Theor Appl Genet (1986) 73:82–87



# Variation and inheritance of the arachin polypeptides of groundnut (*Arachis hypogaea* L.)

**T. G. Krishna, S. E. Pawar and R. Mitra**

Nuclear Agriculture Division, Bhabha Atomic Research Centre, Trombay, Bombay-400 085, India

Received May 28, 1986; Accepted June 22, 1986
Communicated by R. Riley

**Summary.** Variation in the arachin polypeptides of groundnut genotypes was observed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Three regions could be observed on the electropherogram. Region 1, corresponding to conarachin, did not show any variation; region 2, consisting of arachin acidic subunits, showed variation; region 3, containing the arachin basic subunits, did not show any variation. There are four varietal classes of arachin polypeptide patterns: class A comprised three acidic subunits of arachin of molecular weights 47.5, 45.1 and 42.6 kd and a basic subunit of 21.4 kd; class B, with three acidic subunits of molecular weights 47.5, 45.1 and 41.2 kd and a basic subunit of 21.4 kd; class C of an additive pattern of class A and class B; class D, of two acidic polypeptides of 47.5, 45.1 kd and the basic 21.4 kd subunit. Of the 90 genotypes studied, 73% belong to class A, 15% to class B and 6% each to class C and D. Analysis of $F_2$ seeds from a cross between class A and class B genotypes showed that the two polypeptides (42.6 kd and 41.2 kd) are coded by nonallelic genes and also revealed that class C and class D patterns arose as a result of hybridisation between class A and class B. *A. monticola,* the progenitor of *A. hypogaea,* showed a pattern similar to the additive pattern of class A and class B while some diploid *Arachis* species had the 41.2 kd polypeptide. Based on arachin polypeptide patterns the probable origin of *A. hypogaea* has been suggested.

**Key words:** *Arachis hypogaea* – Groundnut – Storage protein – Arachin polypeptides – SDS-PAGE – Variation – Inheritance

## Introduction

In groundnut, about 87% of the total seed protein consists of a globulin fraction (see Basha and Pancholy 1981; Mossé and Pernollet 1983) with two major fractions, arachin and conarachin (Johns and Jones 1916), corresponding to the legumin and vicilin fractions respectively of other legume seeds. Arachin exists as a monomer as well as a dimer. However, both the monomer and dimer consists of 3 kinds of polypeptides of about 60 kd which associate to make a holoprotein of 180 kd (monomer) and 360 kd (dimer) (Yamada et al. 1979). Each polypeptide chain is composed of a 40 kd acidic and a 20 kd basic subunit linked together by disulphide bonds (Krishna et al., unpublished observation). Conarachin is a complex protein consisting of one kind of polypeptide (Yamada et al. 1980).

Savoy (1976) used SDS-PAGE to isolate groundnut seed protein in 45 genotypes. He reported a general electrophoretic pattern with respect to numbers and kinds of polypeptides. The genotypes could be grouped into four classes depending on the relative concentrations of six main polypeptides. However, no attempt was made to identify whether the variation was in the arachin or conarachin protein fractions. Cherry (1977) analysed 21 groundnut cultivars and many wild *Arachis* species for their protein electrophoretic patterns on native polyacrylamide gel. He could not detect significant variability in the arachin region. However, variable patterns were reported in wild *Arachis* species. Basha (1979), using two dimensional gel electrophoresis, showed variation among the polypeptides of groundnut seed protein obtained from 11 cultivars and 4 breeding lines. However, the differences in polypeptide composition were not identified for specific fractions.

During the course of screening the groundnut germplasm and some wild *Arachis* species for arachin polypeptide variation, we detected different patterns. The inheritance of a variant arachin polypeptide in a

**Table 1.** Type of arachin banding pattern

| Class | MW (kd) | Schematic diagram | Genotypes |
|---|---|---|---|
| A | 47.5 | ——— | 'Spanish improved', J-11, JL-24, |
| | 45.1 | ——— | 'Chico', 'Gujaral Dwarf', 'Starr', |
| | 42.6 | ——— | 'Tennessee Red', 'Florunner', 'Dixie Spanish', 'Spanhoma' |
| | 21.4 | ——— | 'Tifspan', TG-2 to TG-6, Tall, Dark green, Small leaf, Virescent, TG-14, TG-15 TG-E-2, TG-20, TG-21 ICGS-1 to 3, ICGS-7 to 16, ICGS-21 to 28, ICGS-30, ICGS-32, ICGS-33, ICGS-35, ICGS-36 to 44, ICGS-52, ICGS-53, ICGS-57 to 61 |
| B | 47.5 | ——— | M-13 'Shulamit', 'Chalimbana', |
| | 45.1 | ——— | TG-1, TG-1A, TG-7 TG-17, TG-19, TG-19A |
| | 41.2 | ——— | TG-E-3, TG-S-2, TG-S-3 ICGS-31, ICGS-51 |
| | 21.4 | ——— | |
| C | 47.5 | ——— | TG-8 (TG-1 × Virescent), |
| | 45.1 | ——— | TG-9 (TG-1 × Virescent), |
| | 42.6 | ——— | TG-10 (TG-1 × Virescent), |
| | 41.2 | ——— | TG-13A (LV3 [large pod × Virescent] × TG-13 (TG-1 × |
| | 21.4 | ——— | Virescent), TG-E-1 (Tall × TG-9). *A. monticola* |
| D | 47.5 | ——— | AU-2, TG-16 (Virescent × |
| | 45.1 | ——— | TG-1), TG-18 (Dark green × TG-1), TG-18A (mutant of TG-18), TG-S-1 (TG-18A × |
| | 21.4 | ——— | M-13) |
| | 45.1 | ——— | *A. duranensis, A. cardenasii* |
| | 41.2 | ——— | |
| | 21.4 | ——— | |

TG ('Trombay Groundnut') cultures developed at this Research Centre (Mouli et al. 1983)
ICGS-material from ICRISAT collection

cross between 'Spanish Improved' (SP) and 'Trombay Groundnut 1' (TG-1) was studied. The result of the screening and inheritance studies are presented here.

## Materials and methods

*Seed material*

The groundnut genotypes used in this study are listed in Table 1.

*Seed meal preparation*

Seeds without testa were defatted with solvent ether or n-hexane at room temperature by grinding in a pestle and mortar. The dried defatted meal was stored under vaccum at room temperature until further use.

*Protein extraction and electrophoresis*

100 mg of defatted seed meal was extracted with 5 ml of extraction buffer [0.1 M Tris-HCl, pH 6.8 containing 2% sodium dodecyl sulphate (SDS) and 10% glycerol] at room temperature for about 3 h. For inheritance studies, a small portion of the seed (about 200 mg) distal from the embryonic axis was cut with a sharp razor blade and used for protein extraction (1 ml buffer). The slurry was centrifuged at 12,000 g for 20 min and the supernatant was used for further analysis. Protein in the extract was estimated by the biuret method using bovine serum albumin (fraction V) as the reference protein (Layne 1957).

Aliquots of the extract were mixed with an appropriate volume of diluent consisting of 0.05 M Tris-HCl, pH 7, 20% glycerol, 1% SDS, 5% 2-mercaptoethanol and 0.25 mg/ml coomassie brilliant blue to obtain a final protein concentration of 1 µg/µl. The mixture was heated at 100 °C for 3 min before loading 20 µg protein in each track on 12% homogeneous, discontinuous, SDS-polyacrylamide slab gel system according to the method of Laemmli (1970). Electrophoresis was performed at constant voltage for 650–700 volt-hours in a Pharmacia GE 2/4 apparatus. After electrophoresis gels were fixed and stained in 0.2% coomassie brilliant blue R 250 in methanol, acetic acid and water (30 : 10 : 60 v/v) and destained in the same solution without the dye. The apparent molecular weights of the polypeptides (average of at least six different separations) were determined from the calibration curve prepared using the Pharmacia molecular weight calibration kit.

*Quantification of polypeptides*

Coomassie brilliant blue dye bound to the polypeptide was eluted from the polyacrylamide gel and quantified according to the method of Hames (1982).

*Hybridisation*

The parental arachin polypeptide pattern was ascertained by electrophoresis using a portion of the seed. The remaining part of the seed was used to raise plants in pots. The hybridisation programme was carried out following the methods described by Patil (1971). $F_1$ and $F_2$ seeds were grown in the field; $F_2$ and $F_3$ seeds were collected from single plants.

## Results and discussion

Variations observed in groundnut seed proteins are shown in Fig. 1. Three regions, designated 1, 2 and 3, were observed after SDS-PAGE. To identify these polypeptides total seed protein extracted from SP with 10 mM phosphate buffer, pH 7.9, was fractionated on a sucrose density gradient a high ionic strength. This separates arachin from conarachin (Yamada et al. 1980). The protein was resolved into three major fractions with s values of 14.1 s, 9.6 s and 3.7 s (Fig. 2). By subjecting the protein around each peak to electrophoresis (Fig. 3) we identified that the 14.1 s protein has polypeptides around 43 and 20 kd; the 9.6 s protein shows a major polypeptide at 66 kd, and the protein of the 3.7 s peak consists mainly of small molecular weight polypeptides. It has been reported earlier that arachin

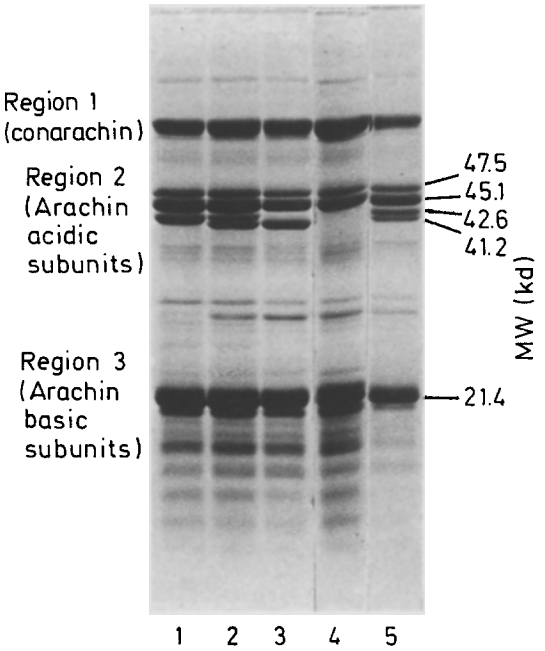


**Fig. 1.** SDS-polyacrylamide gel electrophoretic patterns of seed proteins from groundnut genotypes. The different tracks show (1) SP, (2) TG-8, (3) TG-1, (4) TG-18, (5) mixture of SP and TG-1. 20 μg protein was loaded in each track except for track 5 (10 μg). The numbers in the margin indicate molecular weights of the arachin polypeptides

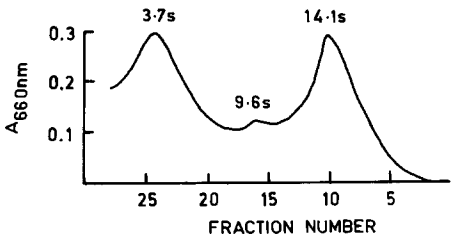


**Fig. 2.** Sedimentation pattern of groundnut seed protein from SP on a 10–30% sucrose density gradient in 10 mM phosphate buffer, pH 7.9 containing 10% NaCl. Protein in the fractions was estimated using the method of Lowry et al. (1951). The direction of sedimentation is from left to right. S values were calculated with reference to catalase following the method of Martin and Ames (1961)

**Table 2.** Arachin-subunit concentration as a percentage of the holoprotein

| | 'Spanish improved' | TG-18 |
|---|---|---|
| Subunit No. 1 (47.5 kd) | 14.3 | 16.9 |
| Subunit No. 2 (45.1 kd) | 28.4 | 45.8 |
| Subunit No. 3 (42.6 kd) | 24.1 | – |
| Subunit No. 4 (21.4 kd) | 33.2 | 37.3 |

SDS-electrophoresis, coomassie blue staining and elution of bound dye from the gel pieces containing the arachin subunits

protein has an s value around 14 s and conarachin an s value around 9 s (Yamada et al. 1980).

No significant qualitative variation in the conarachin polypeptide (region 1, Fig. 1) and the arachin basic subunit (region 3, Fig. 1) was observed among the genotypes studied. The three acidic subunits of the arachin protein (region 2 in Fig. 1) in SP have molecular weights of 47.5, 45.1 and 42.6 kd, respectively. In TG-1 the third arachin polypeptide moves slightly ahead, indicating that the molecular weight of this polypeptide is lower than that from SP. The molecular weight of the third polypeptide of TG-1 was estimated to be 41.2 kd. The fact that the third arachin polypeptide of TG-1 is smaller than the corresponding polypeptide from SP was confirmed by co-electrophoresis (Fig. 1, lane 5). In TG-18 and cv. AU-2 the third acidic polypeptides of arachin protein were absent (Fig. 1, lane 4). Even overloading the sample did not reveal any trace of the third polypeptide. In TG-8 the arachin had four acidic polypeptides (Fig. 1, lane 2). This pattern in TG-8 is similar to the additive pattern of SP and TG-1. Thus, four basic types of arachin polypeptide patterns were observed in the various genotypes studied (Table 1). The number of polypeptides observed is not due to proteolysis since the electrophoretic pattern was identical when the protein was extracted both in the presence and the absence of such proteolytic inhibitors as PMSF.

In TG-18 the second polypeptide has increased in amount as judged by the increased concentration of the bound dye (Table 2) suggestive of compensation for the lack or nonexpression of the gene(s) encoding the third polypeptide. This genotype offers an opportunity to study seed protein gene regulation.

It is inferred that the SP (class A) arachin polypeptide pattern is the more prevalent type (66 out of 90 genotypes) and that the TG-1 pattern (class B) could be a variant form. The class C (TG-8 pattern) and class D (TG-18 pattern) forms could have arisen as a result of hybridization between class A and class B. In class C both third polypeptide specific gene(s) are present while in class D both gene(s) are absent. Evidence for the above conclusion comes from the data obtained from the genetic crosses involving the genotypes showing the variant polypeptides.

When SP was crossed with TG-1 the hybrid seed showed both the 42.6 and 41.2 kd polypeptide of SP and TG-1, respectively (Fig. 4). The reciprocal cross revealed a similar pattern. The $F_2$ seeds derived from the above hybrid showed four phenotypes (Fig. 5). The first resembled the $F_1$ pattern with four bands (Fig. 5, tracks 6, 7, 8, 10 and 11); the second and third were identical to the parent SP and TG-1 (Fig. 5, tracks 1, 2, 3, 5 and 9) respectively; the fourth type had only two arachin subunits (Fig. 5, track 4). Here both the 42.6 and 41.2 kd polypeptides were absent. The 49 $F_2$ seeds

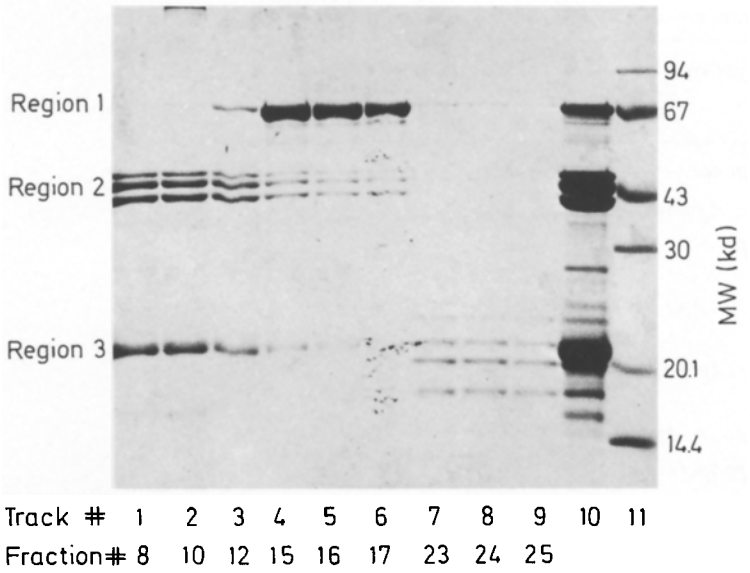


**Fig. 3.** SDS-gel electrophoresis of the protein components fractionated by sucrose density gradient centrifugation. The fraction numbers correspond to those in Fig. 2. Tracks 1 to 9 are sucrose density gradient fractions. Track 10 contains the unfractionated seed protein from SP. Track 11 contains the molecular weight markers. Numbers at the margin indicate the molecular weight in kilodaltons

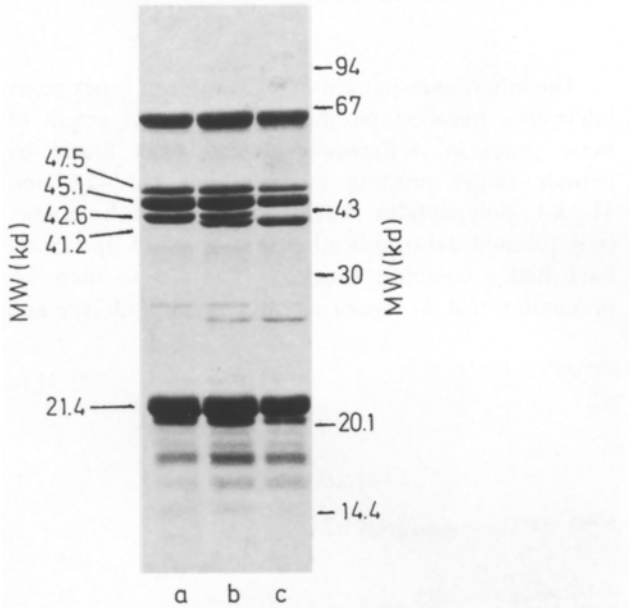


**Fig. 4.** SDS-PAGE of seed proteins from groundnuts. (*a*) SP, (*b*) $F_1$ (SP × TG-1). (*c*) TG-1. 20 µg protein was loaded in each track. Numbers in the *left margin* indicate molecular weights of arachin polypeptides, numbers in the *right margin* indicate the position occupied by marker proteins of these particular weights

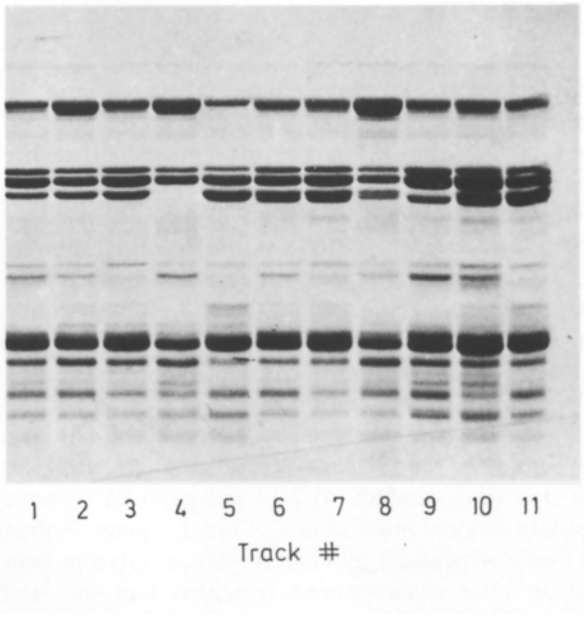


**Fig. 5.** SDS-PAGE of seed proteins from $F_2$ seeds obtained from a cross between SP and TG-1. Each track represents the protein pattern of $F_2$ individual seed. 20 µg protein was loaded in each track

analysed could be grouped into four phenotypes in the ratio of 26 : 11 : 9 : 3 ($F_1$ type : SP : TG-1 : null type, respectively).

When the $F_2$ electrophoretic patterns were scored for the presence or absence of the SP band a ratio of 37 : 12 was obtained while for TG-1 the ratio was 35 : 14.

Some $F_2$ seeds selected after electrophoretic analysis were carried forward to the next generation. Analysis of the $F_2$ seeds having the null type bred true in $F_3$. $F_2$ seeds showing the $F_1$ hybrid pattern segregated into four phenotypic classes as in $F_2$. The parental types from $F_2$ seeds also bred true in the $F_3$.

Thirty-five seeds obtained from the test cross (null × $F_1$) showed phenotypes in the ratio of 7 : 6 : 7 : 15 (SP : TG-1 : $F_1$ : null type, respectively).

Upon careful observation of electrophoretic patterns of the $F_2$ seeds we could observe intensity differences in

**Table 3.** Phenotypic classes and genotypic frequencies from the cross SP×TG-1 in $F_2$ generation

| Phenotypic class | Genotype | Frequencies | |
|---|---|---|---|
| | | Expected | Observed |
| 'Spanish Improved' | *ApcApc/– –* | 3.06 | 6 |
| TG-1 | *– –/ApdApd* | 3.06 | 2 |
| SP pattern with reduced intensity of third polypeptide | *Apc –/– –* | 6.13 | 6 |
| TG-1 patterns with reduced intensity of third polypeptide | *– –/– Apd* | 6.13 | 6 |
| Additive pattern of SP and TG-1 but intensity stronger than $F_1$ | *ApcApc/ApdApd* | 3.06 | 6 |
| $F_1$ pattern | *Apc –/– Apd* | 12.24 | 10 |
| Both bands present but TG-1 band with reduced intensity | *ApcApc/– Apd* | 6.13 | 6 |
| Both bands present but SP band with reduced intensity | *Apc –/ApdApd* | 6.13 | 6 |
| Both band absent | *– –/– –* | 3.06 | 3 |

*P* from $\chi^2_8$ table = 0.70–0.50

the 42.6 and 41.2 kd polypeptides. In other words, we could distinguish between seeds having one or two doses of the gene coding for these polypeptides. The phenotypic classes and the genotypic frequencies of the 42.6 and 41.2 kd polypeptide of SP and TG-1 in $F_2$ generation are given in Table 3. It is suggested that the 47.5 kd arachin acidic polypeptide is coded by a gene *Apa* (Arachin protein-a) while the 45.1 kd polypeptide is coded by the *Apb* gene. Both these genes are present in the wild as well as in the variant type. The 42.6 kd polypeptide is coded by *Apc* in the wild type and is null in the variant type. The gene *Apd* codes for the 41.2 kd polypeptide in the variant type and is null in the wild. The *Apd* gene product (41.2 kd polypeptide) is not a post-translational modification of the *Apc* gene product (42.6 kd polypeptide) by deglycosylation. Glycoprotein staining after electrophoresis revealed that the *Apd* gene product is positive to periodic Schiff stain, just like the *Apc* gene product (unpublished data).

When we score the $F_2$ pattern for the presence or absence of either the *Apc* or *Apd*-coded polypeptide we get a fit for 3 : 1 ratio (for SP, $P = 0.95–0.90$; for TG-1, $P = 0.70–0.50$) showing that the two genes, *Apc* and *Apd*, segregated in simple monogenic ratios. However, if we consider both genes together, the ratio obtained fits a dihybrid ratio (for 9 : 3 : 3 : 1 $P = 0.90–0.80$). The appearance of the null type (both *Apc* and *Apd* absent) in $F_2$ clearly indicated that the two genes are non allelic. The test cross analysis demonstrated that the two genes are not linked (for 1 : 1 : 1 : 1 ratio $P =$ 0.20–0.10). The low probability value could be due to a high frequency of null types which could include some selfed seeds.

The inheritance pattern of *Apc* and *Apd* genes poses interesting question on the homology and origin of these genes in *A. hypogaea*. It has been found by peptide finger printing analysis that the 42.6 and 41.2 kd polypeptides have considerable homology (unpublished data) indicating that *Apc* and *Apd* could have had a common ancestor. If this is so then the progenitor of *A. hypogaea* should contain both *Apc* and

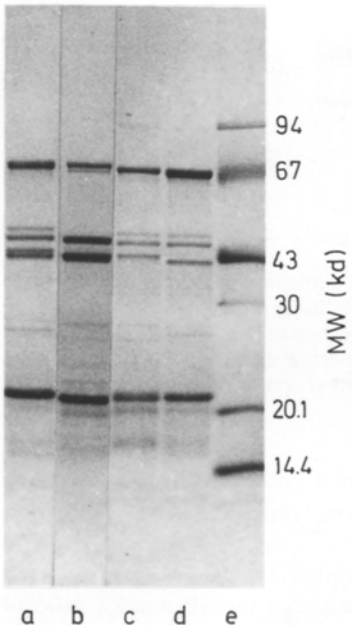


**Fig. 6.** SDS-PAGE of seed extracts from (*a*) *A. monticola*, (*b*) *A. cardenasii* nom nud or *A. duranensis* nom nud, (*c*) SP (*d*) TG-1 (*e*) molecular weight markers. The gel was a gradient of 10 to 20%

*Apd* genes. *A. monticola,* which is an allotetraploid, is believed to be the wild form of *A. hypogaea* (see Smartt et al. 1978; Smartt 1985). The arachin polypeptide pattern in *A. monticola* is similar to the additive pattern of SP and TG-1 (Fig. 6, Table 1). *A. monticola* itself is supposed to have originated by hybridization between two putative diploid parent species followed by doubling of the chromosomes (Smartt et al. 1978). Based on morphological, cytological and phytogeographical evidence one of the most likely genome donors to *A. monticola* is *A. cardenasii* or *A. duranensis* (Smartt et al. 1978). The other putative parent could be *A. batizocoi* (Smartt 1985). The 41.2 kd polypeptide of TG-1 or *A. monticola* is similar to one of the two arachin acidic polypeptides of *A. cardenasii* or *A. duranensis* (Fig. 6, Table 1). This means that *Apd* gene might have come from this diploid gene pool, while *Apc* might be derived from the other unknown gene pool. Screening of wild diploid *Arachis* species for the presence of the *Apc* gene product (42.6 kd polypeptide) might lead to the identification of the other unknown putative parent of *A. monticola,* hence *A. hypogaea.*

*Acknowledgements.* We sincerely thank Mr. S. H. Patil and Mr. Chandramouli (both BARC) for providing the Trombay Groundnut genotypes, Dr. L. J. Reddy and Dr. V. Ramanatha Rao (both ICRISAT, Hyderabad) for ICGS materials and Dr. P. S. Reddy, (NRCG, Junagadh) for the cultivar 'Chalimbana'. We thank Dr. C. R. Bhatia (BARC) for useful comments on the manuscript. Special thanks are due to Mr. D. M. Kale (BARC) for performing the crosses and growing the $F_1$ and $F_2$ materials.

## Preferences

Basha SMM (1979) Identification of cultivar differences in seed polypeptide composition of peanuts (*Arachis hypogaea* L.) by two dimensional polyacrylamide gel electrophoresis. Plant Physiol 63:301–306

Basha SMM, Pancholy SK (1981) Polypeptide composition of arachin and non-arachin proteins from early bunch peanut (*Arachis hypogaea* L.) seed. Peanut Sci 8:82–88

Cherry JP (1977) Potential sources of peanut seed proteins and oil in the genus *Arachis.* Agric Food Chem 25:186–193

Hames BD (1982) In: Hames BD, Rickwood D (eds) Gel electrophoresis of proteins: a practical approach. IRL Press, Oxford and Washington DC, pp 1–91

Johns CO, Jones DB (1916) The globulins arachin and conarachin. J Biol Chem 28:77–87

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

Layne E (1957) In: Colowick SP, Kaplan NO (eds) Methods in enzymology, vol 3. Academic Press, New York, pp 447–454

Lowry OH, Rosebrough NJ, Farr AL, Randall RL (1951) Protein measurement with the Folin-phenol reagent. J Biol Chem 193:265–275

Martin RG, Ames BN (1961) A method for determining the sedimentation behaviour of enzymes: application to protein mixtures. J Biol Chem 236:1372–1379

Mossé J, Pernollet JC (1983) In: Arora SK (ed) Chemistry and biochemistry of legumes. Oxford and IBH Publ, New Delhi, Bombay, Calcutta, pp 111–193

Mouli C, Patil SH, Kale DM (1983) Alterations in subspecific characters of groundnut. In: Induced mutations for improvement of grain legume production, vol III. IAEA-Tecdoc-299, Vienna, pp 205–220

Patil SH (1971) Modified crossing technique for groundnut. Indian J Genet Plant Breed 31:126–127

Savoy CF (1976) Peanut (*Arachis hypogaea* L.) seed protein characterisation and genotype sample classification using polyacrylamide gel electrophoresis. Biochem Biophys Res Commun 68:886–893

Smartt J, Gregory WC, Gregory MP (1978) The genomes of *Arachis hypogaea.* 1. Cytogenetic studies of putative genome donors. Euphytica 26:665–675

Smartt J (1985) Evolution of grain legumes. 5. The oil seeds. Expl Agric 21:305–320

Yamada T, Aibara S, Morita Y (1979) Dissociation-association behaviour of arachin between dimeric and monomeric forms. Agric Biol Chem 43:2549–2556

Yamada T, Aibara S, Morita Y (1980) Accumulation pattern of arachin and its subunit in maturation of groundnut seeds. Plant Cell Physiol 21:1217–1226