Chapter 6

Genomics and Genetic Enhancement of Peanut

Andrew H. Paterson,[a] *H. Thomas Stalker,*[b] *Maria Gallo-Meagher,*[c] *Mark D. Burow,*[d] *Sangam L. Dwivedi,*[e] *Jonathan H. Crouch,*[e] and *Emma S. Mace*[e]

[a]University of Georgia, 111 Riverbend Road, Athens, GA 30602;

[b]Department of Crop Sciences, Box 7620, North Carolina State University, Raleigh, NC;

[c]Department of Agronomy, PO Box 110300, University of Florida, Gainesville, FL;

[d]Texas A&M University Research and Extension Center, Lubbock, TX 29403; and

[e]International Crops Research Institute for the Semi Arid Tropics (ICRISAT), Hyderabad, India

Introduction

Cultivated peanut, also known as groundnut (*Arachis hypogaea* L.), is grown on 25.5 million hectares between latitudes 40° N and 40° S with a total global production of 35 million tons (Mt)(1). Peanuts originated in South America; however, the vast majority of peanut is produced in Asia (66.8%: 23.4 Mt) and Africa (24.6%: 8.6 Mt). The remaining 8.6% (3 Mt) comes from North America, the Caribbean, Europe, and Oceania.

Peanut is among the top five oilseeds grown in the world with 65% of world production used by India, China, and the United States. The cake remaining after peanut oil extraction can be used in human food or incorporated into animal feeds (2). Wild *Arachis* species (*A. pintoi* Krapov. and W.C. Gregory and *A. glabrata* Benth.) have nutritive values higher than those of most commercially important tropical forage legumes, and have been used for pasture improvement in North America, Central America, South America, and Australia (3). The greater adaptability of rhizomatous perennial peanut (*A. glabrata*) to the tropical environment and its high yield when harvested for hay give it the potential of becoming one of the most important forages in the tropics (4).

In the early 1900s, Dr. George Washington Carver of the Tuskegee Institute in Alabama developed more than 300 uses for cultivated peanut. Largely because of his research efforts, peanut became the second-largest row crop in the Southern United States (after cotton). During the 1990s, the U.S. peanut crop had an average value of over $1 billion, ranking it as our second most important seed legume, after soybean. Per capita, Americans consume more than 3 kg of peanuts and peanut products per year, ranking it among our most popular foods.

Peanut offers numerous human health benefits. It contains mostly unsaturated fat, which has been shown to lower LDL-cholesterol levels in the blood, and resveratrol, which leads to improved cardiovascular health. Peanut is also a good source of

folic acid, which helps prevent neural tube defects, and it contains nearly half of the 13 essential vitamins and 35% of the essential minerals. Because of its high nutritional value, the peanut is being widely investigated as a key food source for astronauts during extended space missions.

Unique Botanical Features of Peanut and Their Impact on Crop Productivity

Peanut has one of the most intriguing reproductive systems in the plant kingdom. Peanuts produce flowers above the ground, but their fruits and seeds develop below the ground. The temporal and spatial separation of fertilization and fruit development is probably an adaptation to drought in the Amazon basin that constitutes peanut's center of diversity (5).

The reproductive mechanism of peanut offers new avenues of inquiry that should lead to a better understanding of plant reproductive and developmental biology. Following fertilization, peanut ovules advance to the 8- to 16-cell stage and then become quiescent for up to three weeks. Discovering what controls this discrete cessation of ovule development may lead to future strategies that could be employed in other crops for the production of seedless fruits. Subsequent stages of peanut reproduction, such as the meristematic activity of the ovary, the gravitropic response of the developing peg, and the dark and Ca^{2+} requirements for pod and seed development, also reflect meaningful changes to the standard reproductive program for legumes and most higher plants. They may provide insight into mechanisms at work in other plant biological processes which involve meristematic activity, gravitropism, light regulation, and Ca^{2+} regulation.

Peanut's reproductive mechanism also has major implications for crop productivity and genetic diversity. A primary repercussion of its reproductive strategy is a low success rate for completion of reproduction. After pollination, not all flowers produce pegs, and not all pegs grow sufficiently long to reach the soil. Consequently, a high percentage of pegs yield nothing. Depending on the variety, the reproductive success rate ranges from 11% (6) to 68% (7). Although a number of factors ultimately affect seed set, it is clear from breeding work that peanut reproductive success can be genetically improved. There is a high likelihood that peanut yields could be significantly enhanced if one were able to improve peanut's ability to form pegs that reach the soil and complete development. Additionally, a better understanding of peanut reproduction could substantially increase the extremely limited genetic diversity in cultivated peanut by improving the very low success rate of interspecific hybridization (8).

Genetic Diversity and Chromosomal Relationships of Cultivated and Wild *Arachis*

Unlike many other natural polyploid species for which multiple polyploidization events have been identified and that permit exchange of moderate levels of genetic

variability among polyploids (9, 10), cultivated peanut, *A. hypogaea* is a tetraploid ($2n = 4x = 40$) believed to have originated recently from a single hybridization event (11). Archaeological evidence from excavations in Peru place the origin of *A. hypogaea* at least 3,500 years ago (12). Cytological studies of *A. hypogaea* observed 20 chromosome bivalents at meiosis in 88 to 98% of cells; the exceptions were rare univalents, trivalents, and quadrivalents, which suggested limited homoeologous pairing between the A and B genomes that compose the tetraploid (13,14).

Despite the paucity of DNA-level variability among cultivated genotypes, a wealth of diversity exists in other *Arachis* species. In section *Arachis* of the genus there are 27 species representing three genomes, A, B, and D. Additionally, 42 congeners have been identified in the other eight sections of the genus (22). Abundant DNA marker polymorphisms have been detected between wild species in section *Arachis* (16,23,24) and also within some species such as *A. duranensis* Krapov. and W.C. Gregory (25). This large amount of variation supports the hypothesis that *A. hypogaea* likely originated from a single hybridization event followed by chromosome doubling, with very little subsequent introgression from related diploid species (26). Kochert *et al.* (11) concluded that *A. duranensis* and *A. ipaensis* Krapov. and W. C. Gregory are the most likely diploid progenitors of the cultivated peanut, and *A. duranensis* was probably the female parent.

Only one other species in section *Arachis*, *A. monticola* Krapov. and Rigoni, is tetraploid and readily crossable with *A. hypogaea*, but it is indistinguishable from *A. hypogaea* based on DNA markers (15,16). Thus, *A. hypogaea* and *A. monticola* can be considered the same biological species.

Conventional Genetic Improvement of Peanut

Over 276 peanut cultivars were released between 1920 and 2000 for cultivation in various countries in Asia, Africa, and the Americas. Each has specific adaptation to its respective region of production and cropping system (27–29). A yearly genetic gain of nearly 15 kg/ha (per hectare) has been reported for large-seeded Virginia type cultivars released from the 1950s to the 1970s reaching a maximum yield of 4.6 t/ha in the United States (30). However, since the 1970s there has been increased emphasis on improving pest resistance and quality traits. The result of this emphasis is that the yield potential of cultivars released since that time has not surpassed those of the highest-yielding cultivars released during the 1970s.

Many breeding programs, including the one at ICRISAT, made substantial progress toward improving the yield potential of cultivars adapted to rain-fed or irrigated high-input situations. Some cultivars have produced pod yield of over 9 t/ha in Zimbabwe (31) and China (32). However, there is a wide gap between realized yields at farm level (world average yield of 1.4 t/ha) when compared to the highest average yields in China (3.1 t/ha) and the United States (3.4 t/ha) (1). There is therefore a need to incorporate multiple resistances to biotic and abiotic stresses into improved genetic backgrounds, even if it requires some sacrifice in yield potential, as

experienced by U.S. peanut breeders, to narrow the gap between realized and potential yields (33).

Plant introductions have had increasing importance to peanut improvement, particularly in Virginia-type cultivars which have increased resistance to tomato spotted wilt virus; but relatively few accessions are in the pedigrees of cultivars grown in the U S. (28) Within the cultivated and wild species gene pools of peanut there is a large amount of genetic variation for many agronomically important traits (34,35). Continued evaluation of the U S germplasm collection, consisting of more than 8,000 plant introductions (36), and the collection maintained by the ICRISAT with more than 14,000 accessions (37), will be critical for genetic improvement of cultivated peanut

Recently developed cultivars have reduced vegetative mass, shorter main stem length, and greater reproductive allocation (partition more of their daily assimilate to fruit) than those developed previously, as predicted by Duncan *et al* (38). High yield *in recently released* cultivars appears to be related more to total flower production than to reproductive efficiency. Therefore, future increases in seed yield might be accomplished by developing cultivars with a combination of high reproductive efficiency, harvest index, and total flower count (39).

Consistent with its predominantly inbreeding nature, the most commonly used breeding methods in peanut are pedigree selection, bulk-pedigree selection, and single-seed descent Backcross breeding has not been used extensively because most of the economically important traits in peanut are quantitatively inherited (40,41). Breeders often make single crosses to generate variability However, with increased emphasis on multiple resistance breeding, emphasis is now focused on complex crosses followed by intercrossing of segregants to bring the desired improvement into breeding populations. While selection for resistance to insect pests and diseases is practiced in early generations, selection for yield and yield component traits is delayed until later generations Recurrent selection has also been used for continued genetic enhancement in peanut (42,43). As the application of molecular markers becomes routine, it is likely that greater emphasis will be *placed on* backcross breeding, *particularly* to assist the rapid introgression of disease resistances from wild *Arachis* species and exotic *A. hypogaea* germplasm.

Genetic Mapping of Peanut

The paucity of DNA polymorphism in cultivated peanut (see above) posed a considerable obstacle to genetic *mapping. The use of a synthetic amphidiploid*, TxAG-6 (44,45) made possible the generation of the first molecular map representing the entire tetraploid genome of peanut, as well as study of the transmission genetics of chromatin from a synthetic amphidiploid into cultivated peanut. TxAG-6 was developed through the cross [*A batizocoi* Krapov. and W. C. Gregory × (*A cardenasii* Krapov. and W. C. Gregory × *A diogoi* Hochne)]^{4x}, where *A. cardenasii* and *A. diogoi* are A-genome diploids, and *A. batizocoi* was once considered to be a B-genome diploid ancestor (46–48). This cross has been used to introduce root-knot nematode

resistance into cultivated peanut (49,50) and also harbors many other potentially valuable attributes.

The discovery of a high level of polymorphism between the cultivar Florunner and the parents of TxAG-6 by RAPD analysis (49) was followed by RFLP surveys that reported 83% polymorphism on a per-band basis between Florunner and TxAG-6 (8). Using 78 peanut BC_1 lines generated using TxAG-6 (45,49) as the donor parent and the *A. hypogaea* cultivar Florunner as recurrent parent, 220 cDNA probes were used to map 370 RFLP loci onto 23 linkage groups. A total of 917 bands was observed, for an average of 4.1 bands per probe. A mean 1.68 loci per probe were mapped. The total length of the tetraploid map, 2210 cM (centimorgans), was slightly greater than twice the length (1063 cM) of a map previously reported from a cross between two A-genome diploid species (51). Based on a small number of DNA markers mapped in both crosses, there is a high degree of colinearity between the diploid and tetraploid chromosomes (8).

As expected of a recently formed polyploid, many DNA markers mapped to two different linkage groups, and patterns of duplication shed light on the identities of homoeologous chromosomes. Eighty-nine probes produced markers on both homoeologous linkage group pair members. The 23 linkage groups were comprised of nine pairs of homoeologous linkage groups, one trio representing a homoeologous chromosome pair, one fragment consisting of two markers, and one linkage group that was possibly an artifact. Given that cultivated peanut is a disomic polyploid ($2n = 4x = 40$), 20 linkage groups were expected. By comparison of "alloalleles" in tetraploid peanut to the alleles present in the respective diploid progenitors of TxAG-6, the subgenomic affinities of the linkage groups were readily discerned. There was evidence for large structural differences in the LG1/LG11 pair, and weaker evidence for possible rearrangements in the LG7/LG17, LG4/LG14, and LG5/LG15 pairs. Other homoeologous linkage group pairs appeared to be collinear to the degree of resolution afforded by this experiment. This is in agreement with cytological observations (52,53) of incomplete pairing of chromosomes in *A. cardenasii* and *A. batizocoi* diploid F_1 hybrids resulting in five or more univalents per pollen mother cell.

A consensus molecular map of the peanut tetraploid genome based on SSR and AFLP markers is also being developed (Mace, ICRISAT, unpublished data). However, due to the lack of polymorphism observed between the parental genotypes, to date only approximately 70 SSR and AFLP markers have been mapped to 17 linkage groups covering approximately 420 cM. Future research efforts will aim to map the new SSR markers on mapping populations derived from selected diverse parental genotypes such as TxAG-6.

Marker-Assisted Selection

Marker-assisted selection (MAS) offers great promise for improving the efficiency of conventional plant breeding. Molecular markers are especially advantageous for traits where conventional phenotypic selection is difficult, expensive, or lacks accuracy or precision. This includes resistance to certain pathogens and insect pests plus

tolerance to abiotic stresses, quality parameters, and complex agronomic traits with low heritabilities. The essential requirements for developing marker-assisted selection systems are (i) availability of germplasm with substantially contrasting phenotypes for the traits of interest, (ii) highly accurate and precise screening techniques for phenotyping mapping populations for the trait of interest, (iii) identification of flanking marker(s) closely associated with the loci of interest and the flanking regions on either side, and (iv) simple robust DNA marker technology to facilitate rapid and cost-effective screening of large populations.

In peanut, a few examples of traits that may justify the cost and time required to develop and apply DNA markers include early leafspot (*Cercospora arachidicola* Hori.), late leafspot (*Cercosporidium personatum* Berk. et Curt.) Deighton, nematodes (*Meloidogyne* spp.), leafminer, and *Spodoptera*, for which there are only low to moderate levels of resistance (or tolerance) available in cultivated peanut but high levels available in wild species. Traits associated with seed quality—as measured by oleic/linoleic (O/L) ratio, for which the higher the ratio the better the shelf-life of the peanut products—and drought tolerance (specific leaf area, total transpiration, water use efficiency, and partitioning), which are difficult to measure in large segregating generations and substantially influenced by genotype-by-environment interaction, may also benefit from MAS (54). MAS holds great promise for the improvement of such characters if markers can be identified for genes controlling components of these traits.

MAS within *A. hypogaea* is constrained by the low level of DNA polymorphism found in the cultivated gene pool. The RFLP-based tetraploid map developed by Burow *et al.* (8), based on an interspecific cross, is likely to be useful in terms of locating specific genes of interest in this interspecific cross and also provides valuable information about genome organization and evolution. However, the markers themselves are of less value in elite cultivated germplasm, in which very little polymorphism is found. Recent advances in the development of PCR-based marker protocols have opened new possibilities in the study of complex traits in crop plants. The hybridization-based co-dominant markers (RFLP) and PCR-based dominant markers (RAPD and AFLP) in many crops have been superseded by co-dominant PCR-based markers (SSR). The AFLP assay has been used frequently in diversity and mapping studies in many crop plants. However, efforts to convert AFLP markers into simple co-dominant PCR markers are laborious, expensive, and time consuming, and above all have met with only mixed success, particularly in complex polyploids (55).

Although MAS has been little used within *A. hypogaea*, even with the limitations afforded by present technologies it has much potential use for introgressing genes from closely related *Arachis* species into the cultivated genome. Garcia *et al.* (56) reported introgression of genes from *A. cardenasii* (an A-genome species) into *A. hypogaea* in 10 of 11 linkage groups on the diploid RFLP map developed by Halward *et al.* (51). Garcia *et al.* (57) then used RAPD and SCAR technologies to map two dominant genes conferring resistance to the nematodes (*M. arenaria* (Neal) Chitwood, Race 1) where a tetraploid breeding line derived an *A. hypogaea* × *A. cardenasii* cross was analyzed. Burow *et al.* (49) reported RFLP markers linked to *M. arenaria* resistance in another interspecific cross involving the species *A. hypogaea*,

A. batizocoi, *A. cardenasii*, and *A. diogoi*. Additional linkages of RAPD markers have been found for components of early leafspot and corn rootworm (*Diabrotica undecimpunctata howardi* Barber) resistances in (cultivar NC 7 × NC GP WS 1)^{4x} interspecific crosses. In crosses restricted to the *A. hypogaea* parents NC 7 and PI 109839, linkages of RAPD primers with genes conferring resistance to leafhopper (*Empoasca fabae* Harris), Cylindrocladium black rot and several components of early leafspot were found (58). Among diploids, several AFLP markers were linked to tomato spotted wilt virus resistance (59). Although the above linkages of resistance genes to different molecular markers may prove useful for selecting breeding lines with desirable traits, there have been only a few successes in peanut for utilizing these materials for cultivar development.

Clearly, there is a need to explore new assays with greater power to reveal polymorphisms in peanut, such as single nucleotide polymorphisms (SNP). Recent developments in SNP technology indicate that in the near future there may be additional options available for rapid identification of large numbers of polymorphic markers (60). SNPs comprise the largest set of sequence variants in most organisms (61, 62). SNPs are biallelic markers but occur very frequently within the genome, their mutation rate is low, and they are capable of high throughput genotyping and are often linked to genes (63). For example, a map containing 1.42 million SNPs distributed throughout the human genome has been constructed with an average density of one SNP every 1.9 kb (64). SNPs have also been reported in many crop plants and show much promise for MAS, but to date there have been no substantial explorations of the peanut genome.

Tools and Strategies for Physical Mapping of Peanut

The cultivated peanut genome is estimated to contain about 2.9 pg DNA per haploid genome equivalent (65). The A and B genomes are thought to contribute approximately equal amounts of genetic information to cultivated peanut, based on (1c) estimates of 1.34 pg DNA for *A. duranensis* and 1.42 pg DNA for *A. ipaensis*, respectively (66). Cot analysis of the genomic DNA of *A. duranensis* reveals genomic composition of 27% highly repetitive DNA, 37% middle-repetitive DNA, and 36% low-copy DNA (67). These parameters of size and repetitiveness are similar to those for the human genome, suggesting that a comprehensive physical map of peanut will require a much greater effort than was necessary for botanical models such as *Arabidopsis* or *Oryza*.

The cornerstone of a physical mapping strategy for peanut will be a comprehensive arrayed library of large-insert DNA clones. Early progress toward a bacterial artificial chromosome (BAC) library for the tetraploid peanut cultivar Florunner (68) has been advanced to include about 180,000 clones (roughly six genome-equivalent coverage). Hybridization-based anchoring of genetically-mapped DNA markers, together with fingerprinting and alignment of BACs to one another will help assemble tools suitable for many applications, including positional cloning anywhere in the peanut genome. Utilization of synteny information from other legumes or botanical

models (see Chapter 9) may help identify gene probes that can be targeted to gaps in the peanut physical map and accelerate closure. Selected BAC end sequencing may help to extend contigs, although the abundance of repetitive DNA in peanut suggests that only 36% of peanut BAC ends will be locus-specific.

A fundamental complication in peanut physical mapping will be the presence of two duplicated copies of most genes, one from each of the two diploid progenitors. By hybridization-based approaches, it is likely to be very difficult to distinguish which BACs contain true homologs of a particular probe sequence, and which contain homoeologs, or corresponding sequences from the alternate subgenome. Methods have been established (69) to resolve the subgenome specificity of individual BACs in the tetraploid libraries. An alternative approach would be to produce additional BAC libraries for diploid progenitors of peanut, which would reduce the magnitude of the problem.

Tools and Strategies for Functional Genomics of Peanut

Fundamental to analysis of gene function in an organism is a detailed picture of the transcriptome (the population of DNA sequences that encode mRNA products). For peanut, this information is almost completely lacking but urgently needed. Other cultivated legumes such as *Glycine*, *Medicago*, and *Phaseolus* enjoy extensive resources of 359,130, 194,501, and 21,931 ESTs, respectively, representing diverse sets of tissues and in some cases multiple genotypes. However, as of 15 August 2003, peanut had only 1,540 public EST (gene) sequences. Generation of peanut ESTs for many diverse tissues is a key first step in associating peanut genes with their respective functions. In view of the basic and applied importance of peanut's unusual reproductive system, a particularly high priority is generation of ESTs for reproductive tissues. The addition of a peanut EST resource to the extensive resources for *Glycine* and *Medicago*, and emerging resources for *Phaseolus*, will empower a host of investigations of features that distinguish these four major legume lineages from one another, and indeed, which distinguish the legumes from other well-studied rosids such as the *Brassicales* (crucifers, including *Arabidopsis*) and *Malvales* (cotton), as well as asterids (lettuce and sunflower), core eudicots (ice plant and sugarbeet), and monocots. EST resources are a crucial first step toward most functional genomic approaches, in particular large-scale analysis of the timing and levels of expression of peanut genes.

Prospects for Sequencing the Peanut Genome

The large size and high fraction of repetitive DNA make the complete sequencing of the peanut genome a daunting task, one of similar cost and complexity to sequencing of the human genome. A clear first step is sequencing large populations of expressed sequences (ESTs) to rapidly gain information about many of the regularly expressed genes in peanut. The cost will be a tiny fraction of resources needed to

sequence the entire genome. While this primary goal will provide much information, it will also leave important questions unanswered, such as the identities of key regulatory genes that are expressed only briefly and/or at low levels, and therefore tend not to be sampled by the EST approach. Further, EST sequencing will not provide information about regulatory sequences that determine timing and levels of gene expression, and which may be key to the molecular basis of many morphological changes that distinguish related plant taxa.

Once basic information such as EST resources for peanut is developed, a logical next step may be the exploration of the peanut genome by a recently described approach called Cot-based cloning and sequencing (CBCS) (70,71). CBCS is based on fractionation of total genomic DNA into components that are comprised of populations of DNA elements that occur in the genome in similar copy numbers. For example, the "highly repetitive" component typically contains many retrotransposons, centromeric repeats, and other highly-abundant elements often thought of as "junk DNA." The "middle-repetitive" component may contain lower-copy repetitive elements, ribosomal DNA, and some large multigene families. The "low-copy" component contains most genes. By first fractionating a genome into such components, then cloning each component and sequencing into the clone library to a depth that provides appropriate coverage of the sequence complexity of the respective components, the population of nonredundant DNA sequences in a genome can be obtained at substantially less time and cost than by approaches that are presently in use (71). CBCS offers a particularly appealing approach by which one might efficiently obtain not only the gene-rich sequences in peanut, but also quickly identify the repetitive DNA families that complicate peanut genomics to mitigate the impact of these elements on sequence assembly and annotation.

Summary

Despite its major economic importance both in the United States and internationally, peanut is badly underexplored at the genomic level. This deficiency hinders geneticists from being able to provide intrinsic low-cost and environmentally benign solutions to many challenges that increase the cost and risk of peanut production, and cause peanut to fall short of consumer needs and desires. Urgent needs include the development of large numbers of user-friendly genetic mapping tools; sequencing of substantial populations (100,000 or more) of expressed sequences from diverse tissues and genotypes; the assembly of a genetically anchored physical map and its alignment to the emerging sequences of small-genome legumes such as *Medicago*; and sampling the gene-rich regions to quantify the additional information that may be gained by further sequencing of peanut genomic DNA.

Acknowledgments

We thank many colleagues for valuable contributions, and the Texas Higher Education Coordinating Board, USDA National Research Initiative, Goldkist Inc., and the Georgia, Florida, North Carolina, and Texas Agricultural Experiment Stations for funding.

References

1. FAO, *FAO Production Year Book 55*:118–119 (2001).
2. Savage, G.P., and J.I. Keenan, The Composition and Nutritive Value of Groundnut Kernels, in *The Groundnut Crop: A Scientific Basis for Improvement*, edited by J. Smartt, Chapman and Hall, London, 1994, pp. 173–213.
3. Kerridge, P.C., and B. Hardy, *Biology and Agronomy of Forage* Arachis, CIAT, Cali, Colombia, 1994, No. 240, pp. 206.
4. Ruiz, T.M., R. Ramos-Santana, and A. Sotomayor-Rios, Dry Matter Yield of Rhizoma Perennial Peanut (*Arachis glabrata*) Harvested at Six, Nine and 12 Weeks at Two Semiarid Sites, *J. Agric. Univ. P.R. 84*:115–131 (2000).
5. Sauer, J.D. *Arachis*—Peanut, in *Historical Geography of Crop Plants*, Lewis Publishers, Boca Raton, FL, 1993, pp. 80–83.
6. Gregory, W.C., B.W. Smith, and J.A. Yarbrough, Morphology, Genetics, and Breeding, in *The Peanut—The Unpredictable Legume*, National Fertilizer Assoc., Washington, D.C., 1951, pp. 28–88.
7. Pattee, H.E., H.T. Stalker, and F.G. Giesbrecht, Reproductive Efficiency in Reciprocal Crosses of *Arachis monticola* with *A. hypogaea* Subspecies, *Peanut Sci. 25*:7–12 (1998).
8. Burow, M.D., C.E. Simpson, J.L. Starr, and A. H. Paterson, Transmission Genetics of Chromatin From a Synthetic Amphiploid in Cultivated Peanut (*A. hypogaea* L.): Broadening the Gene Pool of a Monophyletic Polyploid Species, *Genetics 159*:823–837 (2001).
9. Soltis, D.E., and P.S. Soltis, Molecular Data and the Dynamic Nature of Polyploidy, *Crit. Rev. Plant Sci. 12*:243–273 (1993).
10. Soltis, P.S., and D.E. Soltis, The Role of Genetic and Genomic Attributes in the Success of Polyploids, *Proc. Natl. Acad. Sci. USA 97*:7051–7057 (2000).
11. Kochert, G., H.T. Stalker, M. Gimenes, L. Galgaro, C.R. Lopes, and K. Moore, RFLP and Cytological Evidence on the Origin and Evolution of Allotetraploid Domesticated Peanut *Arachis hypogaea* (Leguminosae), *Am. J. Bot. 83*:1282–1291 (1996).
12. Singh, A.K., and C.E. Simpson, Biosystematics and Genetic Resources, in *The Groundnut Crop: A Scientific Basis for Improvement*, edited by J. Smartt, Chapman and Hall, London, 1994, pp. 24–42.
13. Singh, A.K., and J.P. Moss, Utilization of Wild Relatives in Genetic Improvement of *Arachis hypogaea* L. 2 Chromosome Complements of Species of Section *Arachis*, *Theor. Appl. Genet. 61*:305–314 (1982).
14. Wynne, J.C., and T. Halward, Cytogenetics and Genetics of *Arachis*, *Crit. Rev. Plant Sci. 8*:189–220 (1989).
15. Kochert, G., T. Halward, W.D. Branch, and C.E. Simpson, RFLP Variability in Peanut (*Arachis hypogaea* L.) Cultivars and Wild Species, *Theor. Appl. Genet. 81*:565–570 (1991).
16. Halward, T.M., H.T. Stalker, E. LaRue, and G. Kochert, Genetic Variation Detectable with Molecular Markers among Unadapted Germplasm Resources of Cultivated Peanut and Related Wild Species, *Genome 34*:1013–1020 (1991).
17. Hopkins, M.S., A.M. Casa, T. Wang, S.E. Michell, R.E. Dean, G.D. Kochert, and S. Kresovich, Discovery and Characterization of Polymorphic Simple Sequence Repeats (SSRs) in Peanut, *Crop Sci. 39*:1243–1247 (1999).
18. Dwivedi, S.L., S. Gurtu, S. Chandra, W. Yuejin, and S.N. Nigam, Assessment of Genetic Diversity among Selected Groundnut Germplasm I: RAPD analysis, *Plant Breed. 120*:345-349 (2001).

19. He, G., and C. Prakash, Evaluation of Genetic Relationships among Botanical Varieties of Cultivated Peanut (*Arachis hypogaea* L.) Using Markers, *Genet. Resour. Crop Evol. 48*:347–352 (2001).
20. Cifarelli, R.A., M. Gallitelli, and F. Cellini, Random Amplified Hybridization Microsatellites (RAHM): Isolation of a New Class of Microsatellite-Containing Clones, *Nucleic Acids Res. 23*:3802–3803 (1995).
21. Fischer, D., and K. Bachmann, Microsatellite Enrichment in Organisms with Large Genomes (*Allium cepa* L.), *BioTechniques 24*:796–802 (1998).
22. Krapovikas, A., and W.C. Gregory, Taxonomy of Genus *Arachis* (Leguminosae), *Bonplandia 8*:1–186 (1994).
23. Paik-Ro, O.G., R.L. Smith, and D.A. Knauft, Restriction Fragment Length Polymorphism Evaluation of Six Peanut Species within the *Arachis* Section, *Theor. Appl. Genet. 84*:201–208 (1992).
24. Lanham, P.G., S. Fennell, J.P. Moss, and W. Powell, Detection of Polymorphic Loci in *Arachis* Germplasm Using Random Amplified Polymorphic DNAs, *Genome 35*:885–889 (1992).
25. Stalker, H.T., J.S. Dhesi, and G. Kochert, Genetic Diversity within the Species *Arachis duranensis* Krapov. and W. C. Gregory, a Possible Progenitor of Cultivated Peanut, *Genome 38*:1201–1212 (1995).
26. Young, N.D., N.F. Weeden, and G. Kochert, Genome Mapping in Legumes (Fam. Fabaceae), in *Genome Mapping in Plants*, edited by A.H. Paterson, Landes Bioscience Press, Austin, TX, 1996, pp. 211–227.
27. Godoy, I. J., and E. H. Giandana, Groundnut Production and Research in South America, in *Groundnut—A Global Perspective*, edited by S.N. Nigam, ICRISAT, Patancheru, India, 1992, pp. 77–85.
28. Isleib, T.G., C.C. Holbrook, and D.W. Gorbet, Use of Plant Introductions in Peanut Cultivar Development, *Peanut Sci. 28*:96–113 (2001).
29. Isleib, T.G., J.C. Wynne, and S.N. Nigam, Groundnut Breeding, in *The Groundnut Crop: A Scientific Basis for Improvement*, edited by J. Smartt, Chapman and Hall, London, 1994, pp. 552–623.
30. Mozingo, R.W., T.A. Coffelt, and J.C. Wynne, Genetic Improvement in Large-Seeded Virginia-Type Peanut Cultivars Since 1944, *Crop Sci. 27*:228–231 (1987).
31. Smartt, J., Makulu Red—A 'Green Revolution' Groundnut Variety? *Euphytica 27*:605–608 (1978).
32. Yanahao, S., and W. Caibin, Factors Contributing to High Yields of Groundnut in Shandong, China, *Int. Arachis Newsletter 8*:7–9 (1990).
33. Nigam, S.N., Some Strategic Issues in Breeding for High and Stable Yield in Groundnut in India, *J. Oilseeds Res. 17*:1–10 (2000).
34. Stalker, H.T., and C.E. Simpson, Genetic Resources in *Arachis*, in *Advances in Peanut Science*, edited by H.E. Pattee and H.T. Stalker, American Peanut Research Education Society, Stillwater, OK, 1995, pp. 14–53.
35. Holbrook, C.C., and H.T. Stalker, Peanut Breeding and Genetic Resources, in *Plant Breeding Reviews*, edited by J. Janick, John Wiley & Sons, NJ, 2003, Vol. 22, pp. 297–356.
36. Pittman, R.N., Progress and Status of the U.S. Peanut Collection, *Proc. Amer. Peanut Res. Educ. Soc. 32*:49 (2000).
37. Upadhyaya, H.D., M.E. Ferguson, and P.J. Bramel, Status of the *Arachis* Germplasm Collection at ICRISAT, *Peanut Sci. 28*:89–96 (2001).

38. Duncan, W.G., D.E. McCloud, R.L. McGraw, and K.J. Boote, Physiological Aspects of Peanut Yield Improvement, *Crop Sci. 18*:1015–1020 (1978).
39. Coffelt, T.A., M.L. Seaton, and S.W. VanScoyoc, Reproductive Efficiency of 14 Virginia-Type Peanut Cultivars, *Crop Sci. 29*:1217–1220 (1989).
40. Wynne, J.C., and W.C. Gregory, Peanut breeding, *Adv. Agron. 34*:39–71 (1981).
41. Knauft, D.A., and J.C. Wynne, Peanut Breeding and Genetics, *Adv. Agron. 55*:393–445 (1995).
42. Gurok, H.P., J.C. Wynne, and H.T. Stalker, Recurrent Selection within a Population from an Interspecific Peanut Cross, *Crop Sci. 26*:249—253 (1986).
43. Halward, T.M., J.C. Wynne, and H.T. Stalker, Recurrent Selection Progress in a Population Derived from an Interspecific Peanut Cross, *Euphytica 52*:9–84 (1991).
44. Simpson, C.E., Pathways for Introgression of Pest Resistance into *Arachis hypogaea* L., *Peanut Sci. 18*:22–26 (1991).
45. Simpson, C.E., S.C. Nelson, J.L. Starr, K.E. Woodard, and O.D. Smith, Registration of TxAG-6 and TxAG-7 Peanut Germplasm Lines. *Crop Sci. 33*:1418 (1993).
46. Husted, L., Cytological Studies of the Peanut *Arachis*. I. Chromosome Number and Morphology. *Cytologia 5*:109–117 (1933).
47. Husted, L., Cytological Studies of the Peanut *Arachis*. II. Chromosome Number, Number Morphology and Behavior and their Application to the Origin of Cultivated Forms, *Cytologia 7*:396–423 (1936).
48. Smartt, J., W.C. Gregory, and M.P. Gregory, The Genomes of *Arachis hypogaea*. 2. The Implications in Interspecific Breeding, *Euphytica 27*:677–680 (1978).
49. Burow, M.D., C.E. Simpson, A. H. Paterson, and J.L. Starr, Identification of Peanut (*Arachis hypogaea* L.) RAPD Markers Diagnostic of Root-Knot Nematode (*Meloidogyne arenaria* (Neal) Chitwood) Resistance. *Mol. Breed. 2*:369–379 (1996).
50. Simpson, C.E., and J.L. Starr, Registration of Coan Peanut, *Crop Sci. 41*:918 (2001).
51. Halward, T.M., H.T. Stalker, and G. Kochert, Development of an RFLP Map in Diploid Peanut Species, *Theor. Appl. Genet. 87*:379–384 (1993).
52. Smartt, J., W.C. Gregory, and M.P. Gregory, The Genomes of *Arachis hypogaea*. 1. Cytogenetic Studies of Putative Genome Donors, *Euphytica 27*:665–675 (1978).
53. Stalker, H.T., J.S. Dhesi, D.C. Parry, and J.H. Hahn, Cytological and Interfertility Relationships of *Arachis* Section *Arachis*, *Am. J. Bot. 78*:238–246 (1991).
54. Dwivedi, S. L., J.H. Crouch, S.N. Nigam, M.E. Ferguson, and A.H. Paterson, Molecular Breeding of Groundnut for Enhanced Productivity and Food Securities in the Semi-Arid Tropics: Opportunities and Challenges, *Adv. Agron. 80*:153–221 (2003) (in print).
55. Shan, X., T.K. Blake, and L.E. Talbert. Conversion of AFLP Markers to Sequence-specific PCR Markers in Barley and Wheat, *Theor. Appl. Genet. 98*:1072–1078 (1999).
56. Garcia, G.M., H.T. Stalker, and G. Kochert, Introgression Analysis of an Interspecific Hybrid Population in Peanuts (*Arachis hypogaea* L.) using RFLP and RAPD Markers, *Genome 38*:166–176 (1995).
57. Garcia, G.M., H.T. Stalker, E. Shroeder, and G. Kochert, Identification of RAPD, SCAR and RFLP Markers Tightly Linked to Nematode Resistance Genes Introgressed from *Arachis cardenasii* to *A. hypogaea*, *Genome 39*:836–845 (1996).
58. Stalker, H.T., and L.G. Mozingo, Molecular Markers of Arachis and Marker-Assisted Selection, *Peanut Sci. 28*:117–123 (2001).
59. Milla, S., Relationships and Utilization of *Arachis* Germplasm in Peanut Improvement, Ph.D. Thesis, North Carolina State University, Raleigh. 2003.

60. Kanazin, V., H. Talbert, D. See, P. DeCamp, E. Nevo, and T. Blake, Discovery and Assay of Single-Nucleotide Polymorphisms in Barley (*Hordeum vulgare*), *Plant Mol. Biol. 48*:529–537 (2002).
61. Kwok, P.Y., Q. Deng, H. Zakeri, S.L. Taylor, and D.A. Nickerson, Increasing the Information Content of STS-Based Genome Maps: Identifying Polymorphisms in the Mapped STSs, *Genomics 31*:123–126 (1996).
62. Kruglyak, L., The Use of a Genetic Map of Biallelic Markers in Linkage Studies, *Nature Genet. 17*:21–24 (1997).
63. Kwok, P.Y., and Z. Gu, SNP Libraries: Why and How Are We Building Them? *Mol. Medicine Today 12*:538–543 (1999).
64. The International SNP Map Working Group, A Map of Human Genome Sequence Variation Containing 1.42 Million Single Nucleotide Polymorphisms, *Nature 409*:928–933 (2001).
65. Arumunganathan, K., and E.D. Earle, Nuclear DNA Content of Some Important Plant Species, *Plant. Mol. Biol. Rptr. 9*:208–218 (1991).
66. Singh, K.P., S.N. Raina, and A.K. Singh, Variation in Chromosomal DNA Associated with the Evolution of *Arachis* Species, *Genome 39*:890–897 (1996).
67. Dhillon, S. S., A.V. Rake, and J.P. Miksche, Reassociation Kinetics and Cytophotometric Characterization of Peanut (*Arachis hypogaea* L.) DNA, *Plant Physiol. 65*:1121–1127 (1980).
68. Yuksel, B., D. Peterson, S. Lee, M. Burow, and A.H. Paterson, Progress Toward Construction and Characterization of a Peanut BAC library, *International Plant and Animal Genome Conference*, San Diego, CA, www.intlpag.org/pag/11/abstracts/P5f_P499_XI.html (2003).
69. Lin, Y.R., X. Draye, X. Qian, S. Ren, L. Zhu, J. Tomkins, R.A. Wing, Z. Li, and A.H. Paterson, Fine-Scale Mapping and Sequence-Ready Contig Assembly in Highly-Duplicated Genomes, using the BAC-RF Method, *Nucleic Acids Res. 2823*: www3.oup.co.uk/nar/methods/Volume_28/Issue_07/gnd023_gml.abs.html (2000).
70. Peterson, D.G., S.R. Schulze, E.B. Sciara, S.A. Lee, J.E. Bowers, A. Nagel, N. Jiang, D.C. Tibbitts, S.R. Wessler, and A.H. Paterson, Integration of Cot Analysis, DNA Cloning, and High-Throughput Sequencing Facilitates Genome Characterization and Gene Discovery, *Genome Res. 12*:795–807 (2002).
71. Peterson, D.G., S.R. Wessler, and A.H. Paterson, Efficient Capture of Sequence Complexity Using Cot-Based Cloning and Sequencing, *Trends Genet. 18*:547–550 (2002).