Physiol Mol Biol Plants (May–June 2018) 24(3):465–481
https://doi.org/10.1007/s12298-018-0521-8

RESEARCH ARTICLE

Seed protein fraction electrophoresis in peanut (*Arachis hypogaea* L.) accessions and wild species

Apekshita Singh[1] · Soom Nath Raina[1] · Vijay Rani Rajpal[2] · Anurudh K. Singh[3]

Received: 25 May 2017 / Revised: 12 January 2018 / Accepted: 14 February 2018 / Published online: 13 April 2018

Abstract Total seed storage proteins were studied in 50 accessions of *A. hypogaea* (11 *A. hypogaea* ssp. *hypogaea* var *hypogaea*, 13 *A. hypogaea* ssp. *hypogaea* var *hirsuta*, 11 *A. hypogaea* ssp. *fastigiata* var *fastigiata* and 15 *A. hypogaea* ssp. *fastigiata* var. *vulgaris* accessions) in SDS PAGE. These accessions were also analysed for albumin and globulin seed protein fractions. Among the six seed protein markers presently used, it was found that globulin fraction showed maximum diversity (77.2%) in *A. hypogaea* accessions followed by albumin (52.3%), denatured total soluble protein fraction in embryo (33.3%) and cotyledon (28.5%). The cluster analysis based on combined data of cotyledons, embryos, albumins and globulins seed protein fractions demarcated the accessions of two subspecies *hypogaea* and *fastigiata* into two separate clusters supported by 51% bootstrap value, with few exceptions, suggesting the genotypes to be moderately diverse. Native and denatured total soluble seed storage proteins were also electrophoretically analysed in 27 wild *Arachis* species belonging to six sections of the genus. Cluster analysis using different methods were performed for different seed proteins data alone and also in combination. Section Caulorrhizae (C genome) and Triseminatae (T genome) formed one, distantly related group to *A. hypogaea* and other section Arachis species in the dendrogram based on denatured seed storage proteins data. The present analysis has maintained that the section Arachis species belong to primary and secondary genepools and, sections Procumbenetes and Erectoides belong to tertiary gene pools.

Keywords Seed storage proteins · Albumin · Globulin · Intraspecific · Interspecific · Phylogeny

✉ Soom Nath Raina
soomr@yahoo.com

[1] Amity Institute of Biotechnology, Amity University Uttar Pradesh, Sector-125, Noida, Uttar Pradesh 201313, India

[2] Department of Botany, Hans Raj College, University of Delhi, Delhi 110007, India

[3] Gurgaon, India

Introduction

The genus *Arachis* contains around 80 annual and perennial species. Based on morphological similarities, cross-compatibility and viability of the hybrids, the genus has been divided into nine taxonomic sections (Krapovickas and Gregory 1994; Valls and Simpson 2005). The section Arachis is constituted by largest number of species with either $2n = 2x = 20$ or $2n = 4x = 40$. Cultivated *A. hypogaea* ($2n = 4x = 40$) belongs to section Arachis. *A. hypogaea* is further classified into two subspecies, subsp. *fastigiata* Waldron and subsp. *hypogaea* Krap. et Rig. with four botanical varieties, var. *vulgaris*, var. *fastigiata*, var. *peruviana*, and var. *aequatoriana*, and two varieties, var. *hypogaea* and var. *hirsuta*, respectively.

Based on chromosome morphology and cross-compatibility data, the diploid ($2n = 2x = 20$) species of section Arachis have been assigned A/B/D genomes (Smartt et al. 1978b; Gregory and Gregory 1979; Singh and Moss 1982, 1984; Singh 1986; Stalker 1991; Stalker et al. 1990; Fernandez and Krapovickas 1994; Peñaloza and Valls 2005). Two more genomes (F, K) have been reported in species in section Arachis based on physical location of rDNA loci and that of heterochromatin (Robledo and Seijo 2010). Despite the abundance of published information on multiple disciplinary approaches (Singh et al.

1991, 1994, 1996; Raina and Mukai 1999; Kochert et al. 1991; Paik-Ro et al. 1992; Raina et al. 2001; Bechara et al. 2010; Koppolu et al. 2010; He et al. 2014; Cunha et al. 2008) the level of speciation and evolutionary relationships between taxa within section Arachis and that of conclusive identification of two diploid wild progenitors of allotetraploid *A. hypogaea* is as elusive and controversial as ever.

Seed proteins are product of multigene families and are quite stable, uniform, reproducible and seldom affected by environmental fluctuations and management practices (Dunhill and Fowden 1965; Adriaanse et al. 1969; Gray et al. 1973; Basha 1979; Javaid et al. 2004; Masoumi et al. 2012; de Lumen 1990). They have sometimes provided better resolution of taxonomic and phylogenetic problems (Ladizinsky and Hymowitz 1979; Sathaiah and Reddy 1985; Panda et al. 1986; Panigrahi et al. 2007; Bianchi-Hall et al. 1993; Lanham et al. 1994; Singh et al. 1991, 1994; Cherry 1975; Savoy 1976; Basha 1979; Klozova et al. 1983a, b; Bertozo and Valls 2001; Liang et al. 2006; Calbrix et al. 2012; Malik et al. 2009; Tripathy et al. 2016; Emre 2009; Song et al. 2016) in various plant taxa including *Arachis*. The present study, based on the data on total seed storage proteins and that of two fractions, including albumins and globulins, provides new and useful information regarding the evolution, diversification and genomic constitution of the taxa belonging to six of the nine sections of genus *Arachis*.

Materials and methods

Plant material

The list of 78 accessions belonging to 30 species, subspecies and varieties investigated is given in Tables 1 and 2. The seed samples were obtained from International Crop Research Institute for Semi-arid Tropics (ICRISAT), Hyderabad, India and voucher specimens are available from the germplasm resources Unit of ICRISAT.

Protein Extraction

Total soluble proteins in cotyledons

Seeds of each accession were decorticated and the cotyledon and embryo components were separated out. The cotyledon seed components were ground in mortar and pestle. About 5 mg fine powder was used for total soluble protein extraction in 500 µl of 0.1 M Tris–HCl (pH 7.5) buffer. The sample was incubated at 4 °C for 5 h. After the lapse of incubation period, the sample was centrifuged at 4000 rpm for 10 min at room temperature (RT) in a table top Beckman centrifuge. The clear supernatant

Table 1 The *Arachis hypogaea* (2n = 4x = 40, AABB) subspecies and varieties used in the present study

S. no.	Species/subsp.	Variety	Accession no.	Origin
1.	*A. hypogaea* L. ssp. *hypogaea* Krap. & Rig.	*hypogaea*	ICG102	Sudan
2.			ICG495	India
3.			ICG543	Nigeria
4.			ICG623	Senegal
5.			ICG707	USA
6.			ICG745	India
7.			ICG1576	Australia
8.			ICG3799	Zaire
9.			ICG3814	Nigeria
10.			ICG3879	Tanzania
11.			ICG4232	India
12.	*A. hypogaea* L. ssp. *hypogaea* Krap. & Rig.	*hirsuta*	ICG826	India
13.			ICG836	India
14.			ICG844	Unknown
15.			ICG875	India
16.			ICG932	Sudan
17.			ICG934	Uganda
18.			ICG942	India
19.			ICG963	India
20.			ICG965	India
21.			ICG970	India
22.			ICG1003	India
23.			ICG1025	India
24.			ICG1053	India
25.	*A. hypogaea* ssp. *fastigiata* Waldron	*vulgaris*	ICG164	India
26.			ICG221	India
27.			ICG434	USA
28.			ICG1104	India
29.			ICG1151	Unknown
30.			ICG1249	India
31.			ICG1319	India
32.			ICG1362	India
33.			ICG1366	India
34.			ICG1367	India
35.			ICG1369	Zaire
36.			ICG1374	Zaire
37.			ICG2960	India
38.			ICG7827	India
39.			ICG13941	India
40.	*A. hypogaea* L. ssp. *fastigiata* Waldron	*fastigiata*	ICG469	Brazil
41.			ICG1410	Argentina
42.			ICG1588	India
43.			ICG1613	USA
44.			ICG1672	USA

Table 1 continued

S. no.	Species/subsp.	Variety	Accession no.	Origin
45.			ICG1678	Argentina
46.			ICG1700	Brazil
47.			ICG3309	S. Africa
48.			ICG3310	Belgium
49.			ICG3580	India
50.			ICG4890	Argentina

containing the total soluble proteins was taken and kept in − 20 °C till further use. Similar procedure was followed for total soluble protein extraction from cotyledon seed components of each accession.

Total soluble proteins in embryos

The embryo seed components were ground separately in mortar and pestle. A similar procedure as above, was followed for total soluble protein extraction from embryo seed components of each accession.

Albumins and globulins protein fraction

The decorticated seeds were used for extraction of albumin and globulin proteins separately. Their extraction was largely based on the method of Przybylska et al. (1977) with slight modifications made by Schroeder (1982). The cotyledons of each accession were separately ground in pestle and mortar. The fine powder was defatted with petroleum ether (1:9 w/v). About 1 g of the defatted meal was stirred for 1 h in 10 ml of 0.2 M NaCl adjusted to pH 7.0 by 50 mM phosphate buffer. The samples were centrifuged at 10,000*g* for 30 min. The supernatant was collected and the residual meal was again used for protein extraction. Pooled supernatant was filtered through two layers of miracloth pre-wetted with extraction buffer. The samples were then precipitated with solid ammonium sulphate at the rate of 0.7 g/ml and centrifuged at 10,000*g* for 30 min. The pellet was dissolved in an appropriate volume of 0.2 M NaCl (pH 7.0) and dialysed for 72 h at 4 °C against distilled water, using 4–5 changes during 24 h. After centrifuging at 12,000*g* for 30 min at 4 °C, the supernatant contained albumin proteins whereas the pellet contained the globulins which was resuspended in 0.5 M NaCl (pH 7.0) after a wash with distilled water. The albumin proteins were lyophilized.

The protein extracts of soluble proteins, albumins and globulins were estimated for quantity (µg/ml) with Bradford Reagent (Bradford 1976), and stored at 4 °C till further use.

Native PAGE

One dimensional discontinuous polyacrylamide gel electrophoresis (PAGE) was carried out in native system according to Laemmli (1970). A 12.5% separating gel (0.375 M Tris–HCl, pH 8.8) and 4% stacking gel (0.125 M Tris–HCl, pH 6.8) was cast. Protein samples extracted from seed cotyledons, before loading were treated with anchor solution [0.5 M Tris–HCl (pH 6.8), 10% w/v glycerol, 1% bromophenol blue] in 1:1 ratio of anchor solution and protein sample. The electrode buffer was Tris–glycine, pH 8.3 (0.025 M Tris, 0.192 M glycine). The gels were electrophoresed initially at 20 mA and subsequently at 35 mA. The current was stopped when the tracking dye was approximately 1.5 cm above the lower end. The staining of the gel was carried out overnight in a solution containing 0.2% Coomassie Brilliant Blue R 250, 40% methanol and 10% acetic acid. The destaining was done in a solution of 40% methanol and 10% acetic acid, without the dye.

SDS PAGE

Cotyledon seed protein fraction

One dimensional discontinuous Sodium dodecyl sulphate (SDS) PAGE was carried out in the same way as for native PAGE. In the case of SDS PAGE the constituents and the buffer solutions remained the same, except that 10% SDS was added in separating gel, stacking gel, running buffer and anchor solution. The protein samples were heated for 10 min before loading onto the gel.

Equal quantity of protein (50 µg) for each analysed accession was mixed (1:1) with loading buffer and loaded in full in each lane.

Protein molecular weights, in case of SDS gel was estimated by using high molecular weight kit (GIBCO BRL). The standard marker gave seven bands in the order of myosin 200 kd, phosphorylase B 97.4 kd, bovine serum albumin 68 kd, ovalbumin 43 kd, carbonic anhydrase 29 kd, β-lactoglobulin 18.4 kd and lysozyme 14.3 kd.

Embryo seed protein fraction

The SDS PAGE of embryo protein samples was carried out in the same way as for cotyledon protein fractions.

Albumin protein fractions

For albumins, 1 mg protein in each analysed accession was dissolved in 100 µl of sterile distilled water, of which 50 µl

in 50 µl of loading buffer was used for loading. The SDS PAGE of albumin proteins was carried out in the same way as for cotyledon protein fractions.

Globulin protein fractions

For each analysed accession, 20 µg of protein was mixed with loading buffer (1:1) and loaded in each lane. The SDS PAGE of globulin proteins was carried out in the same way as for cotyledon protein fractions.

Scoring and Data Analysis

Based on molecular weights and Rf values of protein or protein fraction bands, Quantity One Software (Bio-Rad Laboratories) was used for data analysis. The presence (1) and absence (0) of the bands, in both cases, was scored as binary data matrix. Similarity/dissimilarity matrices and subsequent clusterings were done using Numerical Taxonomy and Multivariate Analysis system (NTSYS) ver 2.02 k package (Rohlf 1993). Similarity matrices with Jaccard and Dice coefficients were made, using program SimQual of NTSYS. The resultant similarity matrix was used to make dendrograms using Unweighted Pair Group Method with Arithmatic Mean (UPGMA) (Sneath and Sokal 1973) and Neighbour-Joining (NJ) (Saitou and Nei 1987). Consensus tree was made by CONSENSUS option in NTSYS.

The distance matrices were also made with Dice dissimilarity coefficient and phylogenetic trees were made by Neighbor Joining method using DARwin (Dissimilarity Analysis and Representation for Windows) software (Perrier and Jacquemoud-Collet 2006). Consensus tree was made using Consensus and Tree distances option.

Results

A total of 50 accessions belonging to *A. hypogaea* ssp. *hypogaea* var. *hypogaea, A. hypogaea* ssp. *hypogaea* var. *hirsuta, A. hypogaea* ssp. *fastigiata* var. *fastigiata* and *A. hypogaea* ssp. *fastigiata* var. *vulgaris* were investigated for SDS PAGE of protein fractions isolated separately from cotyledons and embryos. The albumins and globulins protein fractions were also analysed.

Variation within *Arachis hypogaea*

SDS PAGE protein profile

Cotyledon seed protein fractions In the SDS PAGE, the number of bands per lane in cotyledon proteins in 50 accessions of *A. hypogaea* consisting of *A. hypogaea* ssp. *hypogaea* var. *hypogaea* (11 accessions), *A. hypogaea* ssp. *hypogaea* var. *hirsuta* (13 accessions), *A. hypogaea* ssp. *fastigiata* var. *fastigiata* (11 accessions)and *A. hypogaea* ssp. *fastigiata* var. *vulgaris* (15 accessions), ranged from 18 to 21 with molecular weights 95.02–12.22 kd (Fig. 1; Table 3). The band of 48.03 kd was highly polymorphic. It was absent in one, seven and eight accessions of *A. hypogaea* ssp. *hypogaea* var *hirsuta* (No. 12 as per Table 1), *A. hypogaea* ssp. *fastigiata* var. *vulgaris*, *A. hypogaea* ssp. *fastigiata* var *fastigiata,* respectively. Band of 28.10 kd was also polymorphic. It was absent in five, one, six and three accessions of *A. hypogaea* ssp. *hypogaea* var *hypogaea, A. hypogaea* ssp. *hypogaea* var *hirsuta*, *A. hypogaea* ssp. *fastigiata* var *vulgaris* and *A. hypogaea* ssp. *fastigiata* var *fastigiata*, respectively. Band of 14.15 kd was absent in single accession of *A. hypogaea* ssp. *hypogaea* var *hirsuta* (No. 20 as per Table 1) and *A. hypogaea* ssp. *fastigiata* var *fastigiata* (No. 46 as per Table 1).

Embryo seed protein fractions In embryo seed protein fractions, the number of bands varied from 17 to 22 with molecular weights ranging from 95.02 to 12.22 kd (Fig. 1;

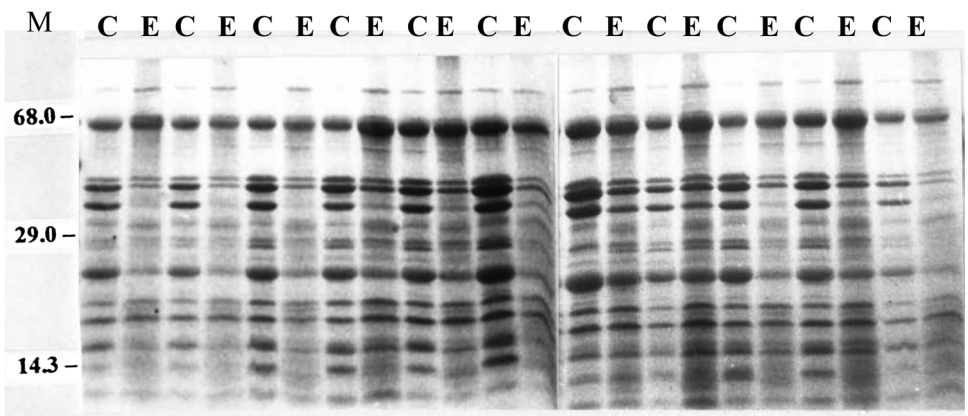

Fig. 1 Gel electrophoresis in cotyledon and embryo seed protein fractions in SDS PAGE, where *C* cotyledon, *E* embryo

Table 2 The *Arachis* species used in the present study

Species	Section	Accession no.	Series	Origin	2n	Genome
A. monticola Krap. et Rig.	Arachis	ICG8197	*Amphiploides*	Argentina	40	AABB
A. monticola Krap. et Rig.	Arachis	ICG8135	*Amphiploides*	Argentina	40	AABB
A. monticola Krap. et Rig.	Arachis	ICG811549	*Amphiploides*		40	AABB
A. batizocoi Krap. et Rig.	Arachis	ICG8210	*Annuae*	Bolivia	20	BB
A. benensis Krap. et Greg. et Simpson	Arachis	ICG11551		Bolivia	20	BB
A. cardensaii Krap. et Greg.	Arachis	ICG11564	*Perennes*		20	AA
A. correntina Krap. et Greg.	Arachis	ICG8132	*Perennes*	Argentina	20	AA
A. diogoi Krap. et Greg.	Arachis	ICG4983	*Perennes*	Paraguay	20	AA
A. duranensis Krap. et Greg.	Arachis	ICG8199	*Annuae*	Argentina	20	AA
A. hoehneii Krap. et Greg.	Arachis	ICG8190			20	AA
A. helodes Krap. et Greg.	Arachis	ICG8952	*Perennes*	Brazil	20	AA
A. ipaensis Greg. et Greg.	Arachis	ICG8206	*Annuae*	Bolivia	20	BB
A. kemepff Krap. et Greg.	Arachis	ICG8164			20	AA
A. magna Krap. et Greg.	Arachis	ICG8960			20	BB
A. stenosperma Greg. et Greg.	Arachis	ICG8137	*Perennes*		20	AA
A. valida Krap. et Greg.	Arachis	ICG8193	*Annuae*		20	BB
A. villosa Benth.	Arachis	ICG13258	*Perennes*	Unknown	20	AA
A. appressipila Krap. et Greg.	Procumbentes	ICG8945	*Procumbensae*	Brazil	20	EE
A. chiquitana Krap.	Procumbentes	ICG11560	*Procumbensae*	Bolivia	20	EE
A. kretschmeri Krap. et Greg.	Procumbentes	ICG13235	*Procumbensae*	Brazil	20	EE
A. matiensis Krap.	Procumbentes	ICG11557	*Procumbensae*	Bolivia	20	EE
A. rigonii Krap. et Greg.	Procumbentes	ICG8904	*Procumbensae*	Bolivia	20	EE
A. stenophylla Krap. et Greg.	Erectoides	ICG8215		Paraguay	20	E_2E_2
A. paraguarensis Chodat & Horsd.	Erectoides	ICG8973		Brazil	20	E_2E_2
A. sylvestris A. Chev.	Heteranthae	ICG13167		Brazil	20	Am
A. pusilla Benth.	Heteranthae	ICG13213		Brazil	20	Am
A. triseminata Krap. et Rig	Triseminatae	ICG8131		Brazil	20	TT
A. pintoi Krap. et Greg.	Caulorrhizae	ICG13222		Brazil	20	CC

Table 3). The percentage polymorphism in cotyledon and embryo profiles was 28.5 and 33.3%, respectively. The most polymorphic protein marker was band of 28.10 kd.

Albumins proteins In SDS PAGE of albumin seed protein fractions, 15–20 bands in the range of 108.8–11.26 kd were scored (Fig. 2; Table 3). In accession number 1003 of *A. hypogaea* ssp. *hypogaea* var. *hirsuta*, five bands of 89.19, 84.02, 68.18, 57.37 and 53.98 kd were absent. The percentage polymorphism in albumin protein fraction was 52.3%.

Globulins proteins By comparison, 22 bands of 96.02–14.42 kd in globulin fraction were scored of which 17 (77.2%) were polymorphic (Fig. 3; Table 3).

Based on SDS PAGE data of cotyledon, embryo, albumin and globulin seed protein fractions, four dendrograms were generated using UPGMA clustering method. One dendrogram based on combined data was also made. Very little difference in topologies was observed. Here dendrogram based on the combined data is described (Fig. 4). Two well defined clusters were identified. Cluster I and Cluster II joined each other with 51% bootstrap value. Cluster I consisted of 28 accessions. In cluster I, all the accessions belonging to *A. hypogaea* ssp. *hypogaea*, except one (No. 22 in Table 1) are grouped together. Of these, eleven and ten accessions were of *A. hypogaea* ssp. *hypogaea* var. *hypogaea* (Nos. 1–11 in Table 1) and *A. hypogaea* ssp. *hypogaea* var. *hirsuta* (Nos. 12–21 in Table 1), respectively. In Cluster I, a subcluster consisted of seven accessions of *A. hypogaea* ssp. *fastigiata* var. *vulgaris* (Nos. 33–39 in Table 1). In Cluster II, 22 accessions were grouped. It consisted of all the accessions from *A. hypogaea* ssp. *fastigiata* except the two from *A. hypogaea* ssp. *hypogaea* var. *hirsuta* (Nos. 23–24 in Table 1). In the cluster II, a subcluster grouped all the accessions of *A. hypogaea* ssp. *fastigiata* var. *fastigiata* together except one of *A. hypogaea* ssp. *fastigiata* var. *fastigiata* (No. 49). In

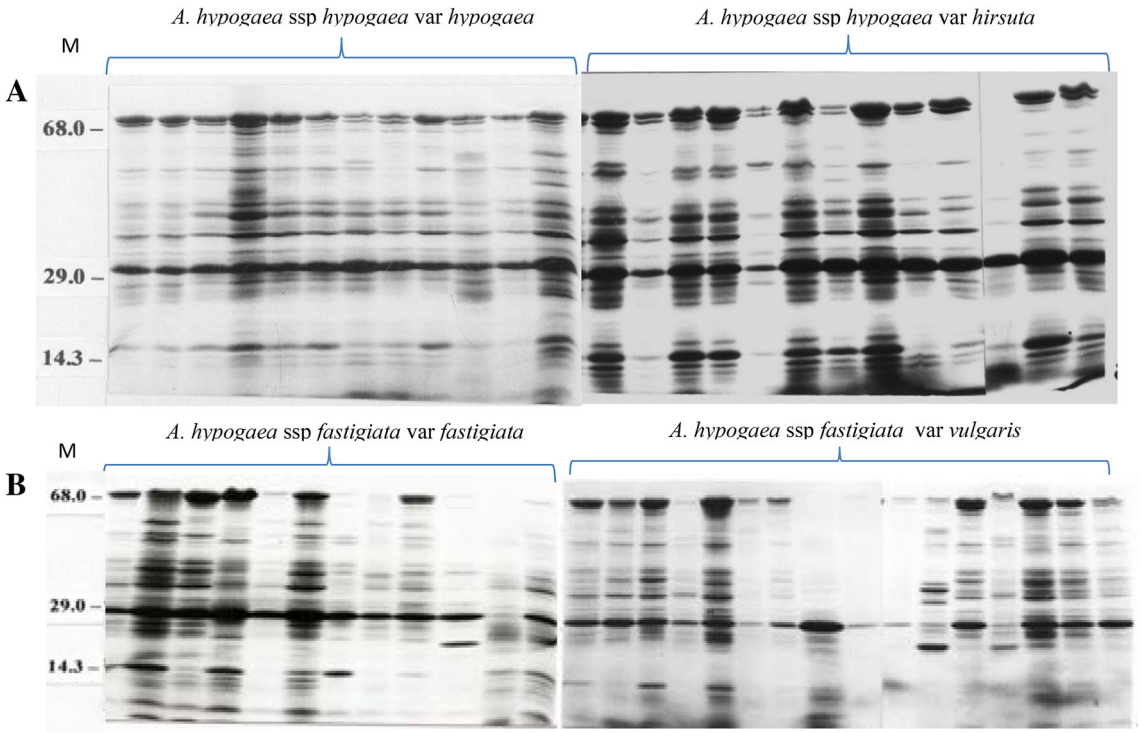

Fig. 2 Gel electrophoresis patterns in Albumin proteins in *A. hypogaea* ssp. *hypogaea* (**a**), *A. hypogaea* ssp. *fastigiata* (**b**)

Table 3 Variation in banding pattern of *Arachis hypogaea* accessions in SDS PAGE

Seed protein fractions	Accession number	Number of bands	Mol wt range (kd)	Polymorphic bands	Percentage polymorphism	Monomorphic bands
Cotyledons	50	21	95.02–12.22	48.03, 35.06, 29.28, 28.10, 21.53, 14.15	28.5	95.02, 68.16, 55.88, 52.55, 40.17, 38.32, 32.37, 30.82, 29.28, 24.50, 20.65, 18.76
Embryos	50	21	95.02–12.22	35.42, 29.28, 28.10, 21.53, 15.92, 14.15, 12.22	33.3	95.02, 68.16, 55.88, 48.03, 40.17, 31.68, 30.82, 24.50, 20.65, 18.76, 16.71
Albumin	50	21	108.8–11.26	108.8, 89.19, 84.02, 68.18, 57.37, 53.98, 34.22, 26.27, 16.96	52.3	96.58, 41.98, 39.63, 37.58, 35.69, 31.68, 30.34, 25.18, 19.53, 11.26
Globulins	50	22	96.08–14.42	69.15, 63.65, 52.65, 41.76, 34.39, 33.06, 28.0, 27.37, 23.31, 21.29, 15.40, 14.42	77.2	96.08, 40.34, 30.47, 24.82, 90.14

another subcluster, remaining seven *A. hypogaea* ssp. *fastigiata* var. *vulgaris* accessions grouped together (Nos. 26–32).

Variation between species

Native PAGE protein profile

In the 15 wild species and *A. hypogaea* (Acc. No. ICG 495) belonging to Section Arachis, the number of bands ranged from 6 to 11 (Fig. 5; Table 4). The least number (6) of bands were detected in *A. villosa* and *A. stenosperma*. There were eight bands each in *A. benensis, A. hoehneii, A. ipaensis* and *A. kempff*. In *A. monticola, A. batizocoi, A. correntina, A. diogoi, A. helodes,* nine bands were scored. *A. cardenasii* and *A. valida,* and *A. duranensis* and *A. hypogaea* were characterized by 10 and 11 bands, respectively. In *A. magna* seven bands were scored. The percentage polymorphism among the section Arachis species was found to be 82.3%.

Both the species *A. pusilla* and *A. sylvestris* in Section Heteranthae had eight bands. However, only four

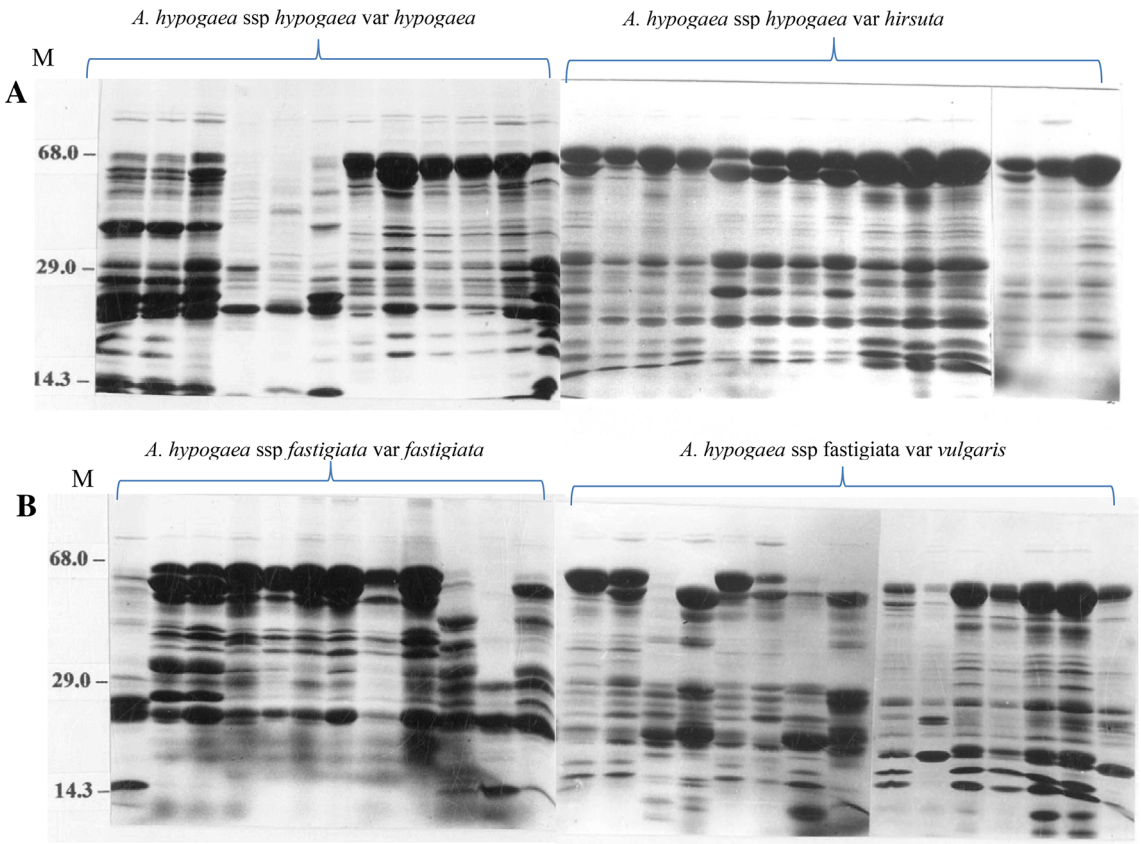

Fig. 3 Gel electrophoresis patterns of Globulin proteins in *A. hypogaea* ssp. *hypogaea* (**a**), *A. hypogaea* ssp. *fastigiata* (**b**)

bands of Rf values 0.02, 0.08, 0.14, 0.25 were common between the two species. It was found that the percentage polymorphism in section Heteranthae was 66.6%. The two species *A. paraguarensis* and *A. stenophylla* of Section Erectoides had nine and eight bands, respectively. *A. paraguarensis* was distinguished by the presence of bands of Rf values 0.05, 0.17, while a band of Rf value 0.35 was present in *A. stenophylla* alone. The percentage polymorphism in section Erectoides was 30%. The number of bands in the five species of Section Procumbentes ranged from seven in *A. chiquitana* and *A. kretschmeri* to 10 in *A. appressipila*. The bands of Rf values 0.07 and 0.35 were only present in *A. appressipila.* The percentage polymorphism in section Procumbentes was 30%. The only representative species *A. triseminata* of Section Triseminatae had eight bands of Rf values 0.02–0.55. Single species *A. pintoi* belonging to Section Caulorrhizae had 10 bands of Rf values of 0.02–0.60. The band of Rf value 0.49 was unique to the species.

The band with Rf value 0.02 was common to all the 16 species in six taxonomic sections. This was followed by a band of Rf value 0.08 in four sections.

SDS PAGE protein profile

The number of bands in 16 species in Section Arachis ranged from 16 to 21 (Fig. 6; Table 5). The maximum 21 bands were scored in *A. hypogaea* and *A. helodes* followed by 20, 18, 17 and 16 in *A. monticola; A. cardenasii, A. benensis*, *A. ipaensis*; *A. correntina, A. batizocoi, A. diogoi, A. duranensis, A. hoehneii*, A. *kempff, A. magma* and *A. valida* and *A. stenosperma* and *A. villosa,* respectively. The band of 61.98 kd was absent in all but *A. correntina* and *A. duranensis*. Band of 12.99 kd was unique to *A. hypogaea*. The band of 19.57 kd was absent in *A. ipaensis* alone. Barring *A. batizocoi,* all other species possessed 25.8 kd band. Similarly, 27.31 kd band was present in all but *A. cardensaii*. The percentage polymorphism in section Arachis was 64%. Both the species of section Heteranthae, *A. pusilla* and *A. sylvestris* had 18 and 21 bands, respectively. Four bands (28.05, 17.35, 14.06 and 12.99 kd) were present in *A. sylvestris* but were absent in *A. pusilla,* whereas the 36.81 kd band was present in only *A. pusilla*. The percentage polymorphism in section Heteranthae was 22.7%. In species of section Erectoides, a total of 16 and 18 bands were present in *A. paraguarensis* and *A. stenophylla,* respectively. Apart from the two bands, 31.79 and 19.57 kd, present in *A. stenophylla,* the protein profiles were

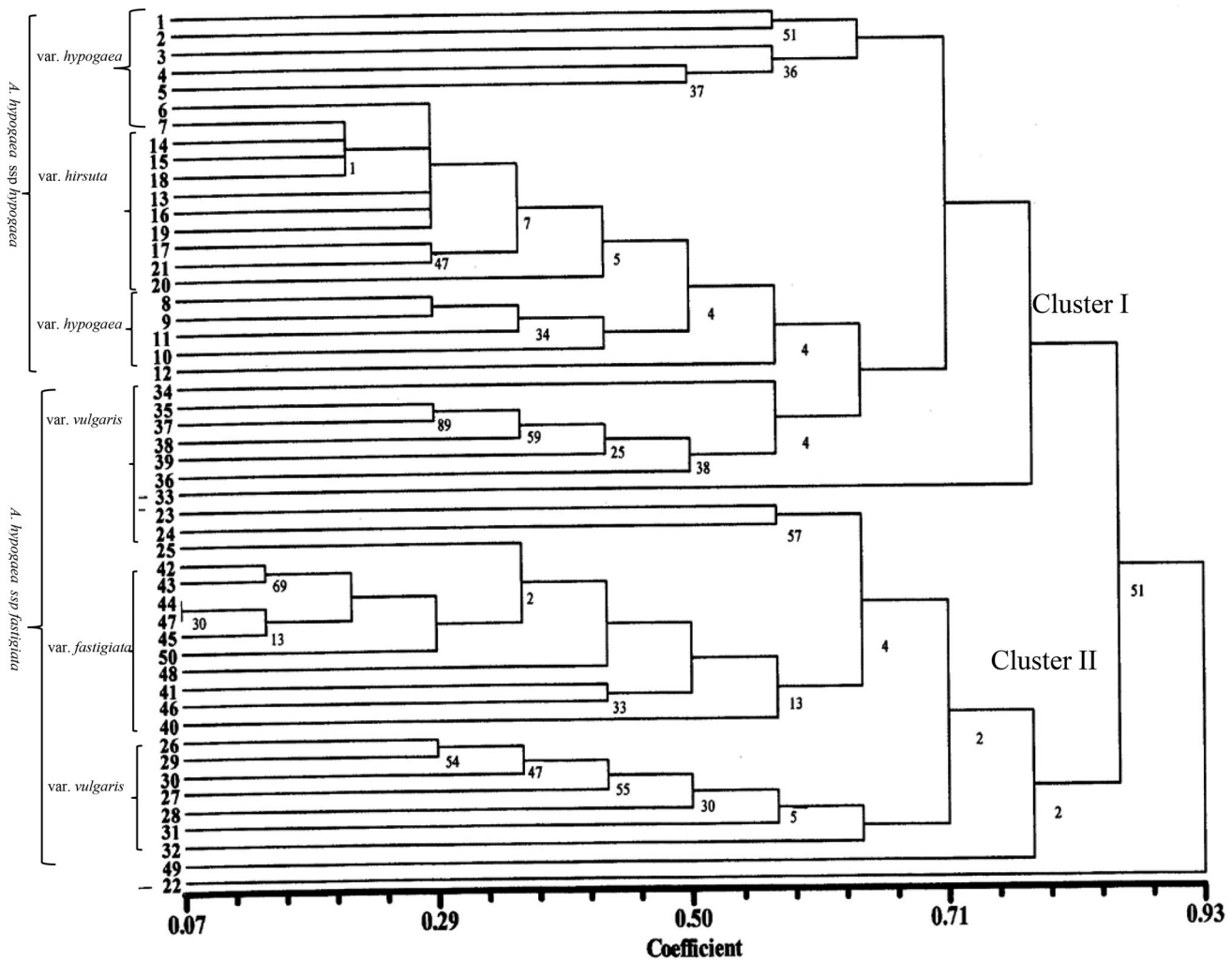

Fig. 4 Dendrogram of *A. hypogaea* accessions based on combined matrices of cotyledon, embryo, albumin and globulin proteins

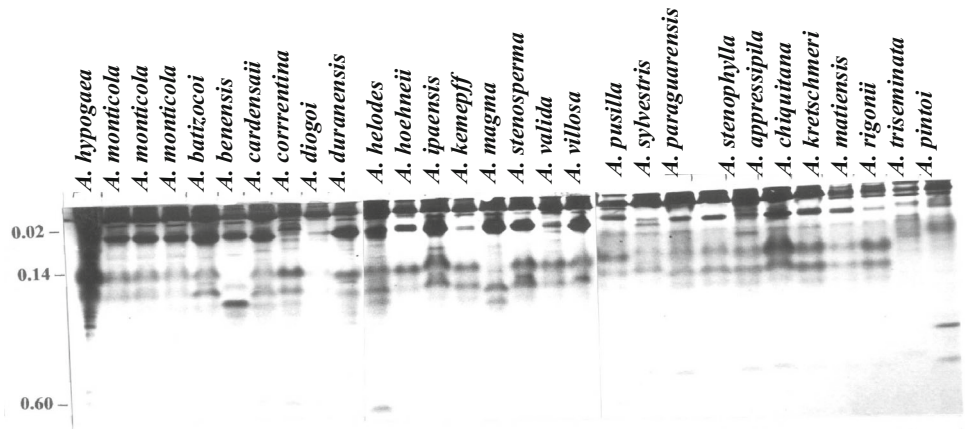

Fig. 5 Gel electrophoresis patterns of Arachis species in Native PAGE

similar in both the species. The percentage polymorphism in section Erectoides was 11.1%. The number of protein bands in the five species of section Procumbentes ranged from 18 in *A. kretschmeri* to 21 in *A. chiquitana.* The bands of 17.35 and 16.88 kd, and that of 32.57 kd were present in *A. chiquitana* and *A. appressipila*, respectively. *A. kretschmeri* was devoid of the 45.23 kd protein fraction band. The percentage polymorphism in section Procumbentes was 33.3%. Sections Triseminatae and Caulorrhizae were represented by single species of *A. triseminata* and *A.*

Table 4 Variation in banding pattern of *Arachis* species in Native PAGE

Section	Species	Rf value	No. of bands	Polymorphic bands	Percentage polymorphism	Monomorphic bands
Arachis	16	0.02–0.60	17	0.07, 0.10, 0.12, 0.14, 0.17, 0.19, 0.20, 0.22, 0.25, 0.29	82.3	0.02, 0.05, 0.08
Procumbentes	5	0.02–0.60	10	0.05, 0.07, 0.35	30	0.02, 0.08, 0.14, 0.19, 0.25, 0.31, 0.60
Erectoides	2	0.02–0.60	10	0.05, 0.17, 0.35	30	0.02, 0.08, 0.14, 0.19, 0.25, 0.31, 0.60
Heteranthae	2	0.02–0.60	12	0.05, 0.10, 0.17, 0.19, 0.20, 0.29, 0.31, 0.60	66.6	0.02, 0.08, 0.14, 0.25
Triseminatae	1	0.02–0.60	8	0.07, 0.12, 0.18, 0.55		0.02, 0.05
Caulorrhizae	1	0.02–0.60	10	0.18, 0.17, 0.25, 0.49		0.02, 0.05

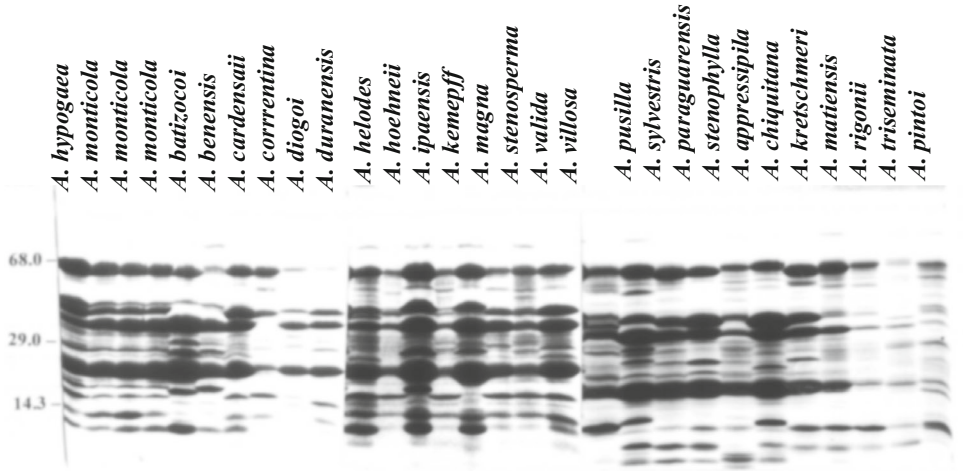

Fig. 6 Gel electrophoresis patterns of Arachis species in SDS PAGE

pintoi, respectively. Nineteen bands were found to occur in *A. triseminata*. It was marked by the absence of bands of 42.84, 45.23, 41.89, 32.57, 17.35, 12.99 and 10.13 kd, which are otherwise present in most other *Arachis* taxa. *A. pintoi* showed unique profile of 21 bands, of which 61.98, 42.84, 41.89, and 17.35 kd were mostly absent in other taxa.

The band of 61.98 kd present in the two species of section Arachis was absent in the sections Heteranthae, Erectoides, and two species of section Procumbentes. The band of 41.89 kd was absent in sections Arachis, Heteranthae, Erectoides, Triseminatae and in some species of section Procumbentes. The band of 36.81 kd was absent in sections Erectoides, Procumbentes and Caulorhizae. The band of 32.57 kd present in *A. appressipila,* was absent in sections Erectoides, Triseminatae and Caulorhizae, and in most species of sections Arachis and Procumbentes. Band of 18.08 kd was absent in section Heteranthae but was present in sections Erectoides, Procumbentes, Triseminatae and Caulorhizae.

Cluster analysis

The cluster analysis of 27 *Arachis* species based on neighbor joining tree with dice coefficient (Fig. 7) using native seed storage proteins data shows two clusters. Cluster I recognises two subclusters. In one subcluster, section Arachis species *A. hypogaea* is grouped with *A. monticola*, *A. correntina* and section Triseminatae species *A. triseminata*, which is very distantly related. In the same subluster, *A. sylvestris* (section Heteranthae) is also grouped. In another subcluster, section Arachis species *A. duranensis*, *A. cardensaii, A. diogoi, A. benensis, A. batizocoi* form one group. *A. hoehneii* is joined as separate unit. Section Caulorrhizae species *A. pintoi* also forms a separate unit. Cluster II groups seven species in one subcluster *A. rigonii, A. kretschmeri, A. chiquitana, A. appressipila, A. matiensis* (all in section Procumbentes), *A. stenophylla* (Erectoides) and *A. kempff* (section Arachis). In another subluster, six species are grouped together. Section Arachis species *A. valida, A. helodes, A. magna, A. ipaensis* and two species *A.*

Table 5 Variation in banding pattern of *Arachis* species in SDS PAGE

Section	Species	Mol wt range (kd)	No. of bands	Polymorphic bands	Percentage polymorphism	Monomorphic bands
Arachis	16	94.85–10.13	25	61.98, 42.84, 40.24, 41.89, 36.81, 32.57, 27.31, 25.8, 17.53, 12.99	64	94.85, 66.80, 53.97, 50.27, 28.05, 23.20, 14.06, 12.17, 10.13
Procumbentes	5	94.85–10.13	24	61.98, 42.84, 17.53, 36.81, 16.88, 32.57, 45.23, 10.13	33.3	94.85, 66.80, 53.97, 50.27, 40.24, 34.03, 31.79, 28.05, 16.88, 12.17
Erectoides	2	94.85–10.13	18	31.79, 19.57	11.1	94.85, 66.80, 53.97, 40.24, 23.20, 12.17
Heteranthae	2	94.85–10.13	22	36.81, 28.05, 17.53, 14.06, 12.99	22.7	94.85, 66.80, 53.97, 50.27, 40.24, 23.20, 12.17, 10.13
Triseminatae	1	94.85–10.13	19	61.98, 36.81, 16.88		94.85, 66.80, 53.97, 50.27, 23.20, 12.17
Caulorrhizae	1	94.85–10.13	21	42.84, 41.98, 18.08		94.85, 66.80, 53.97, 50.27, 23.20, 12.17

paraguarensis, A. pusilla of sections Erectoides and Heteranthae, respectively. Two section Arachis species *A. villosa* and *A. stenosperma* are joined separately.

The cluster analysis of 27 *Arachis* species based on neighbor joining tree with dice coefficient in SDS seed storage proteins data recognises three clusters (Fig. 8). Cluster I groups *A. villosa, A. stenosperma, A. valida, A. kempff, A. ipaensis, A. hoehneii, A. duranensis, A. correntina, A. diogoi*. In a subcluster three species *A. chiquitana, A. sylvestris, A. batizocoi*are grouped together. In

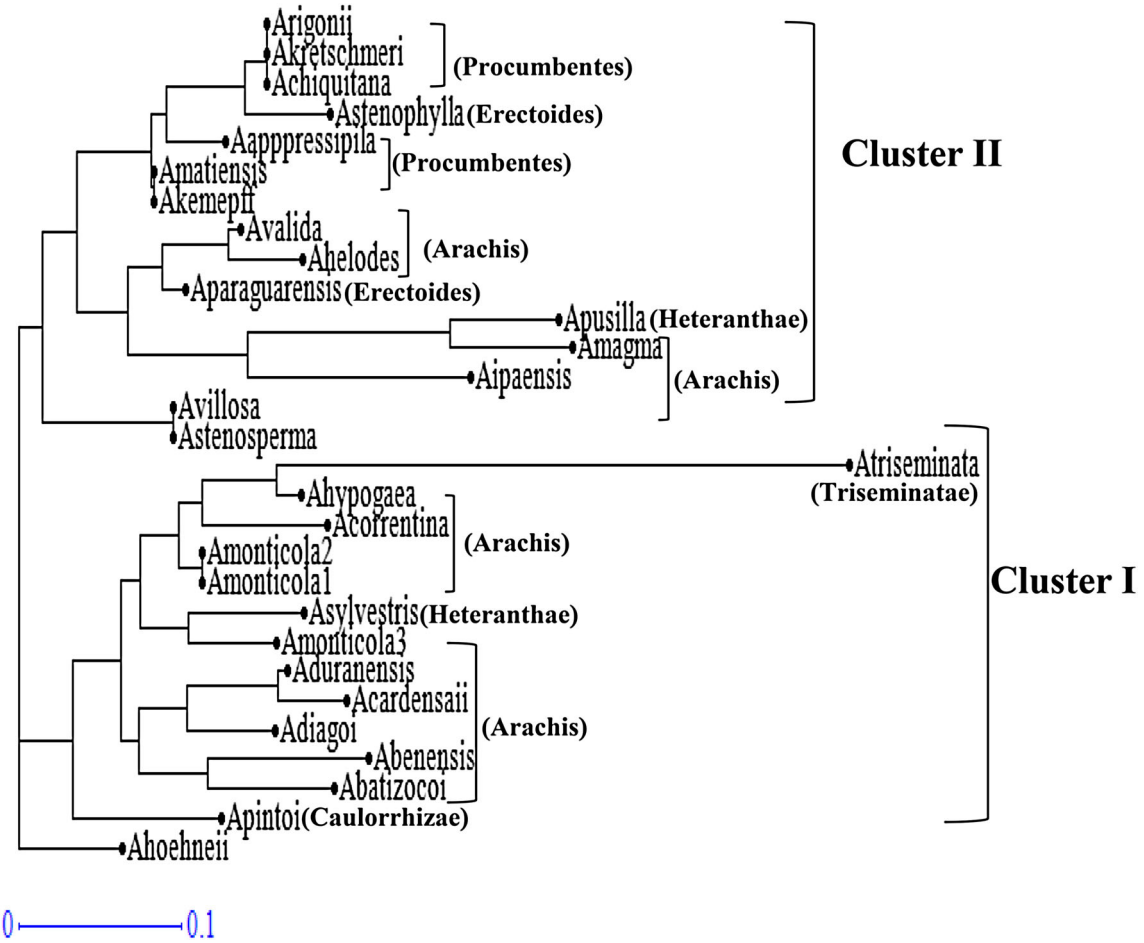

Fig. 7 Neighbor joining tree of Arachis species based on Native PAGE data

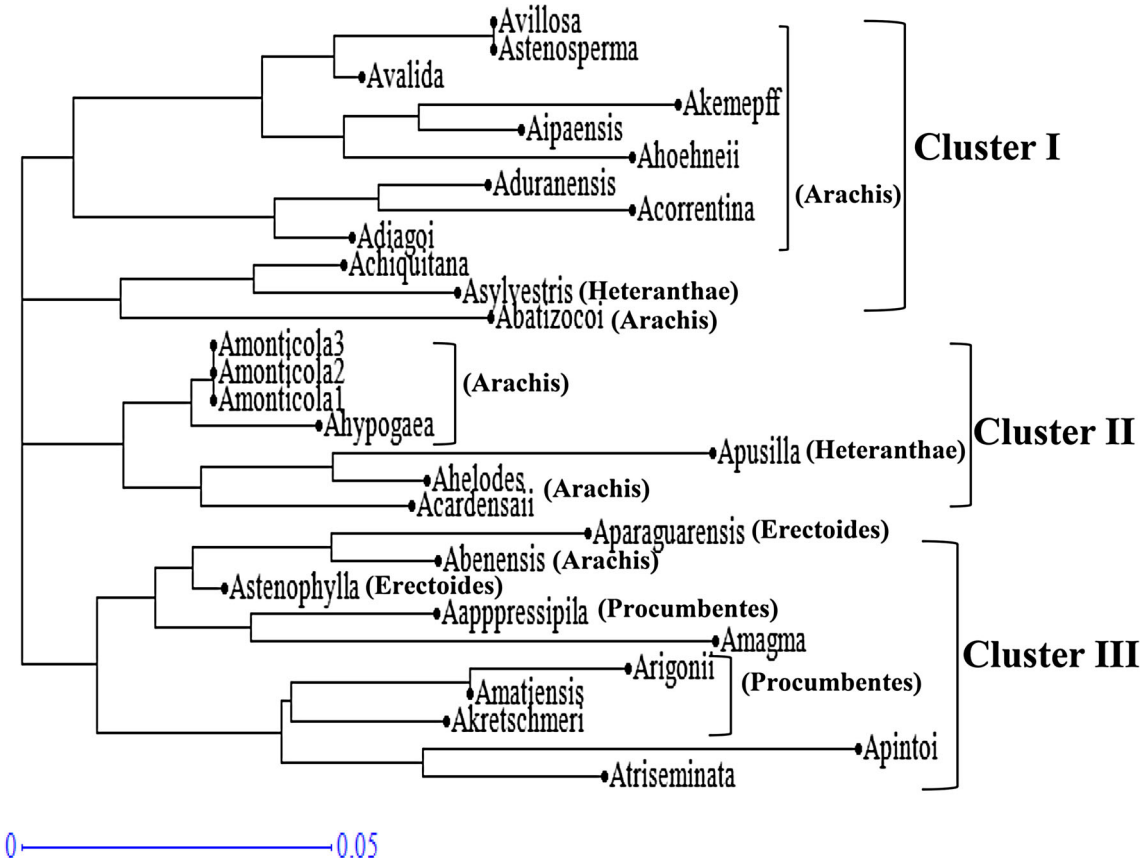

Fig. 8 Neighbor joining tree of Arachis species based on SDS PAGE data

Cluster II, *A. hypogaea* is grouped with *A. monticola, A. pusilla* (section Heteranthae), *A. helodes, A. cardensaii*. In Cluster III there are two small subclusters. In one subcluster, section Erectoides species *A. paraguarensis, A. stenophylla* along with *A. appresipila* (Procumbentes) and two section Arachis species *A. benensis, A. magna* are grouped. In another subcluster, *A. rigonii, A. matiensis, A. kretschmeri* (all in section Procumbentes) are grouped together. The two species *A. triseminata* (section Triseminatae) and *A. pintoi* (section Caulorrhizae) also form part of Cluster III.

Consensus tree (Fig. 9) based on combined data of native and SDS PAGE recognises three main clusters. In Cluster I, mostly species of section Arachis are present (*A. hypogaea, A. monticola, A. duranensis, A. correntina, A. diogoi, A. cardensaii, A. batizocoi* and *A. benensis*). However, it also groups *A. triseminata* (section Triseminatae) and *A. pintoi* (section Caulorrhizae) in the same cluster. Cluster II consists of mainly section Arachis species, *A. ipaensis*, *A. hoehneii, A. villosa, A. stenosperma, A. kempeff*, *A. valida* in one subcluster. In another subcluster section Arachis species, *A. magna* and *A. helodes* together with *A. pusilla* (section Heteranthae) are grouped. In Cluster III, there are two subclusters. In one subcluster, section Procumbentes species *A. rigonii, A matiensis, A kretschmeri, A appressipilla* and two section Erectoides species *A stenophylla, A paraguarensis* are closely grouped. In another small subcluster *A. chiquitana* (section Procumbentes) is joined with *A. sylvestris* (section Heteranthae).

Discussion

In the present study, fifty accessions belonging to two subspecies (*A. hypogaea* ssp. *hypogaea*, *A. hypogaea* ssp. *fastigiata*) of *A. hypogaea* and four varieties (*A. hypogaea* ssp. *hypogaea* var. *hypogaea, A. hypogaea* ssp. *hypogaea* var. *hirsuta, A. hypogaea* ssp. *fastigiata* var. *fastigiata* and *A. hypogaea* ssp. *fastigiata* var. *vulgaris*) were electrophoretically characterised by seed proteins in cotyledon and embryo fractions, as well as for albumin and globulin protein markers. Among the various seed protein markers presently used, it was found that globulin fraction showed maximum diversity (77.2%) in *A. hypogaea* accessions followed by albumin (52.3%), denatured soluble protein fraction in embryo (33.3%) and cotyledon (28.5%). Based on the results of this data, the cluster analysis of total

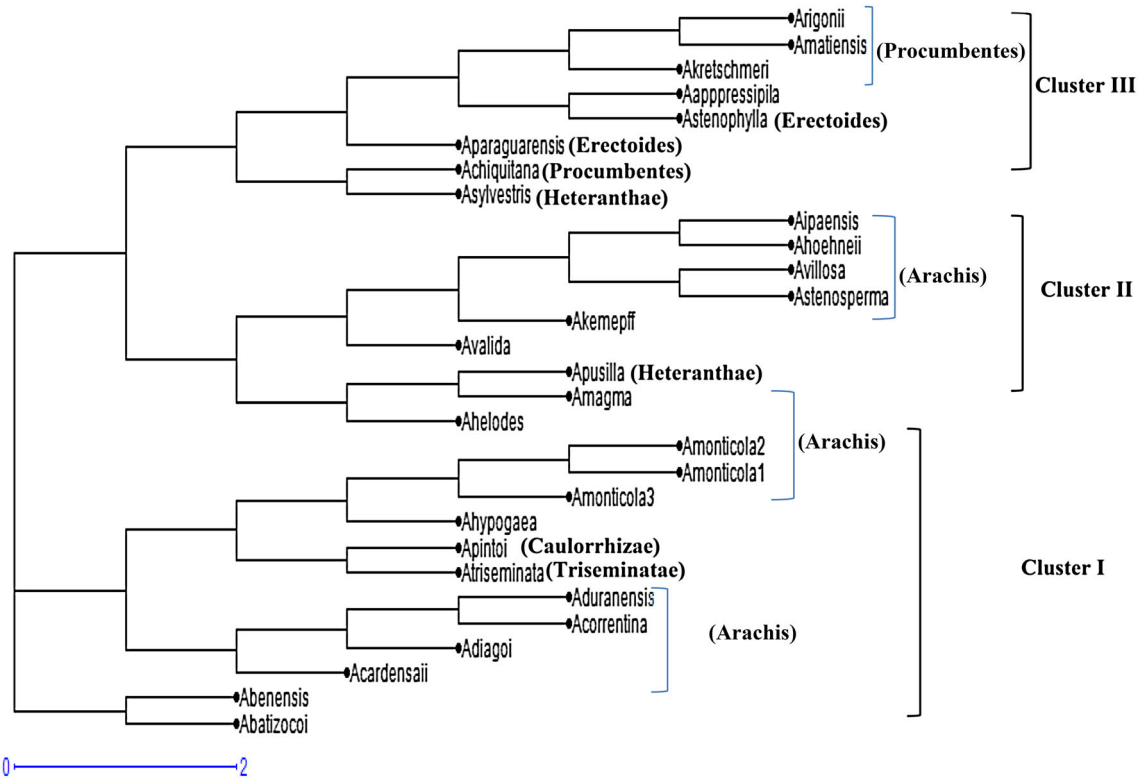

Fig. 9 Consensus tree of Arachis species based on combined Native and SDS PAGE data

soluble proteins analysed in the cotyledon and embryo seed fractions, and albumin and globulin protein fractions, demarcated the two subspecies *A. hypogaea* ssp. *hypogaea* and *A. hypogaea* ssp. *fastigiata* into discrete clusters with bootstrap value confidence of 51% (Fig. 4). Although the grouping of var. *hypogaea*, var *hirsuta*, var *vulgaris* accessions differed to some extent in the analysis, but it was seen that var *fastigiata* accessions grouped together. The ratio of albumin to globulin proteins in the present study was less than one, implying that the amount of globulin proteins per seed sample was much more than the albumin proteins. On the whole, it was observed that polymorphism in globulin protein fraction proved more succesful in demarcating the two subspecies and four varieties. The earlier investigations on seed proteins in *Arachis* (Bianchi-Hall et al. 1993; Lanham et al. 1994; Singh et al. 1991, 1994; Cherry 1975; Savoy 1976; Basha 1979; Klozova et al. 1983a, b; Krishna et al. 1986; Liang et al. 2006) detected low to moderate polymorphism and used a far less number of accessions. Though the number of accessions has been increased in subsequent studies, it is still too low to adequately represent the total variability of world germplasm collections, and to permit drawing of valid conclusions on genetic diversity in the groundnut.

The total seed proteins extracted from 12 different genotypes of cultivated peanut (*A. hypogaea* L.) comprising, *A. hypogaea* ssp. *hypogaea* and *A hypogaea* ssp. *fastigiata* could be separated by electrophoresis on both one-dimensional and two-dimensional PAGE (Liang et al. 2006). Bertozo and Valls (2001) have studied total seed storage protein profiles of 28 accessions of *A. pintoi* and one of *A. repens*, belonging to Caulorrhizae section and compared it by native-PAGE and SDS-PAGE. Seed storage protein profiles between the accessions of the cultivated groundnut found very little variation (Singh et al. 1991, 1994). However in another study, no distinct differences between four peanut cultivars belonging to the two subspecies were found on one dimensional SDS-PAGE, but differences were found in two dimensional electrophoresis of seed proteins (Kottapalli et al. 2008). Similarly, even after analyzing, one hundred and fifty one accessions of groundnut germplasm originated from various sources and differed from the material used by other researchers, low genetic diversity was revealed by SDS PAGE of total soluble proteins (Javaid et al. 2004).

On the basis of wide natural distribution and diverse populations cultivation under diverse agroclimates of the world and the observed wide morphological variation, high incidence of markers polymorphism was expected on screening nucleotide sequences among accessions of *A. hypogaea.* A range of molecular markers such as RFLP, RAPD, AFLP and SSR markers were used for peanut

germplasm characterization in the past (Hilu and Stalker 1995; Singh et al. 2002; Kochert et al. 1996; Subramanian et al. 2000; Dwivedi et al. 2001; He and Prakash 2001; Herselman 2003; Bravo et al. 2006; Koppulu et al. 2010) showing low to moderate polymorphism. Other studies using SSRs have been used for differentiating botanical varieties and the two subspecies (He et al. 2005; Frimpong et al. 2015; Wang et al. 2015; Xiong et al. 2013; Jian et al. 2012; Khera et al. 2013). Similarly, isozyme markers could detect that the two subspecies were distinguished from each other by a few markers (Lu and Pickersgill 1993; Grieshammer and Wynne 1990). Eleven accessions of groundnut germplasm were analysed by SDS PAGE and study showed that diversity exists for seed storage protein profiles but the germplasm was not well characterized into subspecies or varieties (Masoomeh et al. 2015). An attempt was made to characterize 35 peanut cultivars raised by different pedigree using total soluble seed proteins separated by SDS-PAGE (Rao et al. 2013). A wide variation was observed in the pattern of protein bands of studied cultivars (Rao et al. 2013). However, Subramanian et al. (2000) detected highly polymorphic RAPD markers between *A. hypogaea* accessions. Raina et al. (2001) based on RAPD and ISSR data also reported high incidence of polymorphism. The dendrograms created in the latter study precisely organized five botanical varieties of two subspecies into five clusters.

Based on the present results, it can be ascertained that *Arachis hypogaea* is moderately genetically diverse. The genetic diversity analyses showing variation within the species in the present study reveals that seed protein fractions can serve as an important marker to assess the variability among germplasm and also distinguish between the subspecies and varieties.

Despite the abundance of information on morphological, cytogenetic, seed protein, isozyme, genome size, RFLP and RAPD analysis, FISH mapping and rDNA loci ITS sequences (Seetharam et al. 1973; Varsai Muhammad 1973; Krapovickas et al. 1974; Gregory et al. 1973; Gregory and Gregory 1976; Smartt et al. 1978a, b; Gregory et al. 1980; Singh and Moss 1982; Smartt and Stalker 1982; Klozova et al. 1983a, b; Krishna and Mitra 1987; Singh 1986, 1988; Kochert et al. 1996; Singh et al. 1991; Lu and Pickersgill 1993; Lanham et al. 1994; Stalker et al. 1994; Kochert et al. 1996; Singh et al. 1996; Robledo and Seijo 2010; Bechara et al. 2010; Halward et al. 1991, 1992), the phylogenetic relationships between taxa within *Arachis* sections and particularly on ancestry of cultivated *A. hypogaea* remains unclear and inconsistent. However, most study conclude that *A. hypogaea* is an allotetraploid (AABB) originated from wild allotetraploid, *A. monticola* (Krapovickas and Rigonii 1957; Smartt and Gregory 1967). In the present study, wild allotetraploid *Arachis monticola* (AABB) showed the same protein profile as that of cultivated allotetraploid (AABB) *A. hypogaea.* It shared the same subcluster with the cultivated species in the cluster analysis based on native proteins, SDS PAGE and combined data analysis. The present work, therefore, supports the view of previous authors (Gregory and Gregory 1976; Krapovickas and Gregory 1994; Moretzsohn et al. 2004; Seijo et al. 2004, 2007; Koppulu et al. 2010) that *A. monticola* is the most probable ancestor to *A. hypogaea.*

Several diploid species in section Arachis have been proposed as the possible donor of the A genome (Smartt et al. 1978a, b; Klozova et al. 1983a; Krishna and Mitra 1988; Singh et al. 1991; Kochert et al. 1991; Paik-Ro et al. 1992; Lanham et al. 1994; Raina and Mukai 1998, 1999; Raina et al. 2001; Milla et al. 2005; Seijo et al. 2007; Moretzsohn et al. 2013; Calbrix et al. 2012). Similarly, many species have been proposed for B genome (Singh et al. 1991; Stalker and Moss 1987; Kochert et al. 1996; Fernandez and Krapovickas 1994; Raina and Mukai 1998; Moretzsohn et al. 2013) without any firm conclusion.

Our results based on the seed protein profiles also do not show conclusive evidence of proposing two species in section Arachis which might have contributed A and B genomes in the synthesis of allotetraploid (AABB) *A. hypogaea* in nature.

The consensus tree created on the basis of combined data of native and denatured seed storage proteins revealed *A. duranensis, A. correntina, A. diogoi, A. cardensaii*, *A. benensis, A. batizocoi*close to the two tetraploid species (Fig. 9). Thus, in the present study, one of the four A genome species, i.e. *A. duranensis, A. correntina, A. cardenasii, A. diogoi,* and one of the two B genome species *A. batizocoi, A. benensis* appear possible ancestors of *A. hypogaea*. Smartt et al. (1978a), Klozova et al. (1983a), Krishna and Mitra (1988) and Lanham et al. (1994) supported *A. cardensaii* for being one of the most probable donor of A genome. But this claim is contradicted by other works of Kochert et al. (1991, 1996), Moretzschon et al. (2012), Calbrix et al. (2012) in support of *A. duranensis*, *A. helodes*, and *A. simpsonii* (Millow et al. 2005). Cross-compatibility, chromosome pairing, and hybrid fertility suggest that *A. batizocoi* or *A. hoehnei* contributed B genome and *A. duranensis, A. villosa* or *A. cardenasii* contributed A genome (Smartt et al. 1978a; Singh and Moss 1984; Singh 1986, 1988; Mallikarjuna et al. 2006). Out of the 26 diploid peanut species studied, only *A. duranensis* (A genome) and *A. ipaensis* (B genome) contained the correct complement of seed storage protein coding regions (Calbrix et al. 2012).

The cluster analysis (Fig. 7) also shows that B genome species *A. valida*, *A. ipaensis, A. magna* grouped separately from other A genome species. *A. magna* and *A. ipaensis* were closely related by Bechara et al. (2010).

At generic level, two species each were analyzed in sections Erectoides and Heteranthae. Two species *A. pusilla* and *A. sylvestris* of section Heteranthae do not cluster together. The species of section Erectoides and Procumbentes grouped together in the same cluster except one (*A. chiquitana*), supporting the observations of Singh et al. (1994), Lanham et al. (1994) and Koppolu et al. (2010). Among the five species of section Procumbentes, *A. matiensis* and *A. rigonii* were the closest as they shared the same node. Earlier, Procumbentes was considered to be a series of section Erectoides (Bechara et al. 2010). The relationship of sections Procumbentes and Erectoides is to be further resolved as supported by He et al. (2014). On the basis of SDS PAGE of seed proteins, Section Arachis was found to be phylogenetically closest to section Erectoides followed by Procumbensae, Ambinervosae, Caulorhizae, Triseminatae and Extranervosae, respectively (Singh et al. 1994). The present phylogenetic analysis has maintained the assigned primary, secondary genepools to section Arachis and tertiary gene pools to Procumbentes and Erectoides sections among *Arachis* species (Singh and Nigam 2016).

However, the placement of single accessions of *A. triseminata* (section Triseminatae) and *A. pintoi* (section Caulorrhizae) along with section Arachis species in the same cluster is a very different observation here, though both the species have quite distinct seed protein profiles. The two species also form distantly related units in phylogenetic trees based on native and SDS proteins data separately. In many studies *A. triseminata* is not closely related to section Arachis species (Krapovickas and Gregory 1994; He et al. 2014; Bechara et al. 2010). Similarly, Caulorrhizae was most distantly placed section, corroborating the results of many workers (Singh et al. 1994; Friend et al. 2010; Bechara et al. 2010).

It is suggested that the number of accessions within the cultivated and related wild species should be increased for better understanding of genomic relationships. In conclusion, the present seed protein results provide one of the most comprehensive information on genetic diversity within the cultivated peanut (*A. hypogaea*), and the divergence of wild species of other sections, to facilitate conservation and their use in genetic improvement of groundnut.

Acknowledgements The first author is thankful to Council for Scientific and Industrial Research (CSIR), India for financial assistance.

Compliance with ethical standards

Conflict of interest We declare no conflict of interest.

References

Adriaanse A, Klop W, Robbers JE (1969) Characterization of *Phaseolus vulgaris* cultivars by their electrophoretic pattern. J Sci Food Agric 20:647–650

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

Bechara MD, Moretzsohn MC, Palmieri DA, Monteiro JP, Bacci M, Martins J, Valls JFM, Lopes CR, Gimenes MA (2010) Phylogenetic relationships in genus Arachis based on ITS and 5.8S rDNA sequences. BMC Plant Biol 10:255

Bertozo MR, Valls JFM (2001) Seed storage protein electrophoresis in *Arachis pintoi* and *A. repens* (Leguminosae) for evaluating genetic diversity. Genet Res Crop Evol 48:121–130

Bianchi-Hall CM, Keys RD, Stalker HT, Murphy JP (1993) Diversity of seed storage protein patterns in wild peanut (*Arachis*, Fabaceae) species. Plant Syst Evol 186:1–15

Bradford MM (1976) Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254

Bravo JP, Hoshino AA, Angelici CMLCD, Lopes CR, Gimenes MA (2006) Transferability and use of microsatellite markers for the genetic analysis of the germplasm of some Arachis section species of the genus Arachis. Genet Mol Biol 29:516–524

Calbrix RG, Beilinson V, Stalker HT, Nielsen NC (2012) Diversity of seed storage proteins of *Arachis hypogaea* and related species. Crop Sci 52:1676–1688

Cherry JP (1975) Comparative studies of seed protein and enzymes of species and collections of *Arachis* by gel electrophoresis. Peanut Sci 2:57–65

Cunha FB, Nobile PM, Hoshino AA, Moretzsohn MC, Lopes CR, Gimenes MA (2008) Genetic relationships among *Arachis hypogaea* L. (AABB) and diploid *Arachis* species with AA and BB genomes. Genet Resour Crop Evol 55:15–20

de Lumen BO (1990) Molecular approaches to improving the nutritional and functional properties of plant seeds as food sources: developments and comments. J Agric Food Chem 38:1779–1788

Dunhill PM, Fowden L (1965) The amino acids of seeds of Cucurbitaceae. Phytochemistry 4:933–934

Dwivedi SL, Gurtu S, Chandra S, Yuejin W, Nigam SN (2001) Assessment of genetic diversity among selected groundnut germplasm. I. RAPD analysis. Plant Breed 120:345–349

Emre I (2009) Electrophoretic analysis of some *Lathyrus* L. species based on seed storage proteins. Genet Resour Crop Evol 56:31–38

Fernandez A, Krapovickas A (1994) Cromosomas y evolucion en *Arachis* (Leguminosae). Bonplandia 8:187–220

Friend SA, Quandt D, Tallury SP, Stalker HT, Hilu KW (2010) Species, genomes and section relationships in genus *Arachis* (Fabaceae): a molecular phylogeny. Plant Syst Evol 290:185–199

Frimpong RO, Sriswathi M, Ntare BR, Dakora FD (2015) Assessing the genetic diversity of 48 groundnut (*Arachis hypogaea* L.) genotypes in the Guinea savanna agro-ecology of Ghana, using microsatellite-based markers. Afr J Biotechnol 14:2484–2493

Gray JR, Fairbrothers DE, Quirm JA (1973) Biochemical and anatomical population variation in the *Danthonia sericea* complex. Bot Gazzettes 134:166–173

Gregory WC, Gregory MP (1976) Groundnut, *Arachis hypogaea* (Leguminosae-Papilionadae). In: Simmonds NW (ed) Evolution of crop plants. Longman Group Ltd., Harlow, 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) Structure and genetic resources of peanuts. In: Wilson CT (ed) Peanuts—culture and uses. The American Peanut Research and Education Association, Stillwater, OK, 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, and Wynne JC (1990) 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 markets among unadapted germplasm resources of cultivated peanut and related wild species. Genome 34:1013–1020

Halward TM, Stalker HT, Larue EA, Kochert G (1992) Use of single primer DNA amplifications in genetic studies of peanut (*Arachis hypogaea* L). Plant Mol Biol 18:315–325

He GH, Prakash CS (2001) Evaluation of genetic relationships among botanical varieties of cultivated peanut (*Arachis hypogaea* L.) using AFLP markers. Genet Resour Crop Evol 48:347–352

He GH, Meng RH, Gao H, Guo BZ, Gao GQ, Newman M, Pittman RN, Prakash CS (2005) Simple sequence repeat markers for botanical varieties of cultivated peanut (*Arachis hypogaea* L.). Euphytica 142:131–136

He G, Barkley NA, Zhao Y, Yuan M, Prakash CS (2014) Phylogenetic relationships of species of genus Arachis based on genic sequences. Genome 57:327–334

Herselman L (2003) Genetic variation among Southern African cultivated peanut (*Arachis hypogaea* L.) genotypes as revealed by AFLP analysis. Euphytica 133:319–327

Hilu KW, Stalker HT (1995) Genetic relationships between peanut and wild species of *Arachis* sect. Arachis (Fabaceae): evidence from RAPDs. Plant Syst Evol 198:167–178

Javaid A, Ghafoor A, Anwar R (2004) Seed storage protein electrophoresis in groundnut for evaluating genetic diversity. Pak J Bot 36(1):25–29

Jian G, Li-Feng L, Shun-Li C, Huan-Ying C, Xin-Yan W (2012) Studies on genetic diversity of peanut (*Arachis hypogaea* L.) varieties bred in Hebei, Shandong and Henan provinces. J Plant Genet Res 13:201–206

Khera P, Upadhyaya HD, Pandey MK, Roorkiwal M, Sriswathi M, Janila P, Guo YM, Mackain MR, Nagy ED, Knapp SJ, Mack JL, Conner JA, Akins PO, Varshney RK (2013) Single nucleotide polymorphism-based genetic diversity in the reference set of peanut (*Arachis* spp.) by developing and applying cost-effective kompetitive allele specific polymerase chain reaction genotyping assays. Plant Genome. https://doi.org/10.3835/plantgenome2013.06.0019

Klozova E, Turkova V, Smartt J, Pitterora K, Svachulova J (1983a) Immunochemical characterization of seed proteins of some species of the genus *Arachis* L. Biol Planta 25:201–208

Klozova E, Svachulova J, Smartt J, Hadac E, Turkova V, and Hadacova V (1983b) The comparison of seed protein patterns within the genus Arachis by polyacrylamide Gel Electrophoresis. Biolo Planta 25:266–273

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

Kochert G, Stalker HM, Gimenes M, Galgaro L, Lopes CR, Moore K (1996) RFLP and cytogenetic evidence on the origins and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). Am J Bot 83:1282–1291

Koppolu R, Upadhyaya HD, Dwivedi SL, Hoisington DA, Varshney RK (2010) Genetic relationships among seven sections of genus Arachis studied by using SSR markers. BMC Plant Biol 10:15–27

Kottapalli KR, Payton P, Rakwal R, Agrawal GA, Shibato J, Burow M et al (2008) Proteomics analysis of mature seeds of four peanut cultivars using two-dimensional gel electrophoresis reveals distinct differential expression of storage, anti-nutritional and allergenic proteins. Plant Sci 175:321–329

Krapovickas A, Gregory WC (1994) Taxonomical del genera *Arachis* (Leguminosae). Bonplandia 8:1–184

Krapovickas A, Rigonii VA (1957) Neuvasespecies de Arachis vinculades al problema del origin del mani. Dorwiniana 11:431–455

Krapovickas A, Fernanadez A, Seeligman MP (1974) Recuperaction de la fertilidad en un hibgido interspecifico de *Arachis* (Leguminosae). Bonplandia 3:129–142

Krishna TG, Mitra R (1987) Arachin polymorphism in groundnut (*A. hypogaea* L.). Phytochemistry 26:897–902

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

Krishna TG, Pawar SE, Mitra R (1986) Variation and inheritance of the arachin polypeptides of groundnut (*A. hypogaea* L.). Theor Appl Genet 73:82–87

Ladizinsky G, Hymowitz T (1979) Seed protein electrophoresis in taxonomic and evolutionary studies. Theor Appl Genet 54:145–151

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

Lanham PG, Forster BP, McNicol P, Moss JP, Powell W (1994) Seed storage protein variation in *Arachis* species. Genome 37:487–496

Liang XQ, Luo M, Holbrook CC, Guo BZ (2006) Storage protein profiles in Spanish and runner market type peanuts and potential markers. BMC Plant Biol 6:24

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

Malik FA, Qureshi AS, Ashraf M, Khan MR, Javed A (2009) Evaluation of genetic diversity in soybean (*Glycine max*) lines using seed protein electrophoresis. Aust J Crop Sci 3:107–112

Mallikarjuna N, Tandra SK, Jadhav DR (2006) Arachis hoehnei the probable B genome donor of *Arachis hypogaea* based on crossability, cytogenetical and molecular studies. J SAT Agric Res 2:1–2

Masoomeh J, Mehran G, Fatemeh J (2015) Seed storage protein electrophoresis for identification of some groundnut (*Arachis hypogaea* L.). Int Res J Appl Basic Sci 9:1718–1721

Masoumi SM, Kahrizi D, Rostami-Ahmadvandi H, Soorni J, Kiani S, Mostafaie A, Yari K (2012) Genetic diversity study of some medicinal plant accessions belong to Apiaceae family based on seed storage proteins patterns. Mol Biol Rep 39:10361–10365

Milla SR, Isleib TG, Stalker HT (2005) Taxonomic relationships among *Arachis* sect. *Arachis* species as revealed by AFLP markers. Genome 48:1–11

Moretzsohn MC, Hopkins MS, Mitchell SE, Kresovich S, Valls JFM, Ferreira ME (2004) Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis of hypervariable regions of the genome. BMC Plant Biol 4:11

Moretzsohn M, Ediene G, Gouvea EG, Inglis PW, Leal-Bertioli SC, Valls JF, Bertioli DJ (2013) A study of the relationships of cultivated peanut (*Arachis hypogaea*) and its most closely related wild species using intron sequences and microsatellite markers. Ann Bot 111:113–126

Paik-Ro OG, Smith RL, Knaft DT (1992) Restriction fragment length polymorphism evaluation of six peanut species within the Arachis section. Theor Appl Genet 84:201–208

Panda RC, Kumar OA, Rao KGR (1986) The use of seed protein electrophoresis in the study of phylogenetic relationships in chilli pepper (*Capsicum* L.). Theor Appl Genet 72:665–670

Panigrahi J, Kumar DR, Mishra M, Mishra RP, Jena P (2007) Genomic relationships among 11 species in the genus Cajanus as revealed by seed protein (albumin and globulin) polymorphisms. Plant Biotechnol Rep 1:109–116

Peñaloza A, Valls JFM (2005) Chromosome number and satellite chromosome morphology of eleven species of *Arachis* (Leguminosae). Bonplandia 14:65–72

Perrier X, Jacquemoud-Collet JP (2006) DARwin software http://darwin.cirad.fr/darwin. Accessed 11 Jan 2017

Przybylska J, Blixt S, Hurich J, Zimniak-Przybylska Z (1977) Comparative study of seed proteins in the genus Pisum. I. Electrophoretic patterns of different protein fractions. Genet Pol 18:27–38

Raina SN, Mukai Y (1998) Genomic in situ hybridization in *Arachis* (Fabaceae) identifies the diploid wild progenitors of cultivated (*A. hypogaea*) and related wild (*A. monticola*) peanut species. Plant Syst Evol 214:251–262

Raina SN, Mukai Y (1999) Detection of a variable number of 18S-5.8S-26S and 5S ribosomal DNA loci by fluorescent in situ hybridization in diploid and tetraploid *Arachis* species. Genome 42:52–59

Raina SN, Rani V, Kojima T, Ogihara Y, Singh KP, Devarumath RM (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 44:576–586

Rao PS, Bharathi M, Reddy KB (2013) Identification of peanut (*Arachis hypogaea* L.) varieties through chemical tests and electrophoresis of soluble seed proteins. Legume Res 36:475–483

Robledo G, Seijo G (2010) Species relationships among the wild B genome of Arachis species (section Arachis) based on FISH mapping of rDNA loci and heterochromatin detection: a new proposal for genome arrangement. Theor Appl Genet 121:1033–1046

Rohlf EJ (1993) NTSYS-pc: numerical taxonomy and multivariate analysis system, version 1.80. Applied Biostatistics Inc., Setauket

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sathaiah V, Reddy TP (1985) Seed protein profiles of castor (*Ricinus communis* L.) and some Jatropha species. Genet Agraria 39:35–43

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

Schroeder HE (1982) Quantitative studies on the cotyledonary proteins in the genus *Pisum*. J Sci Food Agric 33:623–633

Seetharam A, Nayar KMD, Sreekantaradhya R, Achar DKT (1973) Cytological studies in the interspecific hybrid of *Arachis hypogaea* × *A. duranensis*. Cytologia 38:277–280

Seijo JG, Lavia GI, Fernandez A, Krapovickas A, Ducasse D, Moscone EA (2004) Physical mapping of the 5S and 18S-25S rRNA genes by FISH as evidence that *Arachis duranensis* and *A. ipaënsis* are the wild diploid progenitors of *A. hypogaea* (Leguminosae). Am J Bot 91:1294–1303

Seijo JG, Lavia GI, Fernández A, Krapovickas A, Ducasse DA, Bertioli DJ, Moscone EA (2007) Genomic relationships between the cultivated peanut (*Arachis hypogaea*—Leguminosae) and its close relatives revealed by double GISH. Am J Bot 94:1963–1971

Singh AK (1986) Utilization of wild relatives in the genetic improvement of *Arachis hypogaea* L. Part 8. Synthetic amphidiploids and their importance in interspecific breeding. Theor Appl Genet 72:433–439

Singh AK (1988) Putative genome donors of *A. hypogaea* L. Evidence from crosses with synthetic amphidiploids. Plant Syst Evol 160:143–153

Singh AK, Moss JP (1982) Utilization of wild relatives in genetic improvement of *Arachis hypogaea* L. Part 2: chromosome complements of species in 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. Part 5. Genome analysis in section Arachis and its implications in gene transfer. Theor Appl Genet 68:355–364

Singh AK, Nigam SN (2016) *Arachis* gene pools and genetic improvement in groundnut. In: Rajpal VR et al (eds) Gene pool diversity and crop improvement, sustainable development and biodiversity. Springer, Cham, pp 17–75

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

Singh AK, Gurtu S, Jambunathan R (1994) Phylogenetic relationships in the genus *Arachis* based on seed protein profiles. Euphytica 74:219–225

Singh KP, Raina SN, Singh AK (1996) Variation in chromosomal DNA associated with the evolution of *Arachis* species. Genome 39:890–897

Singh KP, Singh A, Raina SN, Singh AK, Ogihara Y (2002) Ribosomal DNA repeat unit polymorphism and heritability in peanut (*Arachis hypogaea* L.) accessions and related wild species. Euphytica 123:211–220

Smartt J, Gregory WC (1967) Interspecific cross-compatibility between the cultivated peanut (*A. hypogaea*) and other members of the genus Arachis. Oleaginenx 22:455–459

Smartt J, Stalker HT (1982) Speciation and cytogenetics in *Arachis*. In: Pattee HE, Young CT (eds) Peanut science and technology. American Peanut Research and Education Society, Yokum, pp 21–49

Smartt J, Gregory WC, Gregory MP (1978a) b The genomes of *Arachis hypogaea*. 2. The implications of interspecific breeding. Euphytica 27:677–680

Smartt J, Gregory WC, Gregory MP (1978b) a The genomes of *Arachis hypogaea*. 1. Cytogenetic studies of putative genome donors. Euphytica 27:665–675

Sneath PHA, Sokal RR (1973) Numerical taxonomy. W. H. Freeman and Company, San Francisco

Song B, Oehrle NW, Liu S, Krishnan HB (2016) Characterization of seed storage proteins of several perennial glycine species. J Agric Food Chem 64:8499–8508

Stalker HT (1991) A new species section Arachis of peanuts with D genome. Am J Bot 78:630–637

Stalker HT, Moss JP (1987) Speciation, cytogenetics and utilization of *Arachis* species. Adv Agron 41:1–37

Stalker HT, Jones TM, Murphy JP (1990) Isozyme variability among *Arachis* species. Am Peanut Res Educ Soc 22:50

Stalker HT, Phillips TD, Murphy JP, Jones TM (1994) Variation in isozyme patterns among *Arachis* species. Theor Appl Genet 87:746–755

Subramanian V, Gurtu S, Nageswara Rao RC, Nigam SN (2000) Identification of DNA polymorphism in cultivated groundnut using random amplified polymorphic DNA (RAPD) assay. Genome 43:656–660

Tripathy SK, Mohanty P, Jena M, Dash GB, Pradhan K, Nayak PK, Dash S, Lenka D, Mishra D, Mohapatra PM, Swain D, Senapati N (2016) Revealing contrasting genetic variation and study of genetic diversity in urdbean (*Vigna mungo* (L.) Hepper using SDS-PAGE of seed storage proteins. Res Biotechnol 7:11–20

Valls JFM, Simpson CE (2005) New species of *Arachis* L. (Leguminosae) from Brazil, Paraguay and Bolivia. Bonplandia 14:35–63

Varsai Muhammad S (1973) Cytogenetical investigations in the genus *Arachis* L. II. Triploid hybrids and their derivatives. Madras Agric J 60:1414–1427

Wang H, Khera P, Huang B, Yuan M, Katam R, Zhuang W, Shultz KH, Moore KM, Culbreath AK, Zhang X, Varshney RK, Xie L, Guo B (2015) Analysis of genetic diversity and population structure of peanut cultivars and breeding lines from China, India and the US using simple sequence repeat markers. J Integr Plant Biol. https://doi.org/10.1111/jipb.12380

Xiong F, Liu J, Jiang J, Zhong R, He L, Han Z, Li Z, Tang X, Tang R (2013) Molecular profiling of genetic variability in domesticated groundnut (*Arachishypogaea* L.) based on ISJ, URP, and DAMD markers. Biochem Genet 51:889–900