Theor Appl Genet (2005) 111: 1229–1237
DOI 10.1007/s00122-005-0017-0

ORIGINAL PAPER

SP Tallury · KW Hilu · SR Milla · SA Friend
M Alsaghir · HT Stalker · D Quandt

Genomic affinities in *Arachis* section *Arachis* (Fabaceae): molecular and cytogenetic evidence

Received: 5 January 2005 / Accepted: 14 June 2005 / Published online: 27 September 2005

Abstract Section *Arachis* is the largest of nine sections in the genus *Arachis* and includes domesticated peanut, *A. hypogaea* L. Most species are diploids ($x = 10$) with two tetraploids and a few aneuploids. Three genome types have been recognized in this section (A, B and D), but the genomes are not well characterized and relationships of several newly described species are uncertain. To clarify genomic relationships in section *Arachis*, cytogenetic information and molecular data from amplified fragment length polymorphism (AFLP) and the *trn*T-F plastid region were used to provide an additional insight into genome composition and species relationships. Cytogenetic information supports earlier observations on genome types of *A. cruziana*, *A. herzogii*, *A. kempff-mercadoi* and *A. kuhlmannii* but was inconclusive about the genome composition of *A. benensis*, *A. hoehnei*, *A. ipaensis*, *A. palustris*, *A. praecox* and *A. williamsii*. An AFLP dendrogram resolved species into four major clusters and showed *A. hypogaea* grouping closely with *A. ipaensis* and *A. williamsii*. Sequence data of the *trn*T-F region provided genome-specific information and showed for the first time that the B and D genomes are more closely related to each other than to the A genome. Integration of information from cytogenetics and biparentally and maternally inherited genomic regions show promise in understanding genome types and relationships in *Arachis*.

Communicated by S. J. Knapp

SP Tallury (✉) · SR Milla · HT Stalker
Department of Crop Science, North Carolina State University, Raleigh, NC, 27695-7629, USA
E-mail: tallury@ncsu.edu
Tel.: +1-919-5153809
Fax: +1-919-5155657

KW Hilu · SA Friend · M Alsaghir · D Quandt
Department of Biology, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061, USA

Present address: D Quandt
Institute for Botany, Technical University Dresden, 01062 Dresden, Germany

Introduction

Peanut or groundnut, *Arachis hypogaea* L., is a semi-tropical legume native to South America and is currently cultivated in more than 100 countries. The genus *Arachis* contains 70–80 species (Krapovickas and Gregory 1994; C. E. Simpson, personal communication). Most species in the genus are diploid ($2n = 2x = 20$), but tetraploid ($2n = 4x = 40$) and aneuploid ($2n = 2x = 18$) species also have been described (Krapovickas and Gregory 1994; Lavia 1998). These species are grouped into nine sections based on morphological characters and geographical distribution (Krapovickas and Gregory 1994). Section *Arachis* is the largest with 26 species and also includes the allotetraploid cultivated species, *A. hypogaea*. Most species of section *Arachis* possess an A genome, with only a few having a B and one with a D genome (Smartt et al. 1978; Stalker 1991; Holbrook and Stalker 2003). The A-genome species are characterized by the presence of a small chromosome pair that is lacking in B and D genome species (Smartt et al. 1978; Stalker 1991). Further, the D-genome species, *A. glandulifera*, has an asymmetrical karyotype as compared to A and B genome species (Stalker 1991). Historically, genome analysis in section *Arachis* has included somatic karyotyping using rapidly dividing cells from root tip squashes stained with basic fuchsin/aceto-carmine (Stalker and Dalmacio 1981; Singh and Moss 1982). However, such conventional karyotyping is highly laborious and may result in ambiguous conclusions about species relationships. Consequently, a series of studies were conducted to understand the phylogenetic relationships among species in section *Arachis* using characters from morphology (Stalker 1990), isozyme variation (Grieshammer and Wynne 1990; Lu and Pickersgill 1993; Stalker et al. 1994), seed storage proteins (Singh et al. 1991;

Bianchi-Hall et al. 1993), restriction fragment length polymorphisms (RFLPs) (Halward et al. 1991; Kochert et al. 1991; Paik-Ro et al. 1992), simple sequence repeats (SSRs) (Hopkins et al. 1999), randomly amplified polymorphic DNAs (RAPDs) (Halward et al. 1992; Lanham et al. 1992; Hilu and Stalker 1995), amplified fragment length polymorphisms (AFLPs) (Milla et al. 2005) and in situ hybridization (Raina and Mukai 1999). However, ambiguities remain concerning species and genome relationships due to incongruence and/or incomplete resolution caused by the nature of the methods used in these studies. Furthermore, these methods did not generate fingerprinting characters for the three genomes in section *Arachis*.

In theory, the 26 species of section *Arachis* are cross compatible with *A. hypogaea*, but few attempts have been made to hybridize these species with peanut and/or other A and B genome species to clarify genome identities and species relationships. Although such information is available for the originally described section *Arachis* species (Stalker et al. 1991), it is lacking for many of the newly described species. Thus, the objective of this study is to provide insight into genome composition and species relationships in section *Arachis*, with emphasis on newly described species of uncertain genomes. Conventional cytogenetic approaches such as crossability, male fertility and chromosomal associations in F_1 hybrids, and molecular information from AFLPs and the plastid *trn*T-F region are compared and analyzed to assess genome and species relationships.

Materials and methods

Plant materials

Cytogenetic experiments involved 12 of the 13 diploid species of section *Arachis*; excluded one was the D genome species, *A. glandulifera* (Table 1). The *trn*T-F and AFLP studies included all the 13 species, and also representatives of the six botanical varieties of *A. hypogaea* as well as additional accessions of *A. duranensis* (A genome, KGBSPSc 30073) and *A. glandulifera* (D genome, KGSSc 30091 and KGSSc 30098). Additionally, four interspecific F_1 hybrids (Table 1) were also examined in the *trn*T-F study to determine the inheritance of the plastid genome in *Arachis*.

Crossability study

Arachis duranensis and *A. batizocoi* were used as representative A and B genome female tester species, respectively, to make crosses with 10 newly described species of uncertain genome types. The principal reasons for choosing the tester species are due to the fact that *A. duranensis* contains the small chromosome pair characteristic of A-genome species, and that *A. batizocoi* is the original B-genome species described. These species are easily hybridized with other section *Arachis* species when compared to earlier published data which used the same accessions as tester species (Stalker 1991). The crossing

Table 1 Plant materials analyzed for genome identification and affinities

Arachis species	Collector no.	PI no. or line
A. duranensis	K 7988	PI 219823
A. duranensis	KGBSPSc 30073	PI 468319
A. batizocoi	K 9484	PI 298639
A. benensis	KGSPSc 35005	PI475877
A. cruziana	KSSc 36024	PI 476003
A. herzogii	KSSc 36029	PI 476008
A. hoehnei	KG 30006	PI 468150
A. ipaensis	KGBSPSc 30076	PI 468322
A. kempff-mercadoi	KGSSc 30088	PI 468333
A. kuhlmannii	VKSSv 8888	–
A. palustris	VPmSv 13023	–
A. praecox	VSGr 6416	PI 476128
A. williamsii	WiCla 1118	–
A. glandulifera	KGSSc 30091	PI 468336
A. glandulifera	KGSSc 30098	PI 468341
A. hypogaea Var. hypogaea		PI 339954& NC 4
A. hypogaea Var. hirsuta		PI 501296
A. hypogaea Var. fastigiata		PI 339960& NM Valencia C
A. hypogaea Var. vulgaris		PI 261924
A. hypogaea Var. peruviana		PI 590455& PI 275751
A. hypogaea Var. aequatoriana		PI 497615& Grif. 12581
F_1 Hybrids		
A.batizocoi × *A. duranensis* (KGBSPSc 30073)		
A.duranensis (KGBSPSc 30073) × *A. batizocoi*		
A.duranensis (K 7988) × *A. ipaensis*		
A. glandulifera (KGSSc 30091) × *A. herzogii*		

program was conducted during the summer of 1999 in the North Carolina State University greenhouses. Plants were grown in plastic flats. Unopened flower buds of female parents were emasculated between 3:00 and 5:00 P.M. and pollinations were completed between 8:00 and 10:00 A.M. the following morning. Unemasculated flower buds were removed throughout the flowering period. The number of pollinated flowers for each cross was recorded daily and the number of pegs formed after pollinations was recorded weekly (Table 2). The formation of pegs following pollinations is an indication of successful fertilization; and formation of pods with a mature seed is regarded as a measure of compatibilty of parental genomes. About 60 days after the last pollination, pods were harvested, dried and the numbers of seeds were recorded. Crossability percentage was calculated from the total number of pods harvested divided by the total number of pollinations performed (Table 2).

Male fertility (pollen viability)

F_1 hybrid seeds from the crossability study were planted in pots in the greenhouse and a single flower was harvested between 8:00 and 8:30 A.M. in two different days from each plant. Pollen was squeezed out of the anthers onto a glass slide and stained with 2% aceto-carmine for about 2 h. Using a light microscope, darkly stained, well-formed and large pollen grains were considered viable whereas unstained, misshaped pollen grains were classified as non-viable. Percentage of viable pollen was calculated as the average of viable/total pollen grains of the two flowers and used as an indicator of pollen viability or male fertility.

Chromosomal associations in F_1 hybrids

Immature flower buds were collected between 8:00 and 9:00 A.M. and fixed in a 1:3:6 mixture of glacial acetic acid: chloroform: 70% ethyl alcohol for 48 h, transferred to 70% alcohol, and stored in a refrigerator. Anthers were squashed in a drop of 2% aceto-carmine, and chromosomal associations for each cross were observed in microsporocytes at metaphase I and recorded as univalents (I), bivalents (II), trivalents (III) and quadrivalents (IV) (Table 3).

Molecular experiments

Genomic DNA was extracted from fresh leaf material following Johnson et al. (1995) and Milla et al. (2005), and stored at −80°C.

AFLP study

PCR amplification and AFLP analysis was conducted on 13 diploid species of *Arachis* and 10 genotypes of *A. hypogaea* (Table 1) following Milla et al. (2005). Eight primer combinations, viz., E-AAC + M-CAT, E-ACG + M-CAA, E-ACT + M-CAA, E-ACT + M-CAC, E-ACT + M-CAG, E-ACT + M-CAT, E-AGT + M-CTA, and E-ATC + M-CGC (Sigma Genosys, The Woodlands, TX) were used to generate AFLP fingerprints. AFLP fragments were separated on a LI-COR 4200 DNA Analyzer Sequencer on 25 cm gels using 8% denaturing polyacrylamide gels. Near-infrared-labeled size standards (LI-COR Inc. Lincoln, NE)

Table 2 Crossability of *Arachis* species with representative A (*A. duranensis*) and B (*A. batizocoi*) genome tester species

Cross	Pollinations no.	Pegs no.	Pods no.	Pods/pollinations %
A. duranensis (A)				
×				
A. benensis	44	26	17[a]	39[a]
A. cruziana	42	19	21[a]	50[a]
A. herzogii	35	22	8	23
A. hoehnei	38	23	4	11
A. ipaensis	27	15	15	56
A. kempff-mercadoi	26	12	4	15
A. kuhlmannii	29	16	7	24
A. palustris	34	23	12[a]	35[a]
A. praecox	20	8	5[a]	25[a]
A. williamsii	43	24	19	44
A. batizocoi (B)				
×				
A. benensis	38	21	17[a]	45[a]
A. cruziana	41	24	23	56
A. herzogii	33	15	7	21
A. hoehnei	27	14	7	26
A. ipaensis	26	14	14	54
A. kempff-mercadoi	34	12	7	21
A. kuhlmannii	27	15	6	22[b]
A. palustris	31	11	11[a]	35[a]
A. praecox	17	9	9[a]	53[a]
A. williamsii	29	14	8	28

[a]Aborted seeds
[b]Plants did not survive

Table 3 Pollen viability and chromosomal associations in F_1 hybrids of *Arachis* species with representative A (*A. duranensis*) and B (*A. batizocoi*) genome tester species

Cross	Pollen viability[a]	No. of cells analyzed	Mean chromosome associations			
			I	II	III	IV
A. duranensis (A) ×						
A. herzogii	42.9	53	0.5	9.7	0	0.02
A. hoehnei	17.3	8	8.8	5.6	0	0
A. ipaensis	7.1	30	6.3	6.8	0	0
A. kempff-mercadoi	72.0	30	0	10.0	0	0
A. kuhlmannii	58.6	26	0.1	9.9	0	0
A. williamsii	36.8	37	5.4	7.3	0	0
A. batizocoi (B) ×						
A. cruziana	36.4	21	0.3	9.9	0	0
A. herzogii	1.0	10	7.6	6.2	0	0
A. hoehnei	1.0	30	3.9	8.0	0	0
A. ipaensis	4.0	20	5.7	6.9	0.2	0
A. kempff-mercadoi	1.0	12	9.0	5.5	0	0
A. williamsii	2.2	12	8.3	5.8	0	0

[a]Average of viable pollen/total pollen grains of two flowers

were loaded on each gel for sizing of the AFLP fragments. The AFLP-Quantar 1.0 (Keygene Products B.V., Wageningen, The Netherlands) software package was used to score the presence or absence of distinct, reproducible, and unambiguously resolved bands. The scored bands were visually checked to confirm the consistency of the amplified fragments. Fragments were scored as a binary unit character (present = 1; absent = 0). A genetic similarity matrix using simple matching coefficient was generated using NTSYSpc 2.02i and a dendrogram was computed using the unweighted pair-group method with arithmetic averages (UPGMA) clustering method (Rohlf 1998).

The *trn*T-F study

PCR-amplifications (PTC-100, MJ Research, Waltham, MA) were performed in 50 µl–reactions containing 1.5 U Taq DNA polymerase (Qiagen, Inc., Valencia, CA), 1.0 mM dNTPs mix in each 0.25 mM, 1× buffer, 2.0 mM $MgCl_2$ and 10 pmol of each amplification primer, using the primer combination a/b or c/f (Taberlet et al. 1991). Amplification conditions were: 34 cycles of 94°C (1 min) denaturation, 52°C (1 min) annealing, 72°C (2 min) extension, and a 15 min final extension (72°C). The targeted fragments were either resolved on 1% agarose gel, excised and cleaned using the QIAquick extraction kit (Qiagen) or directly purified using the QIAquick PCR purification kit. Cycle sequencing was performed using the Big Dye Terminator Ready Reaction Kit employing the standard protocol. Extension products were electrophoresed on Applied Biosystems 373, 377 or 3100 automated sequencers (Applied Biosystems Inc., Foster City, CA).

Sequences were aligned manually using QuickAlign (Müller 2000), according to the alignment rules laid out in Kelchner 2000 and Borsch et al. 2003. To illustrate the phylogenetic information provided by indels in the data matrix, indels were coded as present = 1 and absent = 0 and included in Table 4. Phylogenetic reconstructions based on maximum likelihood (ML) were conducted with *win*PAUP*4.0b10 (Swofford 2002), assuming a general time reversible model (GTR + G + I) and a rate variation among sites following a gamma distribution (four categories represented by mean). The phylogram is unrooted to avoid problems concerning the choice of potential out-groups; thus, insertion or deletion (indel) events cannot be determined. Measures of support for individual clades were based on heuristic bootstrap searches under likelihood with 500 replicates and the same options in effect.

Results

Cytogenetic analysis

A. duranensis (A genome) × Arachis species hybrids

The crossability percentages of *A. duranensis* with the 10 *Arachis* species ranged from 11% in *A. duranensis* × *A. hoehnei* to 56% in *A. duranensis* × *A. ipaensis* (Table 2). However, four of the 10 crosses (with *A. benensis*, *A. cruziana*, *A. palustris*, and *A. praecox*) produced pods with aborted seeds. Of the remaining six *A. duranensis* hybrids, *A. duranensis* × *A. kempff-mercadoi* had the highest pollen viability (72%) followed by *A. duranensis* × *A. kuhlmannii* (59%), *A. duranensis* × *A. herzogii* (43%), and *A. duranensis* × *A. williamsii* (37%). The pollen viability was only 17.3% in *A. duranensis* × *A. hoehnei*, and the lowest pollen viability of 7.1% was observed in *A. duranensis* × *A. ipaensis* (Table 3). The hybrids with the highest mean bivalent associations included the three species, *A. kempff-mercadoi* (10.0), *A. kuhlmannii* (9.9) and *A. herzogii* (9.7). The mean number of bivalents ranged from 5.6 to 7.3 in the other three species hybrids, with corresponding correlation between low bivalent associations and lower male fertility (Table 3).

Table 4 Summary of characteristic substitutions and indels observed in the *trn*T-F alignment. For illustration, one state for each character (presence = 1; absence = 0) are shown in bold

Species and hybrids	Substitutions																	Indels				
																		21 bp	5 bp	4 bp	6 bp	2 bp
A. duranensis 30073	A	**A**	AAA	**A**	**G**	A	G	C	**C**	A	**A**	**G**	**C**	G	G	G	G	**1**	0	**1**	1	0
A. duranensis 7988	A	**A**	AAA	**A**	**G**	A	G	C	**C**	A	**A**	**G**	**C**	G	G	G	G	**1**	0	**1**	1	0
Hybrid 7988×30076	A	**A**	AAA	**A**	**G**	A	G	C	**C**	A	**A**	**G**	**C**	G	G	G	G	**1**	0	**1**	1	0
A. kuhlmannii	A	**A**	AAA	**A**	**G**	A	G	C	**C**	A	**A**	**G**	**C**	G	G	G	G	**1**	**1**	**1**	1	0
A. kempff-mercadoi	A	**A**	AAA	**A**	**G**	**C**	G	C	**C**	A	**A**	**G**	**C**	G	G	G	G	**1**	**1**	**1**	1	0
A. herzogii	A	**A**	AAA	**A**	**G**	**C**	G	C	**C**	A	**A**	**G**	**C**	G	G	G	G	**1**	0	**1**	1	0
Hybrid 30073×9484	A	**A**	AAA	**A**	**G**	A	G	C	**C**	A	**A**	**G**	**C**	G	G	G	G	**1**	0	**1**	1	0
A. hypogaea	A	**A**	AAA	**A**	**G**	A	**T**	C	**C**	A	**A**	**G**	**C**	G	G	G	G	**1**	0	**1**	1	0
A. hoehnei	A	**A**	AAA	**A**	**G**	A	G	C	**C**	A	**A**	**G**	**C**	G	G	G	G	0	0	**1**	1	0
A. praecox	A	T	AAA	T	C	A	G	C	T	A	C	C	G	G	G	G	G	0	0	0	1	**1**
A. palustris	A	T	**TTC**	T	C	A	G	C	T	A	C	C	G	G	G	G	G	0	0	0	1	**1**
A. benensis	A	T	AAA	T	C	A	G	C	T	**C**	C	C	G	G	G	G	**C**	0	0	0	**0**	0
A. williamsii	A	T	AAA	T	C	A	G	C	C	**C**	C	C	G	G	G	G	**C**	0	0	0	**0**	0
A. glandulifera 30098	A	T	AAA	T	C	A	G	C	T	**C**	C	C	G	G	G	G	**C**	0	0	0	**0**	0
A. glandulifera 30091	A	T	AAA	T	C	A	G	C	T	**C**	C	C	G	G	G	**T**	**C**	0	0	0	**0**	0
Hybrid 30091×36029	A	T	AAA	T	C	A	G	C	T	**C**	C	C	G	G	G	**T**	**C**	0	0	0	**0**	0
A. ipaensis	A	T	**TTC**	T	C	A	G	**G**	T	A	C	C	G	G	G	G	G	0	0	0	**0**	0
A. cruziana	**T**	T	**TTC**	T	C	A	**T**	**G**	T	A	C	C	G	**T**	G	G	G	0	0	0	**0**	0
A. batizocoi 9484	**T**	T	**TTC**	T	C	A	**T**	**G**	T	A	C	C	G	**T**	**A**	G	G	0	0	0	**0**	0
Hybrid 9484×30073	**T**	T	**TTC**	T	C	A	**T**	**G**	T	A	C	C	G	**T**	**A**	G	G	0	0	0	**0**	0

A. batizocoi (B genome) × Arachis species hybrids

The crossability percentages of *A. batizocoi* with the 10 *Arachis* species ranged from 21% in *A. batizocoi* × *A. herzogii* and *A. batizocoi* × *A. kempff-mercadoi* crosses, to 56% in *A. batizocoi* × *A. cruziana* cross (Table 2). The *A. benensis*, *A. palustris* and *A. praecox* crosses resulted in pods with aborted seeds and the hybrid seedlings of *A. batizocoi* × *A. kuhlmannii* died prior to flowering. Thus, no cytological data are available for these hybrid combinations. The remaining six crosses resulted in pods with mature seeds. In contrast to the *A. duranensis* hybrids, the male fertility of hybrids with *A. batizocoi* was significantly lower (Table 3). Only *A. batizocoi* × *A. cruziana* was semi-fertile (36%), whereas each of the remaining five hybrids had less than 5% pollen viability (Table 3). The *A. batizocoi* × *A. cruziana* hybrid had mean chromosomal distribution of 0.3 I and 9.9 II (Table 3). The remaining hybrids displayed more univalents, ranging from 3.9 in *A. batizocoi* × *A. hoehnei* to 9.0 in *A. batizocoi* × *A. kempff-mercadoi*, with correspondingly lower male fertility (Table 3).

AFLP analysis

A total of 239 unambiguously resolved, distinct and reproducible bands were scored from the amplifications of *A. hypogaea* and the 13 diploid *Arachis* species (Table 1). The scored bands ranged in size from 65 to 760 bp and the number of scored loci averaged 30 (range = 14 to 47). The reproducibility of banding patterns was confirmed by duplicate AFLP experiments from 10 random samples with the same eight primer combinations previously mentioned. The mean reproducibility (calculated as percentage of identical bands in the two repeats) ranged from 96 to 100%. The dendrogram generated from AFLP data grouped *A. hypogaea* genotypes in a tight cluster with similarity coefficients (SC) of 0.96–1.0 (Fig. 1). *Arachis ipaensis* and *A. williamsii* clustered closely (SC 0.84) and appeared in the same major cluster with *A. hypogaea*. The remaining species formed three major clusters with *A. benensis* being the only monospecific cluster. These clusters were associated with the *A. hypogaea-A. ipaensisA. williamsii* cluster at the lowest SC value (0.60). One of the three major clusters included the B genome tester species *A. batizocoi* tightly associated (SC 0.92) with the D genome species *A. glandulifera*, and these two species were most closely linked to *A. cruziana* (Fig. 1). In a second major group, one subcluster showed the A genome tester species, *A. duranensis* linked to *A. herzogii*, *A. kempff-mercadoi* and *A. kuhlmannii* and the second subcluster had *A. hoehnei* joining at a relatively low SC (0.74) to a tight cluster of the two aneuploid species *A. palustris* and *A. praecox*.

Phylogenetic analyses of the plastid trnT-F sequences

In total, the sequence alignment comprised of 1949 characters, of which 927 belonged to the *trn*T-L spacer (901–924 bp), 568 to the *trn*L intron, and 369 to the *trn*L-F spacer (344–364 bp). Together, both *trn*L exons contributed 85 characters to the matrix. Of the 1949 characters, only 36 were variable, of which 23 were potentially parsimony informative. Sequence alignment of the *trn*T-F plastid region among species revealed major and minor insertion and deletion events (indels) as well as nucleotide substitutions that are effective as markers for identifying genome types in section *Arachis*

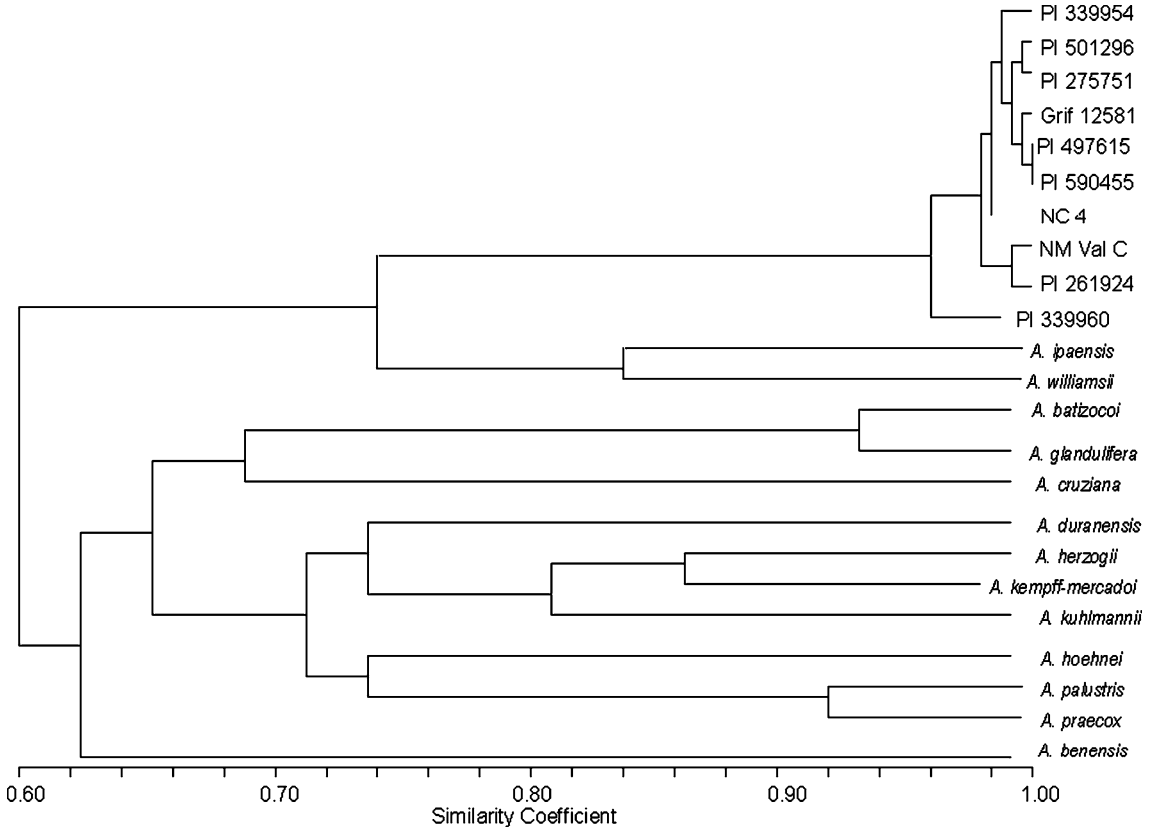

Fig. 1 Dendrogram representing grouping of *Arachis* species based on UPGMA analysis of AFLP data. See details of accession numbers for *A. hypogaea* in Table 1

(Table 4). The A genome is characterized by a large indel of 21 bp in the *trn*T-L spacer, as well as by a 4-bp indel in the *trn*L-F spacer. The B and D genomes share a 6-bp indel in the *trn*L-F spacer whereas the two aneuploid species, *A. praecox* and *A. palustris*, share a 2-bp indel in the *trn*L-F spacer.

The unrooted ML tree (Lscore 2820.28851) resolved three major groups that correspond to the A, B and D genome species (Fig. 2). The B and D genome species were sister groups, with *A. praecox* sister to both. Within the B genome (83%), *A. palustris* retained a basal position, followed by *A. ipaensis* (67%) and *A. cruziana* plus *A. batizocoi* (96%). The D genome clade included *A. glandulifera*, *A. williamsii*, and *A. benensis* with relatively low support (69%). Resolution within the A genome clade was not high. *Arachis kempff-mercadoi* was close to *A. herzogii*, and the two were separated from other A genome species by 64% support. Appearing in a polytomy in this clade are *A. duranensis*, *A. kuhlmannii*, *A. hoehnei* and *A. hypogaea*.

As expected, all hybrids including reciprocals, grouped with their maternal parent. For example, *A. duranensis* (30073) × *A. batizocoi* (9484) hybrid grouped with its maternal parent, *A. duranensis*, and the reciprocal *A. batizocoi* (9484) × *A. duranensis* (30073) grouped with *A. batizocoi* (Fig. 2).

Discussion

An assumption in genome analysis studies is that species with similar genomes (homologous genomes) show high rates of crossability as opposed to those with dissimilar genomic backgrounds. Consequently, the resulting hybrids are expected to have regular meiosis and high male (pollen) fertility. On the basis of these assumptions alone, one can conclude that *A. benensis*, *A. palustris* and *A. praecox* have significantly divergent genomes from either of the A and B genome tester species because none of the harvested pods in the hybrids contained viable seeds. *Arachis benensis* is a 20-chromosome species, and a preliminary karyotype (data not presented) showed that it has a small pair of chromosomes similar to those in the A genome species. Hybrids between this species and the A- and B-genome tester species resulted in pods without seeds. *Arachis benensis* has a unique AFLP fingerprint and appeared as a distinct lineage (Fig. 1), supporting the genomic divergence observed in the cytogenetic study. Although the *trn*T-F analysis resolved the affinity of *A. benensis* with the D genome group, the support for this relationship is low (bootstrap 69% and 56%). Further investigations are needed to determine whether *A. benensis* is a hybrid species or a taxon with a

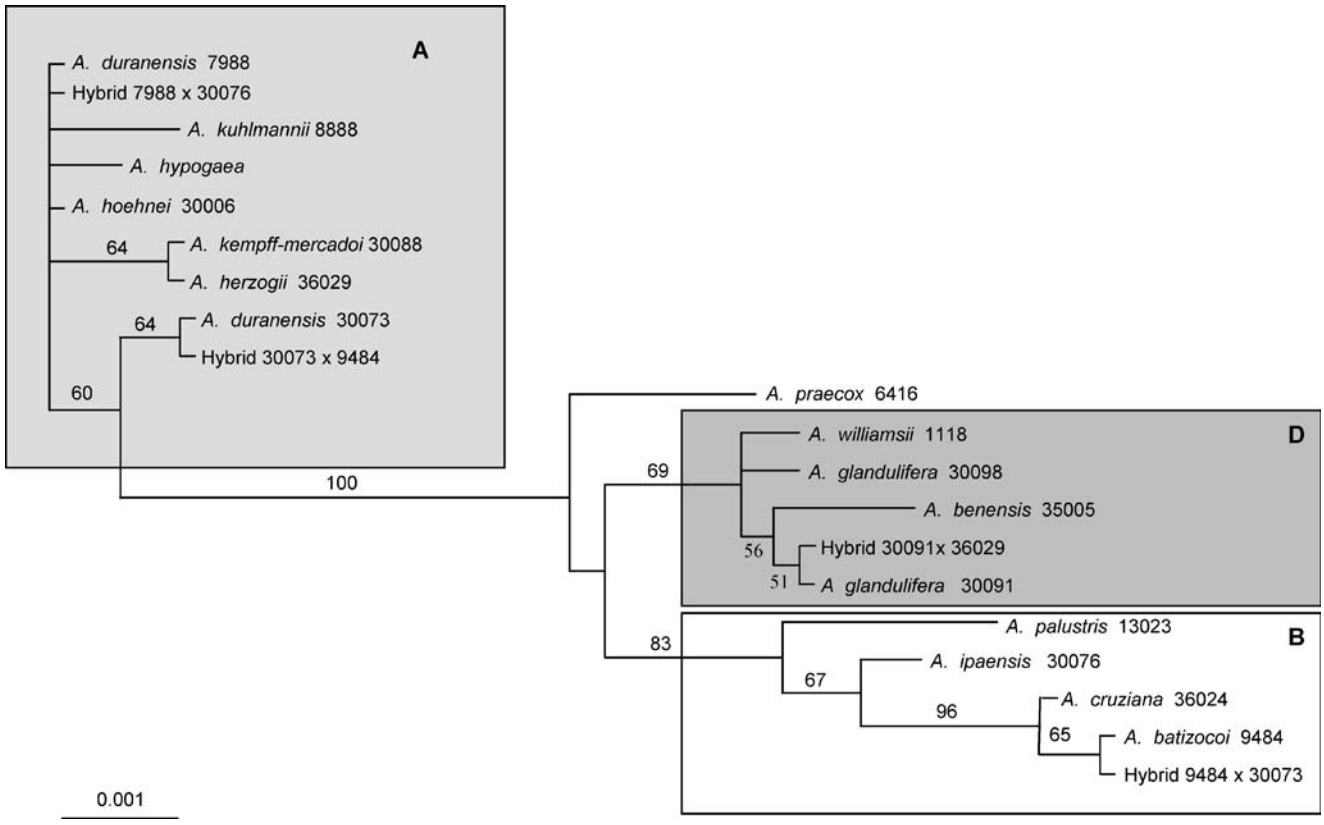

Fig. 2 Maximum likelihood tree (GTR + G+ I, Lscore 2820.28851) showing the relationships of the different genomes (A, B, and D) based on the *trn*T-F region. Bootstrap support >50% is given along the branches

unique genome. On the other hand, *A. palustris* and *A. praecox* are aneuploids with a chromosome number of $2n=2x=18$ (Lavia 1998) as opposed to $2n=2x=20$ in the two tester species. This difference in chromosome number presumably resulted in genetic imbalances in the hybrid embryo and led to abortion. Consequently, cytogenetic results are inconclusive about the genome type and the affinities of these two species. Lavia (1998) concluded that these two species lack the small pair of chromosomes characteristic of the A genome species. Therefore, they could either be aneuploids derived from an A genome species by losing that pair of chromosomes, or they were possibly derived from one of the other genome species. The close grouping in the *trn*T-F phylogram of *A. praecox* and *A. palustris* with the B and D clades indicates a close association with these genomes. Although the pattern of genome evolution cannot be assumed because the ML tree is not rooted in this study, the long branch between them and the A genome lineage underscores high genomic divergence (Fig. 2).

Cytogenetic evidence indicates that *A. duranensis* has relatively more genome similarity to *A. kempff-mercadoi* (10 II and 72% male fertility) than to either *A. kuhlmannii* or *A. herzogii* (Table 3). Both *trn*T-F and AFLP results demonstrate the strong genetic relationships between *A. duranensis* and these three species. Information from cytogenetics, AFLP and *trn*T-F (substitutions and indels) demonstrate that *A. herzogii*, *A. kempff-mercadoi* and *A. kuhlmannii* possess the A genome.

The placement of *A. hoehnei* in the A genome clade in the *trn*T-F phylogram (Fig. 2) is supported by a 4–bp indel shared between *A. hoehnei* and the A genome species (Table 4). This species was linked to the aneuploid species cluster in the AFLP dendrogram, which in turn was close to the A genome species cluster (including *A. duranensis*, *A. herzogii*, *A. kempff-mercadoi*, and *A. kuhlmannii*). The AFLP phenetic association is congruent with the phylogenetic relationship depicted by the *trn*T-F phylogram. In contrast, hybrids between *A. hoehnei* and *A. batizocoi* displayed a mean chromosome association of 8 II compared with 5.6 with the A genome tester species, pointing to relatively higher chromosomal homology with the B genome species. As a result, the genomic affinity of *A. hoehnei* remains unresolved.

The genomes of *A. ipaensis* and *A. williamsii* appear intermediate between the A and B genome tester species (Tables 2, 3). Chromosome associations in hybrids between *A. ipaensis* and both the A and B genome tester species showed similar frequencies of bivalents (Table 3). Previous analyses using molecular markers of *Arachis* species (Kochert et al. 1991; Tallury et al. 2001; Milla et al. 2005) suggested that *A. ipaensis* is genetically different from the B genome species, *A. batizocoi*. In contrast, the *trn*T-F study showed *A. ipaensis* nested

within the B genome lineage and that it lacks the indels characteristic of the A genome species (Table 4). On the other hand, *A. williamsii* appeared in a polytomy at the base of the D-genome species in the in *trn*T-F tree, but the 69% bootstrap support is not conclusive of its D-genome affinity. Cytogenetic data show mean chromosome pairing of 7.3 II in hybrids between *A. williamsii* and the A genome tester species and 5.8 II in the hybrid with the B genome tester. These two values do not point to a B genome affinity for *A. williamsii*. In contrast, AFLP analysis grouped *A. ipaensis* and *A. williamsii* into a tight cluster that is closely associated with *A. hypogaea* (Fig. 1). Both species have been proposed as genomic donors to *A. hypogaea* (Kochert et al. 1996; Krapovickas, personal communication (2001); Tallury et al. 2001). Nevertheless, additional crosses of these two species with other B and D genome species accessions may clarify the genomic grouping of these two species.

Among the *A. batizocoi* (B-genome tester) crosses, *A. batizocoi* × *A. cruziana* produced pods with normal seeds, whereas *A. duranensis* × *A. cruziana* resulted in aborted seeds. Therefore, cytogenetic information provides strong evidence for *A. cruziana* as a B-genome species. The *trn*T-F data unequivocally resolved the placement of *A. cruziana* within the B genome group and points to a strong phylogenetic relationship with *A. batizocoi* (BS 96%). In the AFLP dendrogram, *A. cruziana* was linked to the *A. batizocoi* and *A. glandulifera* cluster. These results are in agreement with the crossability-based suggestion of Simpson et al. (1997) that *A. cruziana* contains the B genome.

The phylogenetic analysis of the *trn*T-F region in the hybrids and their parental species demonstrates the maternal inheritance of the plastid genome in *Arachis*. Because the chloroplast genome is maternally inherited, the *trn*T-F phylogram represents the maternal tree. As a result, none of the B or D genome species evaluated appears to be a potential maternal donor of *A. hypogaea*, as the latter was nested in the A genome clade. The *trn*T-F results clearly demonstrate closer affinities between the B and D genome species and their relative distinctness from the A genome species. This is well demonstrated by the long branch between the A genome lineage and the node leading to the B and D genome lineages. This relationship is supported not only by the ML analysis but also by the presence of structural changes (discussed above) in the *trn*T-F region that are synapomorphic to the B and D genomes (Table 4). The strong grouping revealed by the AFLP markers (SC 0.92; Fig. 1) of *A. batizocoi* (B genome) with *A. glandulifera* (D genome) provides further support for the close genetic relationships between these genomes. However, the branch length separating the three-genome groups shows that the node separating the A genome lineage from the others is long, and those leading to the B and D genome lineages are relatively short. In addition, the bootstrap values supporting these two lineages (83 and 69%) are moderate to low whereas the A lineage is supported by 100%. Therefore, although the A genome group appears well diverged, the B and D genomes are less differentiated from each other.

Although crossability, male fertility and chromosomal associations provided clarification of section *Arachis* genome affinities, there remain several problem areas because of failure to obtain hybrids or sterility. This may be due to a genetic control mechanism that regulates crossability as was reported in wheat (Riley and Chapman 1967), a possibility that is worth investigating in *Arachis*. In triticale and rye crosses, the maternal genotype, which influences the rate of crossability and possible meiotic disturbances occurring in particular genotypic backgrounds, can reduce the hybridization frequencies (Guedes-Pinto et al. 2001). Unfortunately, reciprocal crosses were not made in the present study. Further, sterility of hybrids could be due to genetic factors such as cryptic structural differentiation of genomes rather than to a total genome divergence.

Unlike cytogenetic approaches, molecular information from AFLP and *trn*T-F is not influenced by genetic control mechanisms. The *trn*T-F phylogram represents a maternal relationship whereas the AFLP dendrogram represents the biparental pattern. The mode of inheritance, the type of characters scored, and the phenetic analysis of the AFLP data versus the phylogenetic analysis of the *trn*T-F characters may account for some of the discrepancies between the two approaches.

When combined, the three approaches used in this study have provided valuable insight into the genomic identities and species relationships in section *Arachis*. For instance, the data supported earlier observations of the A genome identities of *A. kempff-mercadoi*, *A. herzogii*, and *A. kuhlmannii*, and the B genomic composition of *A. cruziana*. The *trn*T-F data clearly show that the B and D genomes are more related to each other than to the A genome. The *trn*T-F plastid sequences also provided genome-specific markers and fingerprints of section *Arachis* species. The appearance of the two aneuploid species *A. praecox* and *A. palustris* sister to the B and D genome lineage in the *trn*T-F analysis needs further investigation to determine the exact pattern of chromosome and genome evolution in section *Arachis*.

Acknowledgements This work was partially supported by the North Carolina Agricultural Research Service, National Peanut Foundation, and the Virginia Agricultural Experiment Station.

References

Bianchi-Hall C, Keys RD, Stalker HT (1993) Diversity of seed storage proteins in wild peanuts (*Arachis* species). Plant Syst Evol 186:1–15

Borsch T, Hilu KW, Quandt D, Wilde V, Neinhuis C, Barthlott W (2003) Non-coding plastid *trn*T- *trn*F sequences reveal a well resolved phylogeny of basal angiosperms. J Evol Biol 16: 558–576

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

Grieshammer U, Wynne JC (1990) Isozyme variability in mature seeds of U. S. peanut cultivars and collections. Peanut Sci 17:72–75

Guedes-Pinto H, Lima-Brito J, Ribeiro-Carvalho C, and Gustafson JP (2001) Genetic control of crossability of triticale with rye. Plant Breed 120:27–33

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

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

Holbrook CC, Stalker HT (2003) Peanut breeding and genetic resources. Plant Breed Rev 22:297–356

Hopkins MS, Casa AM, Wang T, Mitchell SE, Dean RE, Kochert GD, Kresovich S (1999) Discovery and characterization of polymorphic simple sequence repeats (SSRs) in cultivated peanut (*Arachis hypogaea* L.). Crop Sci 39:1243–1247

Johnson ES, Miklas PN, Stavely JR, Martinez-Cruzado JC (1995) Coupling- and repulsion-phase RAPDs for marker-assisted selection of PI 181996 rust resistance in common bean. Theor Appl Genet 90:659–664

Kelchner SA (2000) The evolution of non-coding chloroplast DNA and its application in plant systematics. Ann Mo Bot Gard 87:482–498

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

Kochert G, Stalker HT, Ginenes M, Galgaro L, Moore K (1996) RFLP and cytogenetic evidence for the progenitor species of allotetraploid cultivated peanut, *Arachis hypogaea* (Leguminosae). Am J Bot 83:1282–1291

Krapovickas A, Gregory WC (1994) Taxonomy of the genus *Arachis* (Leguminosae). Bonplandia 8:1–186

Lanham PG, Fennell S, Moss JP, Powell W (1992) Detection of polymorphic loci in *Arachis* germplasm using random amplified polymorphic DNA. Genome 35:885–889

Lavia GI (1998) Karyotypes of *Arachis palustris* and *A. praecox* (section *Arachis*), two species with basic chromosome number $x = 9$. Cytologia 63:177–181

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

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

Müller JJ (2000) QuickAlign 1.5.5. (Computer program). Available from http://www.nees.uni-bonn.de/downloads/QA

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

Raina SN, Mukai Y (1999) Genomic *in situ* hybridization in *Arachis* (Fabaceae) identifies the diploid wild progenitors of cultivated (*A. hypogaea*) and related wild (*A. monticola*) peanut species. Plant Sys Evol 214:251–262

Riley R, Chapman V (1967) The inheritance in wheat of crossability with rye. Genet Res 9:259–267

Rohlf FJ (1998) NTSYS-pc, Numerical taxonomy and multivariate analysis system. Version 2.02i.Exeter software, Setauket, NY

Simpson CE, Burow MD, Patterson AH, Valls JFM, Williams DE, Vargas IG, Starr JL (1997) Cross-compatibilty data of additional B - genome *Arachis* accessions. Proc Am Peanut Res Educ Soc 29:56 (abstract)

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, Krishnan SS, Mengesha MH, Ramaiah CD (1991) Phylogenetic relationships in section *Arachis* based on seed protein profiles. Theor Appl Genet 82:593–597

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

Stalker HT (1990) A morphological appraisal of wild species in section *Arachis* of peanuts. Peanut Sci 17:117–122

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

Stalker HT, Dalmacio RD (1981) Chromosomes of *Arachis* species, section *Arachis* (Leguminosae). J Hered 72:403–408

Stalker HT, Moss JP (1987) Speciation, cytogenetics, and utilization of *Arachis* species. Adv Agron 41:1–40

Stalker HT, Dhesi JS, Parry D, Hahn JH (1991) Cytological and interfertility relationships of *Arachis* section *Arachis*. Am J Bot 78:238–246

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

Swafford DL (2002) PAUP: Phylogenetic Analysis Using Parsimony, Version 4.0b6 (computer program). Sinauer, Sunderland, MS

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding chloroplast regions. Plant Mol Biol 17:1105–1109

Tallury SP, Milla SR, Copeland SC, Stalker HT (2001) Genome donors of *Arachis hypogaea* L. Proc Am Peanut Res Educ Soc 33:60 (abstract)