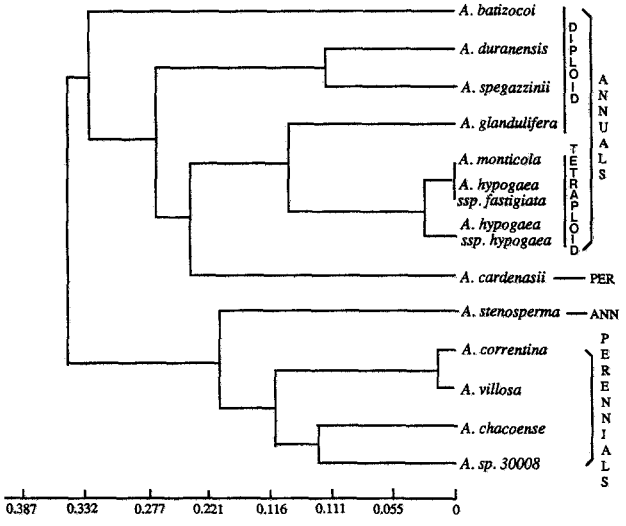

Fig. 2. Dendrogram among 12 species of the section *Arachis* based on Nei's genetic distances calculated from data on 21 isozyme-coding loci

base of Virginia peanuts through intersubspecific crossing.

When identical alleles were present at both loci in the tetraploids, these were usually (but not always) most widespread among the diploid species. No diploid species or accession could be unequivocally identified as a progenitor of the tetraploids. Those most similar to the tetraploids were *A. spegazzinii* (A genome) and *A. glandulifera* (D genome).

Genetic distance among species in the section Arachis

Genetic distances were calculated between all pairs of species (Table 3). Cluster analysis of this matrix produced a dendrogram which separates the species into two groups, annuals versus perennials, though one perennial species (*A. cardenasii*) has been grouped with the annuals, and *A. stenosperma* has been included with the perennials (Fig. 2). *A. stenosperma* is placed in the series Annuae by taxonomists but is considered by Stalker (personal communication) to be intermediate between the true annuals and the perennials. The two subspecies of *A. hypogaea* and the wild tetraploid *A. monticola* are separated by very small distances, as are *A. villosa* and *A. correntina*. Smartt and Stalker (1982) have previously suggested that *A. correntina* be treated as a subspecies of *A. villosa*. Similarly *A. monticola* and *A. hypogaea* are probably not distinct species (Pickersgill 1986).

The most distinct species in this analysis is the only diploid representative of the B genome, *A. batizocoi*. The single D-genome diploid, *A. glandulifera*, clusters with the tetraploids rather than with either the A or the B genome diploids.

Discussion

Allelic diversity and differentiation

Wild populations were not sampled directly for these studies, and the accessions received from germplasm collections are likely to contain much less diversity than the natural populations. The data reported here must therefore be interpreted very cautiously. However, the limited heterozygosity and limited intra-accession polymorphism, compared to polymorphism within or between the species studied, do agree with what is known of the breeding systems of these species.

The annual species, both diploid and tetraploid, have stigmas which are automatically self-pollinated, often be-

fore the flower opens (Lu et al. 1990), though insect visits after anthesis may result in some cross-pollination. Recorded levels of outcrossing in the cultivated peanut are about 2.5% (Norden 1973). Only one heterozygote (other than the fixed heterozygotes found in the tetraploids) occurred in all the annual plants analysed. The perennial species have minute stigmas guarded by hairs which prevent self pollen reaching the stigma unless the flower is manipulated in some way (Lu et al. 1990). Any seeds set by the perennials presumably result from insect-mediated cross-pollination, although many of these pollinations may be geitonogamous. We found several heterozygotes amongst the perennials although we cannot determine whether both alleles at these loci occurred in the original wild populations or whether one was acquired by outcrossing with other accessions in the germplasm collections from which our seeds were obtained.

Polymorphism occurs between, rather than within, accessions of both annuals and perennials in the section *Arachis*, as would be expected if both groups of species are effectively inbred (Gottlieb 1981). This suggests that in *Arachis*, as in other predominantly inbred herbaceous plants (Marshall and Brown 1975, 1983), genetic diversity will be captured more efficiently by increasing the numbers of populations sampled rather than the numbers of plants per population.

Only a few alleles are diagnostic of particular species. However, all species (except those of doubtful taxonomic validity) could be identified by their multilocus genotypes when several enzymes were surveyed simultaneously. Species-specific markers are particularly valuable in *Arachis* because many wild species still lack formal descriptions. Isozyme phenotypes could provide one means of ensuring that accessions in living collections are correctly identified and remain true to type.

Origin(s) of A. hypogaea

The fixed heterozygosity present at some, but by no means all, of the duplicate loci in the tetraploids agrees with the widespread view that the tetraploids are allo- or segmental allo-tetraploids. It is not yet possible to determine which allele at these duplicate loci belongs to which genome, hence it is not possible to propose the zymotypes to be expected in the diploid progenitors of the tetraploids. However, none of the accessions included in this study proved a perfect match for the isozyme characters expected in the ancestors of the tetraploids.

Cytological data (Smartt et al. 1978; Singh and Moss 1984; Singh 1988) suggest that the tetraploids contain the A and B genomes. The only known B-genome diploid is *A. batizocoi*. The tetraploids contain three alleles (*Mnr1*^{c}, *Mnr1*^{d} and *Skdh*^{c}) which we found otherwise only in *A. batizocoi*. However, the tetraploids lack two alleles (*Ap2*^{a} and *Prx2*^{a}) which characterised all accessions of *A. batizocoi* that we studied and should therefore mark the *batizocoi* genome if this is present in the tetraploids. The tetraploids are also monomorphic for one allele (*Pgm1*^{a}) which was not present in any accession of *A. batizocoi* available to us. Studies on RFLPs (Kochert et al., 1991) agree with the isozyme evidence in not supporting *A. batizocoi* as a genome donor to the tetraploids. On isozyme evidence alone, the D-genome diploid *A. glandulifera* is a more plausible source of the non-A genome than *A. batizocoi*, because it differs consistently from the tetraploids at only two of the loci studied (*Mnr1* and *Skdh*). However, cytological data and crossing behaviour appear to eliminate *A. glandulifera* as a progenitor of the tetraploids (Stalker 1991).

The DNA studies of Kochert et al. (1991) and Halward et al. (1992) suggest that the tetraploids might have arisen from a cross between two A-genome diploids, possibly related to the annual species *A. ipaensis*, *A. duranensis* and *A. spegazzinii*. *A. ipaensis* was not included in our analysis, but *A. spegazzinii* is the A-genome species which is isozymically most similar to the tetraploids. It differs at two loci, *Mnr1* and *Pgm1*, both of which varied within, as well as between, species in this study. Only one accession of *A. spegazzinii* was available to us, so it is possible that other accessions of this species may contain the *Mnr1* and *Pgm1* alleles present in *A. hypogaea*. The isozyme data do not implicate the A-genome perennials as possible ancestors of the tetraploids, but do not rule them out either.

The isozyme genotypes of the tetraploids may have diverged from those of their progenitor diploids by either mutation or hybridization. Mutation seems unlikely, since the tetraploids contain no unique alleles, but only alleles found in one or more of the diploid species. Diploid and tetraploid species of the section *Arachis* can be crossed and the F_1 may produce occasional viable offspring (Singh and Moss 1984). Krapovickas et al. (1974) described such a hybrid, involving *A. batizocoi* and *A. hypogaea*, which after 4 years generated some fully fertile tetraploid derivatives morphologically similar to *A. monticola*. They suggested that *A. monticola* could therefore have originated through crossing between *A. hypogaea* and a wild diploid.

This hypothesis could explain the isozyme differences between the two subspecies of *A. hypogaea*. The original tetraploid could have arisen from two progenitor diploid species, then spread in association with man into the range of a third diploid, with which it hybridised to produce a taxonomically distinct subspecies. An alternative possibility, which cannot be eliminated on the basis of our data, is that there were two independent origins of the tetraploids, possibly involving the same B-genome diploid but different A-genome species (Singh and Moss 1984; Singh 1988).

Much remains to be learned about the distribution and diversity of the diploid relatives of the cultivated peanut. Certainly the samples which we have studied are too limited to support far-reaching conclusions on the origin of the tetraploids or on species relationships within the section *Arachis*. Recent collections, which have not yet been multiplied or distributed, reportedly include more undescribed species, as well as more accessions of some species hitherto known from very limited material. Study of this new material will advance our understanding of the genetic diversity and genetic resources in this group and may alter the picture suggested here.

Acknowledgements. We thank Dr. H. T. Stalker, North Carolina State University, U.S.A., and Dr. J. P. Moss, ICRISAT, India for supplying the species used in this study. Some of the work reported here was supported by a scholarship to J. Lu from the Department of Agriculture, P. R. China, while the later stages were supported by the International Board for Plant Genetic Resources, to both of whom we are extremely grateful.

References

Bunting AH (1955) A classification of cultivated groundnut. Emp J Exp Agric 23:158–170

Bunting AH, Elston J (1980) Ecophysiology of growth and adaptation in the groundnut: an essay on structure, partition and adaptation. In: Summerfield RJ, Bunting AH (eds) Advances in legume science. Royal Botanic Gardens, Kew, pp 495–500

Cherry JP (1975) Comparative studies of seed proteins and enzymes of species and collections of *Arachis* by electrophoresis. Peanut Sci 2:58–65

Conkle MT, Hodgskiss PD, Nunnally LB, Hunter SC (1982) Starch-gel electrophoresis of conifer seeds: a laboratory manual. General Tech Rep PSW-64. Pacific Southwest Forest and Range Experiment Station, US Department of Agriculture. Berkeley, California

González de León DR (1986) Interspecific hybridization and the cytogenetic architecture of two species of chili pepper (*Capsicum* – Solanaceae). PhD thesis, Reading University, England

Gottlieb LD (1981) Electrophoretic evidence and plant populations. Prog Phytochem 7:1–46

Gregory WC, Gregory MP (1976) Groundnuts. In: Simmonds NW (ed) Evolution of crop plants. London, Longmans, pp 151–154

Gregory MP, Gregory WC (1979) Exotic germplasm of *Arachis* L.: interspecific hybrids. J Hered 70:185–193

Gregory WC, Gregory MP, Krapovickas A, Smith BW, Yarbrough JA (1973) Structures and genetic resources of peanuts. In: Am Peanut Res Ed Assoc (ed) Peanut: culture and uses. Stillwater, Oklahoma, pp 47–133

Gregory WC, Krapovickas A, Gregory MP (1980) Structure, variation, evolution and classification in *Arachis*. In: Summerfield RJ, Bunting AH (eds) Advances in legume science. Royal Botanic Gardens, Kew, pp 469–481

Grieshammer U, Wynne JC (1990a) Mendelian and non-Mendelian inheritance of three isozymes in peanut (*Arachis hypogaea* L.). Peanut Sci 17:101–105

Grieshammer U, Wynne JC (1990b) Isozyme variability in mature seeds of U.S. peanut cultivars and collections. Peanut Sci 18:72–75

Halward TM, Stalker HT, LaRue EA, Kochert G (1991) Genetic variation detectable with molecular markers among unadapted germ-plasm resources of cultivated peanut and related wild species. Genome 34:1013–1020

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

Kochert G, Halward T, Branch WD, Simpson CE (1991) RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. Theor Appl Genet 81:565–570

Krapovickas A (1969) The origin, variability and spread of the groundnut (*Arachis hypogaea*). In: Ucko PJ, Dimbleby GW (eds) The domestication and exploitation of plants and animals. Duckworth, London, pp 427–441

Krapovickas A (1973) Evolution of the genus *Arachis*. In: Moav R (ed) Agricultural genetics, selected topics. National Council for Research and Development, Jerusalem, pp 135–151

Krapovickas A, Fernandez A, Seeligman P (1974) Recuperación de la fertilidad en un híbrido interspecífico estéril de *Arachis* (Leguminosae). Bonplandia 3:129–142

Krishna TG, Mitra R (1988) The probable genome donors to *Arachis hypogaea* L. based on arachin seed storage-protein. Euphytica 37:47–52

Lu J, Mayer A, Pickersgill B (1990) Stigma morphology and pollination in *Arachis* L. (Leguminosae). Ann Bot 66:73–82

Marshall DR, Brown ADH (1975) Optimum sampling strategies for genetic conservation. In: Frankel OH, Hawkes JG (eds) Crop genetic resources for today and tomorrow. University Press, Cambridge, pp 53–80

Marshall DR, Brown ADH (1983) Theory of forage plant collection. In: McIvor JC, Bray RA (eds) Genetic resources in forage plants. CSIRO, Melborne, Australia, pp 135–148

Nei M (1972) Genetic distance between populations. Am Nat 106:283–292

Norden AV (1973) Breeding of the cultivated peanut (*A. hypogaea* L.). In: Peanuts – cultures and uses. Am Peanut Res Ed Assoc, Stillwater, Oklahoma, pp 175–208

Pickersgill B (1986) Evolution of hierarchical variation patterns under domestication and their taxonomic treatment. In: Styles BT (ed) Infraspecific classification of wild and cultivated plants. Clarendon Press, Oxford, pp 191–209

Roose ML, Gottlieb LD (1978) Stability of structural gene number in diploid species with different amounts of nuclear DNA and different chromosome numbers. Heredity 40:159–163

Singh AK (1988) Putative genome donors of *Arachis hypogaea* (*Fabaceae*), evidence from crosses with synthetic amphidiploids. Plant Syst Evol 160:143–151

Singh AK, Moss JP (1982) Utilization of wild relatives in genetic improvement of *Arachis hypogaea* L. 2. Chromosome complements of species in the section *Arachis*. Theor Appl Genet 61:305–314

Singh AK, Moss JP (1984) Utilization of wild relatives in the genetic improvement of *Arachis hypogaea* L. 5. Genome analysis in the section *Arachis* and its implications in gene transfer. Theor Appl Genet 68:355–364

Singh AK, Sivaramakrishnan S, Mengesha MH, Ramaiah CD (1991) Phylogenetic relations in the section *Arachis* based on seed protein profile. Theor Appl Genet 82:593–597

Smartt J, Stalker TH (1982) Speciation and cytogenetics in *Arachis*. In: Pattee HE, Young CT (eds) Peanut science and technology. Am Peanut Res Ed Assoc, Yoakum, Texas pp 21–49

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

Soltis DE, Haufler CH, Darrow DC, Gastony GJ (1983) Starch-gel electrophoresis of ferns: a compilation of grinding

buffers, gel and electrode buffers and staining schedules. Am Fern J 73:9–27
Stalker HT (1990) A morphological appraisal of wild species in the section *Arachis* of peanuts. Peanut Sci 17:117–122
Stalker HT (1991) A new species in the section *Arachis* of peanuts with a D genome. Am J Bot 78:630–637
Stalker HT, Dalmacio RD (1986) Karyotype analysis and relationships among varieties of *Arachis hypogaea* L. Cytologia 51:617–629
Vallejos CE (1983) Enzyme activity staining. In: Tanksley SD, Orton TJ (eds) Isozymes in plant genetics and breeding, Part A. Elsevier, Amsterdam, pp 469–516
Varisai Muhammad S (1973) Cytogenetical investigations in the genus *Arachis* L. III. Tetraploid interspecific hybrids and their derivatives. Madras Agric J 60:1428–1432
Weeden NF, Wendel JF (1989) Genetics of plant isozymes. In: Soltis DE, Soltis P (eds) Isozymes in plant biology. Dioscorides Press, Portland, Oregon pp 46–72