

Development of an RFLP linkage map in diploid peanut species

T. Halward^{1,*}, H. T. Stalker², G. Kochert¹

¹ Department of Botany, University of Georgia, Athens, GA 30602-7271, USA

² Department of Crop Science, North Carolina State University, Raleigh, NC 27695-7629, USA

Received: 14 July 1992 / Accepted: 11 February 1993

Abstract. An RFLP linkage map of peanut has been developed for use in genetic studies and breeding programs aimed at improving the cultivated species (*Arachis hypogaea* L.). An F₂ population derived from the interspecific hybridization of two related diploid species in the section *Arachis* (*A. stenosperma* × *A. cardenasii*) was used to construct the map. Both random genomic and cDNA clones were used to develop the framework of the map. In addition, three cDNA clones representing genes coding for enzymes involved in the lipid biosynthesis pathway have been mapped in peanut. Of the 100 genomic and 300 cDNA clones evaluated, 15 and 190, respectively, revealed polymorphisms among the parents of our mapping population. Unfortunately, a large number of these produced complex banding patterns that could not be mapped. Of the 132 markers analyzed for segregation, 117 are distributed among 11 linkage groups, while 15 have not yet been associated with any other marker. A total map distance of approximately 1063 cM has been covered to-date.

Key words: *Arachis hypogaea* – *Arachis* sp. – Restriction fragment length polymorphism – Linkage map

Introduction

The genus *Arachis*, a member of the family *Leguminosae*, is native to South America, with central Brazil believed to be the center of origin (Krapovickas

1969). Although a complete taxonomic treatise has not been published to-date, Gregory et al. (1973) divided the genus into seven sections based on morphology, geographic distribution, and cross compatibilities. Cultivated peanut (*Arachis hypogaea* L.), an allotetraploid ($2n = 4x = 40$), has been assigned to the section *Arachis* which also contains the proposed allotetraploid progenitor *A. monticola*, Krap. et Rig. In addition, 20–25 diploid species ($2n = 2x = 20$) have been described in the section *Arachis*. Abundant germplasm resources of both the cultivated species and related wild species are available to peanut breeders; however, most peanut breeding programs have traditionally relied on the crossing of elite breeding lines for developing improved cultivars. As a result, the germplasm base of domesticated peanut is relatively narrow. Although considerable levels of morphological variability have been observed among the germplasm resources of cultivated peanut (ICRISAT 1982; Wynne and Halward 1989), very little genetic polymorphism has been detected with molecular markers among cultivars and exotic germplasm lines within *A. hypogaea* (Grieshammer and Wynne 1990; Halward et al. 1991, 1992). In addition *A. monticola*, the proposed allotetraploid progenitor of cultivated peanut, invariably produced identical banding patterns to *A. hypogaea* when evaluated with molecular markers (Halward et al. 1991; Kochert et al. 1991). Conversely, abundant polymorphism has been observed among related wild species of peanut in the section *Arachis* with isozymes (Stalker et al. 1990), RFLPs (Kochert et al. 1991), and RAPD markers (Halward et al. 1992).

In light of the relatively narrow germplasm base of cultivated peanut, a renewed emphasis should be placed on the evaluation and utilization of related wild species in order to enhance the genetic variability

Communicated by G. E. Hart

* Present address: Department of Agronomy, Colorado State University, Fort Collins, CO 80523, USA

Correspondence to: T. Halward

available for the development of improved cultivars. Extensive screening of wild *Arachis* species has revealed these genetic resources to be valuable as sources of disease and insect resistance (Moss 1980; ICRISAT 1982; Subrahmanyam et al. 1982), tolerance to environmental stresses (ICRISAT 1982), and variation for protein and oil quality (Young et al. 1973; Amaya et al. 1977; Cherry et al. 1977). The exploitation of these valuable genetic resources would be greatly enhanced through the use of molecular markers to tag and follow the introgression of chromosome segments containing desirable traits from the wild species into cultivated peanut, and through the development of a genetic linkage map in peanut to expedite the location and transfer of these chromosomal regions.

At present, there is no conventional genetic linkage map available for peanut. In addition, no linkage groups or genetic maps based on molecular marker evaluations have been reported in the literature. However, a collaborative effort is currently underway to develop a genetic linkage map based on a combination of molecular and conventional markers. As molecular markers accumulate in peanut, the linkage map will become a useful tool for applied breeding programs, as breeders will be able to tag and follow the introgression of specific chromosome segments containing desirable genes from related wild species into breeding lines of cultivated peanut.

Materials and methods

Preparation of DNA libraries

A random genomic DNA library was constructed using the peanut cultivar 'GK-7' (*Arachis hypogaea*, ssp. *hypogaea*) as the genomic DNA source. The library was constructed by cloning gel-isolated *Pst*I fragments (1.0–2.0 kb in length) into pUC8 plasmids and transforming the recombinant plasmids into *E. coli* DH5 α cells according to the procedures of Sambrook et al. (1989). Recombinant clones were selected on IPTG-Xgal plates, and plasmids were isolated by a miniprep procedure (Wilimzig, 1985). Recombinant plasmids were digested with *Pst*I and subjected to electrophoresis on 0.8% agarose gels. The molecular weight of each insert was determined by comparison to molecular weight standards, and blots were prepared on nylon filters (Gene Screen Plus, Du Pont) following the method of Southern (1975). Total peanut DNA was labeled with ³²P-dCTP by nick-translation (Rigby et al. 1977) and hybridized to the filters to detect inserts containing repeated sequences. Cotton chloroplast DNA (compliments of G. Galau, Department of Botany, University of Georgia) was used to screen the library for members containing chloroplast DNA inserts. Those inserts that showed no signal with either chloroplast or total peanut DNA were assumed to represent low-copy number nuclear sequences and were selected for RFLP analysis.

Two separate cDNA libraries were prepared, one from root tissue and one from shoot tissue, using the Stratagene ZAP-cDNA GigapackII Gold Cloning Kit, according to the manufacturer's instructions. Approximately 200 seeds of the peanut cultivar 'GK-7' were germinated in moist paper towels to yield a

total of 10–15 g each of root and shoot tissue for the two libraries. Poly A⁺ RNA was isolated independently from young shoots and young roots according to the procedures described in Hong et al. (1989) and used to construct the two cDNA libraries.

Generation of mapping population

An F₂ population derived from the interspecific hybridization between two diploid wild *Arachis* species (*A. stenosperma* × *A. cardenasii*) was chosen for construction of the linkage map. The population, consisting of 87 F₂ individuals, was developed at North Carolina State University as part of an ongoing interspecific hybridization program.

Detection of RFLPs and linkage analysis

Genomic DNA was isolated from the parents, the F₁ and F₂s of the mapping population using a crude nuclear preparation as described in Kochert et al. (1991). The parents (*A. stenosperma* and *A. cardenasii*) were screened for clones that revealed polymorphisms between them. Survey filters consisted of genomic DNA from each of the parents digested with seven restriction enzymes (*Bam*HI, *Dra*I, *Eco*RI, *Eco*RV, *Hae*III, *Hind*III, *Rsa*I). Clones that were found to reveal polymorphisms between the parents were used to evaluate the F₂ mapping population for segregation.

To analyze segregation, the F₂ progeny were scored as either 'A' [homozygous like parent A (*A. stenosperma*)], 'B' [homozygous like parent B (*A. cardenasii*)], or 'H' (heterozygous). Chi-square analyses were conducted to determine goodness-of-fit of the segregation data to the expected 1:2:1 Mendelian ratio. The segregation data obtained were analyzed using the MAPMAKER computer package. The locus arrangements and map distances for each linkage group were determined based on the output from the MAPMAKER program using the Kosambi mapping function and the constraints of a minimum LOD score of 3.0 and a maximum recombination fraction (theta) of 0.25. RFLP loci detected by different probes were assigned different numbers using the following convention: Xuga.pg (or cr, or cs) ###, where pg indicates a random genomic clone, while cr and cs indicate cDNA clones from the root and shoot libraries, respectively. Clones designated cc ### indicate cDNA clones which represent genes in the lipid biosynthesis pathway. These clones were graciously contributed by Bert Abbott, Clemson University. Multiple-segregating loci detected by a single probe were assigned the same number followed by a letter (a, b, etc.) to indicate each duplicated locus.

Results

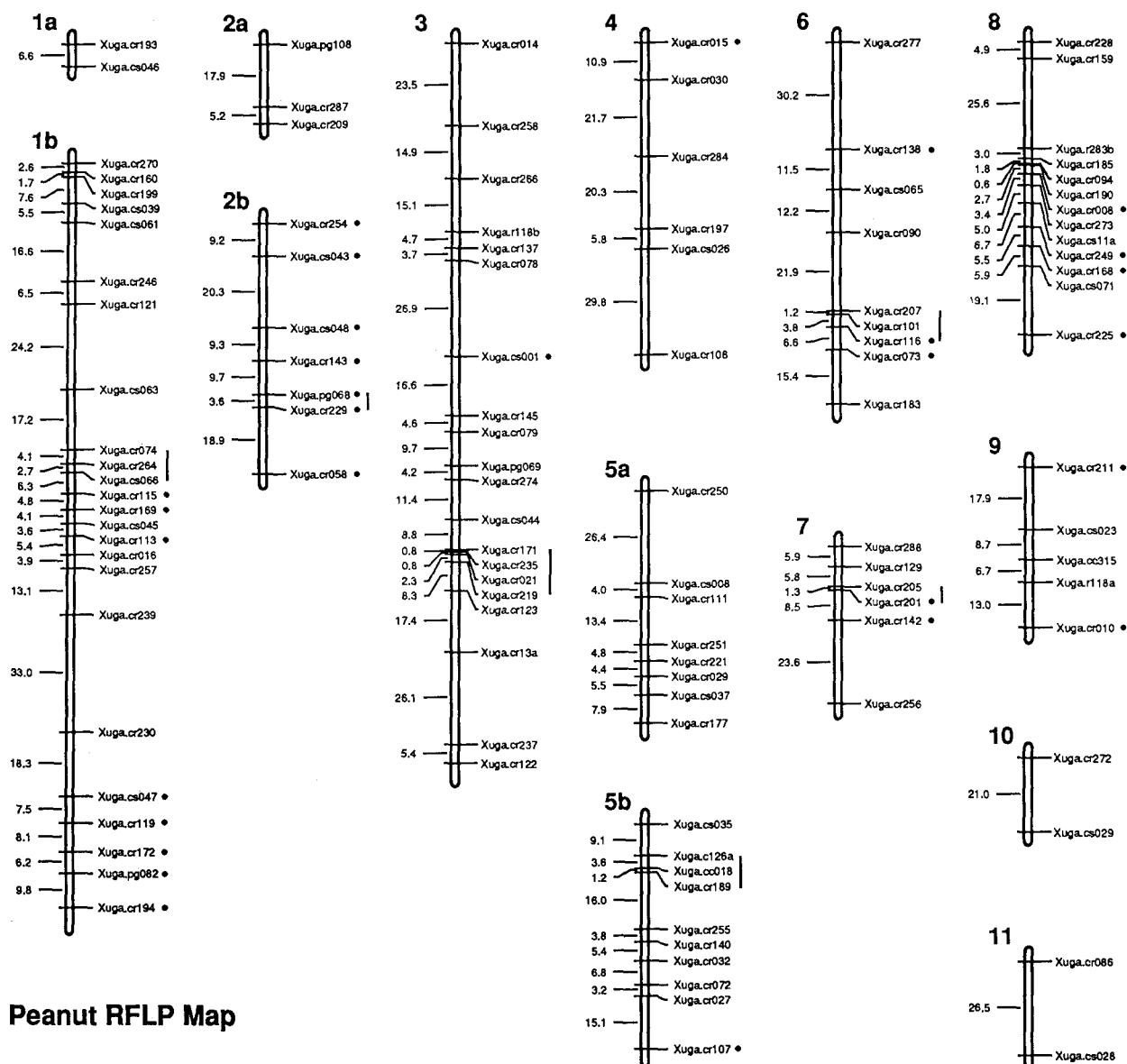
Description of linkage map and clones evaluated

A total of 100 random genomic clones and 300 cDNA clones have been evaluated to-date for polymorphism in our mapping population. Of these, 15 (15%) of the genomic clones and 190 (63%) of the cDNA clones were polymorphic with one or more enzyme. We concentrated our mapping efforts using cDNA clones, as a greater percentage were found to reveal polymorphisms in our mapping population when compared to the random genomic peanut clones evaluated. Of the 205 polymorphic markers, 132 were analyzed for segre-

gation, while the other 73 produced complex banding patterns which could not be mapped.

Chi-square analyses were conducted to determine goodness-of-fit of the segregation data to the expected 1:2:1 Mendelian ratio. Thirty three of the one hundred and thirty two scored loci showed deviation from the expected ratio ($P < 0.05$). Most of these loci were found to have an excess of one or the other parental type,

while four had an excess number of heterozygotes. Those loci with distorted segregation patterns are indicated on the map by a dot (.) to the right of the marker designation (Fig. 1). Loci with distorted segregation ratios were generally scattered among linkage groups and unlinked loci, with two notable exceptions: a small cluster of markers in linkage group 1b showed an excess of bands of the parental 'B' type; while linkage



Peanut RFLP Map

Fig. 1. Peanut RFLP linkage map developed in an F_2 population derived from the interspecific hybridization of two wild diploid species in the section *Arachis* (*A. stenosperma* \times *A. cardenasii*). A total of 117 markers are arranged in 11 linkage groups based on the output from the MAPMAKER computer package using the following constraints: minimum LOD of 3.0, maximum recombination fraction of 0.25. RFLP markers are listed on the right of each linkage group using the nomenclature described in Materials and methods ('pg' indicates a random genomic clone; 'cr' and 'cs' indicate cDNA clones from the root and shoot libraries, respectively; and 'cc' indicates a cDNA clone obtained from Clemson University). Map distances, in centiMorgans, are listed on the left. Closely-linked markers with ambiguous orders, as described in Results, are denoted on the map with a line next to them. Loci with distorted segregation ratios from the expected 1:2:1 are denoted by a (.) to the right of the marker designation

Table 1. Polymorphic markers that have not yet been associated with any other marker in an established linkage group

Random genomic clones	cDNA, roots	cDNA, shoots	cDNA, Clemson
Xuga. pg070. Xuga. pg111.	Xuga. cr013b Xuga. cr061 Xuga. cr081 Xuga. cr105 Xuga. cr167 Xuga. cr242 Xuga. cr243 Xuga. cr269 Xuga. cr283a	Xuga. cs011b Xuga. cs023 Xuga. cs116	Xuga. cc126b

group 2b consists entirely of markers with an excess of bands of the parental 'A' type and may, therefore, reflect an artificial linkage arrangement. As additional loci are added to the map and the level of marker saturation increases, this potentially artificial clustering should be resolved.

When F_2 segregation data were analyzed using the MAPMAKER computer package, 117 segregating loci were distributed among 11 linkage groups (Fig. 1), while 15 have not yet been associated with any other marker (Table 1). Although groups 1a, 1b; 2a, 2b; and 5a, 5b were 239-, 316-, and 562-times, respectively, more likely to be linked together as single linkage groups (i.e., 1, 2, and 5) than to be unlinked using the MAPMAKER 'linked' command, we chose not to accept anything below a LOD of 3.0 ($1000 \times$ more likely) as true linkage. Therefore, we have kept these groups separate until additional markers are added to the map which confirm their status. Two of the linkage groups (10 and 11) are quite small, containing only two markers each. These may eventually link up with one of the larger linkage groups as additional markers accumulate. Ultimately, the map should consist of ten linkage groups corresponding to the haploid chromosome number of diploid *Arachis* species.

A large number of probes (44%) hybridized to more than one polymorphic locus, indicating the presence of a number of duplicated loci within the peanut genome. Unfortunately, the majority of those clones detecting duplicated loci produced complex banding patterns which were difficult to score and could not be mapped with accuracy. As such, we chose not to apply these markers to the linkage map. Five clones which hybridized to multiple loci were scored and mapped. However, multiple loci detected with a single clone could be arranged into linkage groups for only one of these (Xuga.cr118a,b), with 118a mapping to linkage group 9 and 118b mapping to linkage group 3. For the other four clones detecting multiple loci which were applied to the map, only one of the polymorphic loci mapped to an established linkage group (Fig. 1), whereas the other polymorphic loci detected by those clones were distributed among the unlinked markers (Table 1).

The majority of probes (64%) detected polymorphisms with more than one restriction enzyme, suggesting that many of the RFLP markers identified in peanut resulted from DNA rearrangements. Similar results have been observed in a number of other species including soybean (Apuya et al. 1988) and rice (McCouch et al. 1988). Of the cDNA clones evaluated, 38 produced ladders with one or more restriction enzyme suggesting the presence of a number of tandem repeat sequences within the peanut genome. Unfortunately, no variability was observed between the parents of our mapping population for these clones; therefore, they could not be mapped.

In some instances, the order of closely-linked markers within a given linkage group could not be resolved with absolute certainty. Using the 'compare' command of multipoint analysis, the preferred orders of two or three closely-linked markers often had LOD scores differing by a value of 0.1 to 1.0 indicating that any of the orders were equally likely. Closely-linked markers with an ambiguous order have been denoted on the map with a line out to the side (Fig. 1). In addition to random genomic clones and cDNA clones from our root and shoot libraries, three cDNA clones representing genes in the lipid biosynthesis pathway have been mapped in peanut. To-date, a total map distance of approximately 1063 cM of the diploid peanut genome has been covered.

Discussion

Mapping population

A number of factors were taken into consideration when selecting an appropriate population for developing an RFLP map in peanut. The most important factor being insufficient variability observed for RFLPs (Halward et al. 1991) or RAPD markers (Halward et al. 1992) within *A. hypogaea* germplasm to allow construction of a genetic linkage map directly in cultivated peanut. In addition, although most traits in cultivated peanut follow a diploid pattern of inherit-

ance, many appear to be under the influence of duplicate loci (Wynne and Halward 1989). As a result, segregation analysis of RFLP or RAPD markers is more difficult in the allotetraploid cultivated species than in the related diploid species. Finally, a complete series of aneuploids, often used in the construction of RFLP maps (Beckmann and Soller 1986), is not available in peanut. For these reasons, the diploid species of *Arachis* were considered the best choice for a mapping population. We were fortunate to have available several F_2 populations derived from interspecific crosses between various diploid species of *Arachis*. One such population (*A. stenosperma* \times *A. cardenasii*), developed at North Carolina State University, was chosen for map construction.

RFLP maps that have been developed using populations derived from interspecific crosses between a cultivated species and a related wild species have proven useful for cultivar improvement in such crops as tomato, *Lycopersicon esculentum* (Miller and Tanksley 1990) and soybean *Glycine max* (Keim et al. 1990). Although the peanut RFLP map was developed using a population derived from a cross between two wild *Arachis* species, it should also be useful for cultivar improvement programs, especially when applied to following the introgression of chromosome segments from wild species into cultivated peanut. Normal segregation has been reported in populations derived from interspecific hybridization between diploid *Arachis* species with the same genomic complement; only hybridizations involving *A. batizocoi* behave abnormally during meiosis (Ressler and Gregory 1979; Stalker and Wynne 1979; Singh and Moss 1984). This suggests that the arrangement of genes on the chromosomes is homosequential among species of the section *Arachis*. As such, a linkage map developed in one population should be applicable when transferred to other peanut populations. However, additional populations involving crosses among different *Arachis* species should be evaluated using the same set of clones employed to develop the present map in order to confirm linkage relationships across species. We have recently initiated such a study using an F_2 population derived from the interspecific cross between *A. duranensis* and *A. diogeni*.

The map

Several enzymes were used in the initial screening of the parents for polymorphism, as previous results (unpublished data) indicated that a given clone often revealed polymorphism with some enzymes but not others. By simultaneously screening the parents with a number of probe-enzyme combinations, a greater number of polymorphic clones could be detected. Clones that were found to reveal polymorphisms between the parents were used to evaluate the F_2 mapping population

for segregation. Initially, only those clones that revealed polymorphism at a single locus were used in construction of the RFLP linkage map. Single-locus probes are less ambiguous, especially when transferring a linkage map to additional populations. Now that a framework is in place, clones detecting multiple loci can be added with greater confidence to the developing map. The high level of duplication observed within the peanut genome, with 44% of the probes hybridizing to more than one polymorphic locus, may have important implications regarding genome evolution in the genus. Therefore, it would be desirable to add such clones to the map to determine if duplicated regions are clustered within particular linkage groups or randomly distributed throughout the genome.

Applications of the linkage map to peanut breeding and genetics

Currently, no classical genetic linkage map exists for peanut, and no trait has yet been mapped to a specific peanut chromosome (Stalker 1991). Without the availability of a genetic map, it has not been possible to utilize molecular markers in peanut breeding, or to combine molecular and conventional genetic techniques in peanut improvement programs. As the peanut RFLP map presented here continues to develop, a number of potential advantages will be realized in peanut improvement programs. Our understanding of genetic segregation and linkage relationships will be greatly enhanced as molecular markers accumulate in peanut. The chromosomal location and distribution of duplicated regions within the peanut genome will allow for a better understanding of the underlying basis for the complex patterns of segregation often observed in peanut, particularly among the progeny of interspecific crosses (Wynne and Halward 1989; Wynne and Coffelt 1980).

Given the relatively narrow germplasm base available in *A. hypogaea*, and the abundance of polymorphism observed for both morphological traits and molecular markers in wild species of *Arachis*, increased emphasis should be placed on the utilization of wild relatives in peanut improvement programs. This is perhaps the area for which a comprehensive RFLP map will be of most benefit to peanut breeders. Only limited success has been achieved with a variety of methods for attempting to introgress germplasm from wild species into cultivated peanut as part of a conventional peanut breeding program (Singh 1986a, b; Simpson 1991; Stalker 1992). The development of a genetic linkage map in peanut will greatly enhance the ability of breeders to tag and follow the introgression of specific chromosome segments linked to desirable traits from wild species into breeding lines of cultivated peanut. The peanut RFLP map will be useful for monitor-

ing introgression following the mapping of valuable agronomic traits in wild species.

We are currently collaborating with a number of researchers in other laboratories to screen our mapping population for several traits of interest including resistance to rootknot nematode, *Meloidogyne araniaria* (with Jim Noe, Dept. of Plant Pathology, Univ. of Georgia), white mold (*Sclerotium rolfsii*) (with Mel Fuller, Dept. of Plant Pathology, Univ. of Georgia), and both early (*Cercospora arachidicola* Hori) and late [*Cercosporidium personatum* (Berk. and Curt.) Deighton] leafspot (with Kim Moore, AgraTech Research). The goal is to eventually add these traits to the peanut map such that RFLPs can be used as markers to follow their introgression from wild *Arachis* species into breeding lines of cultivated peanut through backcross breeding programs.

References

- Amaya F, Young CT, Hammons RO, Martin G (1977) The tryptophan content of the U.S. commercial and some South American wild genotypes of the genus *Arachis*: A survey. *Oleagineux* 32:225–229
- Apuya N, Frazier BL, Keim P, Roth EJ, Lark KG (1988) Restriction length polymorphisms as genetic markers in soybean, *Glycine max* (L.) Merr. *Theor Appl Genet* 75:889–901
- Beckmann JS, Soller M (1986) Restriction fragment length polymorphisms and genetic improvement of agricultural species. *Euphytica* 35:111–124
- Cherry JP (1977) Potential sources of peanut seed proteins and oil in the genus *Arachis*. *J Agric Food Chem* 25:186–193
- Gregory WC, Gregory MP, Krapovickas A, Smith BW, Yarbrough JA (1973) Structure and genetic resources of peanuts. In: *Peanuts – culture and uses*. Am Peanut Res Educ Assoc, Stillwater, Oklahoma, pp 47–133
- Griesshammer U, Wynne JC (1990) Mendelian and non-Mendelian inheritance of three isozymes in peanut (*Arachis hypogaea* L.). *Peanut Sci* 17:101–105
- Halward TM, Stalker HT, LaRue EA, Kochert G (1991) Genetic variation detectable with molecular markers 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. *Plant Mol Biol* 18:315–325
- Hong JC, Nagao RT, Key JL (1987) Characterization and sequence analysis of a developmentally regulated putative cell wall protein gene isolated from soybean. *J Biol Chem* 262:8367–8376
- ICRISAT (1982) Annual Report. Patancheru, A.P., India
- Keim P, Dier BW, Olson TC, Shoemaker RC (1990) RFLP mapping in soybean: association between marker loci and variation in quantitative traits. *Genetics* 126:735–742
- Kochert G, Halward TM, Branch WD, Simpson CE (1991) RFLP variability in peanut cultivars and wild species. *Theor Appl Genet* 81:565–570
- Krapovickas A (1969) Evolution of the genus *Arachis*. Seminario Avanzado de Genetica Agricola para America Latina, Maracay, Venezuela SAGA/B(d):1–4
- McCouch SR, Kochert G, Yu ZG, Wang ZY, Khush GS, Coffman WR, Tanksley SD (1988) Molecular mapping of rice chromosomes. *Theor Appl Genet* 76:815–829
- Miller JC, Tanksley SD (1990) RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. *Theor Appl Genet* 80:437–448
- Moss JP (1980) Wild species in the improvement of groundnuts. In: Summerfield RJ, Bunting AH (eds) *Advances in legume science vol. 1*. Royal Botanic Gardens, Kew
- Ressler PM, Gregory WC (1979) A cytological study of three diploid species of the genus *Arachis*. *J Hered* 70:13–16
- Rigby P, Dieckmann M, Rhodes C, Berg P (1977) Labeling deoxyribonucleic acid to high specific activity in vitro by nick-translation with DNA polymerase I. *J Mol Biol* 113:237–251
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
- Simpson CE (1991) Global collaborations find and conserve the irreplaceable genetic resources of wild peanut in South America. *Diversity* 7:59–61
- Singh AK (1986a) Utilization of wild relatives in the genetic improvement of *Arachis hypogaea* L. 7. Autotetraploid production and prospects in interspecific breeding. *Theor Appl Genet* 72:164–169
- Singh AK (1986b) Utilization of wild relatives in the genetic improvement of *Arachis hypogaea* L. 8. Synthetic amphidiploids and their importance in interspecific breeding. *Theor Appl Genet* 72:433–439
- Singh AK, Moss JP (1984) Utilization of wild relatives in genetic improvement of *Arachis hypogaea* L. 5. Genome analysis in section *Arachis* and its implication in gene transfer. *Theor Appl Genet* 68:355–364
- Southern EM (1975) Detection of specific enzyme sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 98:503–517
- Stalker HT (1991) A morphological appraisal of wild species in section *Arachis* of peanuts. *Peanut Sci* 17:117–122
- Stalker HT (1992) Utilizing *Arachis* germplasm resources. In: *Proc 2nd Int. Wkshp on Groundnut*, ICRISAT, Patancheru, A.P., India, pp 24–29 (in press)
- Stalker HT, Jones TM, Murphy JP (1990) Isozyme variability among *Arachis* species. *Proc Am Peanut Res Educ Soc* 22:50
- Stalker HT, Wynne JC (1979) Cytology of interspecific hybrids in section *Arachis* of peanuts. *Peanut Sci* 6:110–114
- Subrahmanyam P, McDonald D, Gibbons RW, Nigam SN, Nevill DJ (1982) Resistance to rust and late leafspot diseases in some genotypes of *Arachis hypogaea*. *Peanut Sci* 9:9–14
- Wilimzig R (1985) LiCl method for plasmid minipreps. *Trends Genet* 1:158
- Wynne JC, Coffelt TA (1982) Genetics of *Arachis hypogaea*. In: Pate HE, Young CT (eds) *Peanut science and technology*. Am Peanut Res Educ Soc, Yoakum, Texas, pp 50–94
- Wynne JC, Halward TM (1989) Cytogenetics and genetics of *Arachis*. *Crit Rev Plant Sci* 8:189–220
- Young CT, Waller GR, Hammons RO (1973) Variations in total amino-acid content of peanut meal. *J Am Oil Chem Soc* 50:521–523