PEANUT SCIENCE

VOLUME 36 JANUARY–JUNE 2009 NUMBER 1

Genomics: An Evolving Science in Peanut

H.T. Stalker[1]*, A.K. Weissinger, S. Milla-Lewis, and C.C. Holbrook[2]

ABSTRACT

Genomic spcience offers new research tools to explore the function of genes and their effects on plants and animals. *Arachis hypogaea* is a polyploid species of relatively recent origin and molecular analyses with technologies available in the 1980s and 1990s resulted in little progress in the cultivated species because of apparent lack of molecular variation. Large numbers of polymorphisms existing in wild *Arachis* species led to evolutionary and gene introgression studies. High throughput genomic sequencing technologies have greatly expanded the possibilities for investigating gene function, but techniques are sufficiently expensive that most federal funding has been directed toward model species and 'major' crops. Peanut has lagged behind many other crops, but the number of researchers working on the species in the U.S. and internationally has greatly increased during recent years. In an effort to bring researchers who work with a number of legume crops together to discuss common goals, a national strategic planning workshop was held in 2001 which led to the U.S. Legume Crops Genomics Initiative. A second workshop was held in 2004 to develop a plan with specific objectives for cross-legume genomics research and to outline milestones for accomplishments. Specifically for peanut, a genomics strategic planning workshop was organized at Atlanta in 2004 by the American Peanut Council. A broad view of genomic science was adopted and goals were set by participants to include (a) improving the utility of genetic tools for peanut genomics research, (b) improving the efficacy of technology for gene manipulation in genomics, (c) developing a framework for assembling the peanut genetic blueprint, (d) improving knowledge of gene identification and regulation, and (e) providing bioinformatic management of peanut biological information. Teams of researchers, including molecular biologists, plant breeders, pathologists, and many other disciplines need to be developed to fully utilize the potential of genomics for peanut improvement.

Key Words: Genomics, plant breeding, crop improvement, molecular techniques, strategic planning.

Genomic science is an approach of investigating and understanding the highly complex structures and processes that make up a phenotype. The goal of genomics is to integrate knowledge about how the genome is organized, regulated, and interacted to create structures, products and activities. Genomic science takes advantage of large-scale and high throughput analyses to unravel the genetic make-up of an organism. Techniques to perform these tasks are continuing to be developed very rapidly, with the human genome project leading the way for other animal and plant systems. In plants, Arabidopsis has served as the model system, with rice (*Oryza sativa* L.), wheat (*Triticum aestivum* L. Thell.), maize (*Zea mays* L.), *Medicago trancatula* Gaertner, and a few other species leading the effort to develop methodologies and to sequence their genomes.

[1]Professor, Professor, and Research Assistant Professor, respectively, Department of Crop Science, Box 7620, North Carolina State Univ., Raleigh, NC 27695.

[2]Research Geneticist, U. S. Department of Agriculture, Agricultural Research Service, PO Box 748, Tifton, GA 31793.

*Corresponding author: Tom_Stalker@ncsu.edu

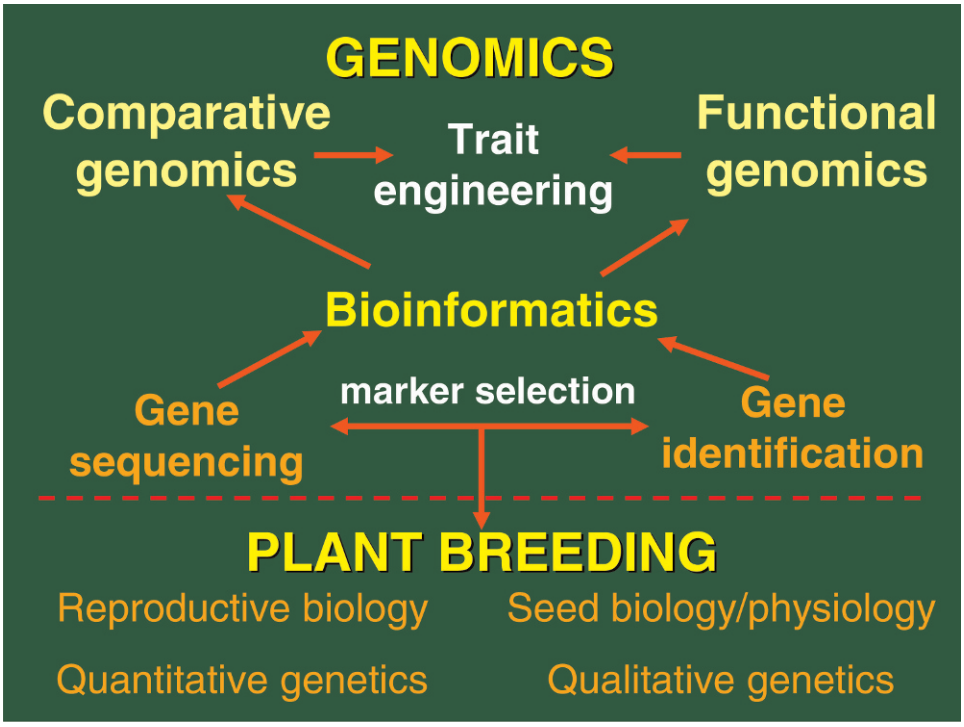

Fig. 1. Areas of research in genomic science.

Large data sets are compiled by the high thoughput instruments used to analyze genomes, and extensive bioinformatic analyses are required to sort, compare, and help interpret results. Unfortunately for the discipline, lists of analytical tools used in molecular biology and subsequent presentation of large data sets have tended to obscure fundamental significance of genomic science.

The research conducted in genomic science is highly labor intensive and expensive. Classical plant breeding programs, which are relatively inexpensive, are not well adapted for utilizing advanced technologies associated with genomics. Further, a large percentage of scientists who perform genomic research are mainly interested in the molecular function of specific genes or processes and are usually less interested in marker development for phenotypic selection applications. On the other hand, plant breeders need markers to facilitate selection and are generally not interested in developing large data sets for sequencing specific genes. Although the gap between the producer of genomic information (molecular biologist) and the user (plant breeders) is very wide, there is enormous potential for interactions among disciplines for plant improvement.

A fundamental question is "what can genomics deliver to improve agriculture productivity, quality, and safety"? Although efforts are needed to apply knowledge gained in molecular genetic research, data is generated to understand basic systems for genetic manipulation through transformation technologies, marker assisted selection, gene knockouts, and mutagenesis (Fig. 1). Genomics can identify relevant genetic pathways through transcriptional profiling and proteomics, thus leading to a better understanding of the plant functions that lead to phenotypic traits, plant tolerances, and pest resistances. Utilization of information from other plant species is possible by comparative genomic analysis, thus allowing the knowledge base of the majority of crops that are under-funded to be expanded with little cost. Lastly, high throughput techniques can be exploited to maximize research opportunities when a relatively small number of researchers are working on specific systems or problems. This paper will summarize the major events leading up to the current directions of genomic research in peanut and the current status of genomic-related science in the species.

Background Information about *Arachis hypogaea*

The cultivated peanut, *Arachis hypogaea* L. is a polyploid species ($2n = 2x = 40$) with a large genome consisting of duplicate chromosomes. A hybridization event between two diploid species is believed to have occurred about 3500 years ago that led to the species we commonly know as peanut. Many species have been hypothesized as possible progenitors of *A. hypogaea*, and additional ones continue to be proposed from a variety of sources. However, the most likely ancestors are *A. duranensis* Krapov. and W. C. Gregory (A-

genome) and *A. ipaensis* Krapov.and W. C. Gregory (B-genome) (Kochert *et al.*, 1996). *Arachis duranensis* also is believed to be the female parent of the ancestral hybrid based on analyses of cytoplasmic genes (Hilu and Stalker, 1995). Temsch and Greihuber (2000) estimated that *A. hypogaea* has 5.91 pg DNA; and *A. duranensis* ranges between 2.49 and 2.87 pg DNA, depending on elevation and latitude of the accession (Temsch and Greihuber, 2001). Singh *et al.* (1996) concluded that the A and B genomes contributed nearly equal amounts of DNA to the domesticated peanut.

Although the genomes of several plant species are in the process of being sequenced, producing a complete DNA sequence-based map of peanut is unlikely because of its large genome size, the species being polyploid with duplicate genes in the two genomes, and the large resource base that will be required to complete the task (Paterson *et al.*, 2004). The *Arachis* genome also has a large fraction of repetitive DNA, for example Dhillon *et al.* (1980) estimated that *A. duranensi*s has 27% highly repetitive, 37% middle-repetitive, and 36% low-copy DNA. Alternatives to genome sequencing are available that can provide useful information about the plant genome, however, such as developing expressed sequence tag (EST) libraries. These libraries supply information about the genes being expressed (transcribed) in different tissues; although they do have limitations such as being a snapshot in time for gene activation, not providing information about gene regulation, and rare transcripts may not be detected. Importantly, the genes (regulatory and low abundance) may be important for many of the traits that distinguish peanut genotypes or separate it from other legumes (Paterson *et al.*, 2004). Reviews have previously been published relating the history of molecular research in peanut (Stalker and Mozingo, 2001; Paterson *et al.*, 2004) so only a brief account will be presented in this manuscript.

Marker Development in Peanut

Research with molecular aspects of the peanut genome began in the 1980s when protein and isozyme variation in *A. hypogaea* was determined to be of little use for characterizing variation within the cultivated peanut. Although large numbers of polymorphisms were detected among other species in the genus (Lu and Pickersgill, 1993; Stalker *et al.*, 1994), the number of markers was too small to be routinely used in breeding programs. Restriction Fragment Length Polymorphisms (RFLPs) significantly expanded the number of data points available for analyses; but similar to isozymes and proteins, non-significant amounts of variation were found within the *A. hypogaea* gene pool (Kochert *et al.*, 1991; Halward *et al.*, 1991). Greater numbers of polymorphisms were found among *Arachis* species (Kochert *et al.*, 1991; Paik-Ro *et al.*, 1992) and were useful for clustering species of *Arachis* (Kochert *et al.*, 1991). Further, intraspecific variation was found within the diploid *A. duranensis* (Stalker *et al.*, 1995) and selected markers are useful for identifying individual accessions. The first molecular map of peanut was produced with RFLPs (Halward *et al.*, 1993).

Development of Polymerase Chain Reaction (PCR) technologies, allowed DNA analyses with smaller aliquots of material and without the use of radioactivity. The commonly used Randomly Amplified Polymorphic DNA (RAPD) analysis for gene identification utilizes PCR technologies; but the method has the disadvantage of only detecting dominant markers. Observable differences by RAPD technologies within the cultivated peanut and among *Arachis* species are similar to reports for other marker systems (Halward *et al.*, 1992). Amplified Fragment Length Polymorphisms (AFLPs) also are PCR-based technologies, but with greatly increased numbers of data points on electrophoretic gels. He and Prakash (1997) used 28 primer pairs to generate 111 AFLPs for DNA markers in *A. hypogaea*, with about 3% of the primers amplifying polymorphisms. AFLPs were the first molecular marker system used to differentiate closely related peanut cultivars (Herselman, 2003).

Hopkins *et al.* (1999) isolated the first Single Sequence Repeat (SSR) markers in peanut He *et al.* (2003) later reported that microsatellite markers developed from SSRs are more variable than other types of markers in peanut and 19 markers were polymorphic among *A. hypogaea* genotypes. Many hundreds of SSR makers have been developed (S. Knapp, personal commun.; Luo *et al.*, 2005a), with about 30% being polymorphic among *A. hypogaea* lines (Ferguson *et al.*, 2004; He *et al.*, 2005; Moretzsohn *et al.*, 2005). SSRs can be used to separate the cultivated lines (Moretzsohn *et al.*, 2005) and they hold great potential for developing useful marker systems for improving selection efficiency in peanut. To date, the number of genes associated with molecular markers in peanut is small (Table 1), but the large number of molecular markers becoming available have great potential for utilizing in a crop improvement program. The next step is for plant breeders to develop appropriate plant populations for testing markers to establish associations with traits of interest. This will require cooperation between breeding and molecular genetic programs. These technologies have the greatest potential for traits that are

Table 1. Genes associated with molecular markers in peanut.

Trait	Marker system*	Origin	Reference
M. arenaria resistance	RFLP	Interspecific cross with *A. hypogaea*	Burow *et al.* (1996)
M. arenaria resistance	RFLP	Interspecific cross (diploid species)	Garcia *et al.* (1996)
M. arenaria resistance	SCAR	Interspecific cross with *A. hypogaea*	Chu *et al.* (2007)
Components of *Cercospora arachidicola* resistance	RAPD	Interspecific cross with *A. hypogaea* and an *A. hypogaea* cross	Stalker and Mozingo (2001)
Empoasca fabae Harris (leafhopper) resistance	RAPD	Interspecific cross with *A. hypogaea*	Stalker and Mozingo (2001)
Diabrotica undecimpunctata howardi Barber (corn rootworm) resistance	RAPD	*A. hypogaea* cross	Stalker and Mozingo (2001)
Cylindrocladium black rot resistance	RAPD	Interspecific cross with *A. hypogaea*	Stalker and Mozingo (2001)
Plant color	RAPD	*A. hypogaea* cross	Stalker and Mozingo (2001)
Aflatoxin contamination	AFLP	Interspecific cross (diploid species)	Milla *et al.* (2005)
Tomato spotted wilt virus resistance	AFLP	Interspecific cross (diploid species)	Milla *et al.* (2004)
Aphid (*Aphis craccivora*) resistance	AFLP	*A. hypogaea* lines	Herselman *et al.* (2004)
High oleic acid fatty acid content	CAP		Lopez and Burow (2004)
Sclerotinia minor	SSR	*A. hypogaea* lines	Chenault and Maas (2006)
Drought resistance	RTPCR	*A. hypogaea* lines	Jain *et al.* (2001)

*RFLP = Restriction Fragment Length Polymorphism; SCAR = Sequence characterized amplified region; RAPD = Randomly Amplified Polymorphic DNA; AFLP = Amplified Fragment Length Polymorphism; CAP = Cleaved amplified polymorphism; SSR = Single Sequence Repeat; RTPCR = Reverse transcriptase polymerase chain reaction.

complex, multigenic, and/or that are difficult to select in the greenhouse or field.

In addition to markers being useful for associating with specific traits, they also may be useful for following introgression from *Arachis* species to *A. hypogaea*. This is important because recombination between the cultivated genomes and those of other species is rare, thus restricting selection for desired traits in interspecific hybrid derivatives (Holbrook and Stalker, 2003). Introgression is possible, however, because Garcia *et al.* (1995) observed genes from the A-genome species *A. cardenasii* Krapov. and W. C. Gregory in advanced generation *A. hypogaea* hybrids in 10 of 11 linkage groups on the diploid RFLP map developed by Halward *et al.* (1993).

Genomic Affinities and Speciation

Studies to understand the phylogenetic relationships among species in section *Arachis*, and to a lesser extent in other groups, have been made using isozyme variation (Stalker *et al.*, 1994; Lu and Pickersgill, 1993), seed storage proteins (Singh *et al.*, 1991; Bianchi-Hall *et al.*, 1993; Liang et al. 2006), RFLPs (Kochert *et al.*, 1991; Halward *et al.*, 1991; Paik-Ro *et al.*, 1992), SSRs (Hopkins *et al.*, 1999), RAPDS (Halward *et al.*, 1992; Lanham *et al.*, 1992; Hilu and Stalker, 1995), and *in situ* hybridization (Raina and Mukai, 1999). The accumulated molecular data supports the hypothesis that *A. hypogaea* originated from a single hybridization event followed by chromosome doubling, and since its origin there has been very little to no introgression from diploid species into the cultivated peanut. Differences exist among marker systems when associating individual accessions within a species and for determining genetic distance between species; but in general, species cluster into genomic groups, with the B and D genomes being more related to each other than to the A genome. Polyploidy has evolved independently in the section *Arachis* (i.e., *A. hypogaea* and *A. monticola* Krapov. and Rigoni) and in species of section *Rhizomatosae* (Smartt and Stalker, 1982); and Nelson *et al.* (2006) used RAPD methodology to conclude that the *Rhizomatosae* species *A. glabrata* Benth. and *A. pseudovillosa* (Chodat and Hassl.) Krapov. and W. C. Gregory also may have evolved independently.

Molecular Maps of Peanut

Several molecular maps have been produced in peanut with different marker systems. The first map used RFLPs and utilized variation between the diploid species *A. stenosperma* Krapov. and W. C. Gregory × *A. cardenasii* where a total of 117 RFLP markers were mapped into 11 linkage groups (Halward *et al.*, 1993). Garcia *et al.* (2005) developed a RAPD-based linkage map of peanut based on a backcross population [*A. stenosperma* × (*A. stenosperma* × *A. cardenasii*)] where 167 RAPD loci and 39 RFLPs were mapped to 11 linkage

groups and all common markers mapped to the same linkage groups and mostly in the same order as in the map previously developed by Halward *et al.* (1993).

Another RFLP map was created by using progenies of a cross between *A. hypogaea* and TxAG-6 {a polyploidy of one B-genome and two A-genome diploid species: 4*x*[*A. batizocoi* Krapov. and W. C. Gregory × (*A. cardenasii* × *A. diogoi* Hoehne)]} where 383 markers were mapped (Burow *et al.*, 1996). Genetic inheritance was disomic (with the exception of 1 linkage group), and the marker R239 was associated with nematode resistance. This marker mapped to the same linkage group in the map produced by Halward *et al.* (1993). Although the Halward *et. al.* (1993), Burow *et al.* (1996), and Garcia *et al.* (2005) maps represent a good starting point for unraveling the peanut genome, they do not have sufficient numbers of markers to be highly useful for genetic studies and, because they are based on diploid species, there are problems relating the information to the cultivated peanut.

A microsatellite map with markers grouped into 11 linkages was made by Moretzsohn *et al.* (2005) by utilizing a cross between *A. duranensis* and *A. stenosperma* (both A-genome species). A second map with the B-genome species *A. ipaensis* and *A. magna* Krapov., W. C. Gregory, and C. E. Simpson used the same markers and a comparison between the A- and B-genome maps indicated that they were generally collinear (Moretzsohn *et al.*, 2005). Lastly, Herselman *et al.* (2004) mapped 12 AFLP markers into five linkage groups by using *A. hypogaea* crosses.

Creating a comprehensive physical map of *A. hypogaea* will require a very large-scale effort (Paterson *et al.*, 2004) and will require a library of large-insert DNA clones. Yüksel and Paterson (2005) produced a bacterial artificial clone (BAC) library with about 180,000 clones using the tetraploid peanut cultivar 'Florunner'. The BAC library should serve as a highly useful resource for developing a physical map, but a limitation is the presence of duplicated copies of genes from the two progenitor species and Paterson *et al.* (2004) concluded that it will be difficult to distinguish true homologes vs. homoeologs from related species in BACs by hybridization-based molecular approaches. Although Lin *et al.* (2000) developed a method to determine the subgenome-specificity of individual BAC clones in polyploid species libraries, an alternative may be to produce BAC libraries for diploid progenitors of peanut and then compare them with the *A. hypogaea* libraries (Paterson *et al.*, 2004).

TILLING Research in Peanut

Targeting Induced Local Lesions IN Genomes (TILLING) is a method developed to find genes of interest in a mutant population of a species. Because allergens in peanut cause a significant health problem in the human population, a high priority of the peanut industry is to eliminate or suppress the proteins that cause allergens. There are multiple seed storage proteins that give rise to allergens, with the major ones being *Ara h 1*, *Ara h 2*, and *Ara h 3* (Burks *et al.*, 1998); *Ara h 2* is the most important peanut allergen (Koppelman *et al.*, 2004). Ozias-Akins (pers. commun.) is developing a TILLING population in peanut with the goal of eliminating *Ara h 2* and possibly other allergen genes.

Peanut Transformation

Ozias-Akins *et al.* (1993) reported the first successful transformation of peanut, with accompanying plant regeneration, by utilizing the micro-bombardment technique. Micro-bombardment has since been completed in peanut with a number of genes conferring disease resistance (Ozias-Akins and Gill, 2001; Magbanua *et al.*, 2000; Yang *et al.*, 1998; Higgins *et al.*, 2004; Dar *et al.*, 2006). Efficiency levels remain low and the process takes many months from when the initial transformation event is induced until plant maturity (Egnin *et al.*, 1998). A high-efficiency, rapid technique to transform peanut is greatly needed and Agrobacterium-mediated transformation offers the possibility to achieve this goal. Cheng *et al.* (1996) used this method on a Valencia-type peanut, but other investigators have been unable to expand the methodology to other genotypes, thus restricting its usefulness. To date, biolistic methodologies are more reliable in peanut than other transformation methodologies and single constructs can be inserted into the peanut genome. Additional research is needed to insert multiple constructs and BAC clones to produce stable progenies. The regulatory process of germplasm release for consumption also must be explored as well as industry acceptance of transformed products.

Gene Sequencing in *Arachis*

Because the peanut is a polyploid with a large genome size, complete sequencing will be too expensive and labor intensive to perform with current resources. However, expressed sequence tags (ESTs) can give significant information about the function of genes. As of June, 2007, 12,832 long-sequence ESTs have been deposited in GenBank, with the majority from seeds. A large number of ESTs have been produced from seeds and leaves which will deposited in GenBank (Guo *et al.*, 2004, 2008; Stalker and N. Nielson, unpublished data; Chen *et al.*, 2006), and more

than 25,000 unigenes have been identified from existing datasets (S. Knapp, personal commun.). In addition, Jayshree *et al.* (2005) identified 1312 short sequences that were isolated from SSR-enriched libraries and (S. Knapp, pers. commun.) is sequencing large numbers of short DNA segments with methylation-filtration techniques.

In addition to developing long and short read ESTs, several investigators have sequenced specific genes found in peanut. Two examples relevant to peanut are the isolation and characterization of the Δ^{12}-fatty acid desaturase gene by Lopez *et al.* (2000) and sequencing the *Ara h 2* gene which gives rise to one of the proteins causing allergens in humans by Ramos *et al.* (2006). These genes were sequenced to produce very specific markers to detect the genes in breeding or TILLING populations.

Genomic sequencing and microarray-based screening has been used to identify putative genes that may be associated with resistance to *Aspergillus parasiticus* and drought stress (Luo *et al.*, 2005b) and for aflatoxin contamination (Guo *et al.*, 2003). In addition, Yüksel *et al.* (2005) evaluated bacterial artificial clones from the library and found 250 putative resistance gene loci in peanut.

Strategic Planning – Peanut and Other Legumes

High-throughput genomic analyses are expensive and the vast majority of the financial resources for plants have been directed toward model systems. Recognizing the problem that significantly more resources will be required to make progress in agronomic plants (specifically in legumes) a landmark meeting was organized as the *U.S. Legume Crops Genomics Workshop* during 2001. Scientists involved in genomic research on soybean, peanut, common- and dry beans, alfalfa, peas and lentils, and model legumes met to develop a set of goals and strategy to present the case for increasing support for legume genomic research. They took a very broad view of genomics to include transformation-related research and recognized that molecular biology needs to be integrated with plant breeding efforts to be of long-term value. The outcome of the meeting was the development of a common set of goals across legume species in six areas of research, including: (a) genome sequencing of strategic legume species, (b) physical map development and refinement, (c) functional analysis including both transcriptional and genetic, (d) develop DNA markers for comparative mapping, (e) characterization and utilization of legume biodiversity, and (f) development of a legume data resource. Significantly, as a direct result of the meeting, the Legume Genomics Steering Committee was formed to organize the U.S. Legume Crops Genomics Initiative (USLCGI) to develop a strategy for additional funding for legumes. Members of this committee consisted of leaders in the United Soybean Board, American Soybean Council; National Peanut Foundation, American Peanut Council; National Dry Bean Council; USA Dry Pea and Lentil Council; Alfalfa and Clovers; Model Legumes; a USDA-ARS representative, and a researcher associated with each crop. National Research Initiative funding directed to legumes in 2005 was believed to be a result of their efforts. Additional outcomes of this committee was the book *Legume Crop Genomics* (Wilson *et al.*, 2004) which documented research strategies, development of genomic tools, resources in legumes, and the legume-community consensus on the genomic research objectives.

The conference *Legumes as a Model Plant Family: Genomics for Food and Feed*, was held during 2004 in Santa Fe, NM. At this conference, 50 researchers developed a plan for genomic research with model species and across the legume crops alfalfa (*Medicago sativa* L.), common bean (*Phaseolus vulgaris* L.), pea (*Pisum sativum* L.), lentil (*Lens culinaris* Med.), chickpea (*Cicer arietinu* L. Gram.), peanut, and soybean (*Glycine max* L. Merr.). At the time of the conference, *M. truncatula* and *Lotus japonicus* (Regel) K. Larsen had been studied extensively as model species in the legume family (and soybean to a lesser extent), but other legume species were lagging behind in data generation and analyses (VandenBosch and Stacey, 2003). The goal of the conference was to develop a plan with specific objectives for cross-legume genomics research and to outline milestones for achieving these objectives (http://catg.ucdavis.edu; Gepts *et al.*, 2005). A list of websites related to legumes is presented in Table 2. In addition to being economically important, peanut was of scientific interest to scientists working with other species because it is unique among cultivated legumes in mode of reproduction, is phylogenetically distinct, and has a unique nodule type.

To follow-up on the cross-legume workshop, on March 23, 2004, The Peanut Foundation/American Peanut Council hosted 26 scientists to document the current status of peanut genomic research and then establish priorities for the next 5 yr. Their efforts led to a *Strategic Plan for Peanut Genomics 2004 to 2008* and an accompanying action plan to complete research (http://www.peanutbioscience.com/) that outlines priorities and milestones for peanut genomic research. The scientists took a very broad viewpoint of genomic technologies to include

Table 2. Species-specific information resources.

AlfaGenes: http://ukcrop.net/perl/ace/search/AlfaGenes
Australian Nat. Univ. - 2D-PAGE Database: http://semele.anu.edu.au/
BeanGenes: http://beangenes.cws.ndsu.nodak.edu/
BeanGenes: http://beangenes.cws.ndsu.nodak.edu/
CNRS-INRA Medicago truncatula EST Program: http://medicago.toulouse.inra.fr/Mt/EST/DOC/MtB.html
CoolGenes: http://ukcrop.net/perl/ace/search/CoolGenes
European Union Medicago truncatula Research Program: http://medicago.toulouse.inra.fr/EU/mtindex.htm
Kazusa *L. japonicus*: http://www.kazusa.or.jp/lotus/
Medicago truncatula Consortium: http://www.medicago.org/
MtDB: http://www.medicago.org
Phaseomics: *Phaseolus* genome initiative: http://www.phaseolus.net
S. R. Noble Foundation Center for Medicago Genomic Research: http://www.noble.org/medicago/index.htm
Soybase: http://soybase.org/
Data Mining from Other Species
Legumes.org: http://www.legumes.org/
Legume Information System: http://comparative-legumes.org/
National Center for Biotechnology Information: http://www.ncbi.nih.gov/
Plant Genome DataBase: http://www.plantgdb.org/
The TIGR Plant Repeat Databases: http://www.tigr.org/tdb/e2k1/plant.repeats/index.shtml
Conferences and Other
1st International Conference on Legume Genetics and Genomics: http://www.agro.agri.umn.edu/iclgg/index.htm
2nd International Conference on Legume Genomics and Genetics: http://www.grainlegumes.com/aep/events/5th_gl_conference_dijon_2004
Cross-Legume Advances through Genomics (CATG) Conference: http://catg.ucdavis.edu/
International Legume Database and Information Service: http://www.ildis.org/
Legumes as a Model Plant Family: http://www.plantphysiol.org/cgi/content/full/137/4/1228
International Conference on Groundnut Aflatoxin Management and Genomics: http://www.aflatoxin.info/Conference/Abstract%20book_Poster3.pdf
Peanut Genomics Websites
Legume Information System (LIS): http://www.comparative-legumes.org
Peanut Bulletin Board at LIS: http://www.comparative-legumes.org:2000/forum
PeanutMap: an Online Genome Database for Comparative Molecular Maps of Peanut http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=16904007
Peanut Genomics: http://peanut.ccgb.umn.edu/
Peanut Genomics Initiative: http://www.PeanutBioscience.com

(a) simple markers to high-through put analysis, (b) gene discovery, (c) translational genetics, (d) transformation technologies, and (e) a recognition of the importance of germplasm and populations. The group set strategic goals to: (a) improve the utility of genetic tools for peanut genomics research, (b) improve the efficacy of technology for gene manipulation in genomics, (c) develop a framework for assembling the peanut genetic blueprint, (d) improve knowledge of gene identification and regulation, and (e) provide bioinformatics management of peanut biological information (http://www.peanutbioscience.com/). They also stated that a priority area for research in peanut is to increase peanut safety, especially by eliminating peanut allergens.

Summary of Progress in Peanut Genomic Research

Molecular markers should be useful for mapping and genetic studies, for estimating associations with plant and agronomic traits important to plant breeding efforts, to insure genes have been transferred from *Arachis* species to *A. hypogaea* in interspecific hybrid populations, to determine evolutionary relationships among species, and for plant identity assurances. The molecular data analyzed to date indicates that (a) variation within *A. hypogaea* ranges between 1 and 5% for different marker systems, (b) a large amount of variation exists among *Arachis* species for all marker systems, (c) very low-density maps have been created, (d) few genes of agronomic importance have been associated with various molecular markers, (e) introgression from *Arachis* species to *A. hypogaea* is possible, (f) transformation technologies are available for inserting genes from distantly related species to peanut, and (g) new molecular technologies promise to solve many of ''low genetic variability'' problems in peanut. To have effective use of genomic technologies in peanut (and other crop species), cooperative

efforts must be developed between the molecular biologist and plant breeder to answer relevant biological questions and to utilize the tools made available by genomics. Successful utilization of molecular genetics for crop improvement will require integration of genomic science with classical methodologies to investigate problems that are too complex to be solved by traditional methodologies.

Literature Cited

Bianchi-Hall, C., R.D. Keys, H.T. Stalker, and J.P. Murphy. 1993. Diversity of seed storage proteins in wild peanuts (*Arachis* species). Pl. Syst. Evol. 186:1-15.

Burks, W., H.A. Sampson, and G.A. Bannon. 1998. Peanut allergens. Allergy 53:725-730.

Burow, M.D., C.E. Simpson, A.H. Paterson, and J.L. Starr. 1996. Identification of peanut (*Arachis hypogaea* L.) RAPD markers diagnostic of root-knot nematode (*Melooidogyne arenaria* (Neal) Chitwood) resistance. Molec. Breed. 2:369-379.

Chen, H., C.C. Holbrook, and B.Z. Guo. 2006. Peanut seed transcriptome: construction of six peanut seed cDNA libraries from two peanut cultivars. Amer. Peanut Res. Educ. Soc. 38:73 (abstr.).

Chenault, K.D. and A.L. Maas. 2006. Identification of a simple sequence repeat (SSR) marker in cultivated peanut (*Arachis hypogaea* L.) potentially associated with Sclerotinia blight resistance. Proc. Amer. Peanut Res. Educ. Soc. 37:24-25 (abstr.).

Cheng, M., R. Jarret, Z. Li, A. Xing, and J. Demski. 1996. Production of fertile transgenic peanut (*Arachis hypogaea* L.) plants using *Agrobacterium tumefaciens*. Plant Cell Rep. 15:653-657.

Chu, Y., C.C. Holbrook, P. Timper, and P. Ozias-Akins. 2007. Development of a PCR-based molecular marker to select for nematode resistance in peanut. Crop Sci. 47:841-845.

Dar, W.D., B.V.S. Reddy, C.L.L. Gowda, and S. Ramesh. 2006. Genetic resources enhancement of ICRISAT-mandate crops. Cur. Sci. 91:880-884.

Dhillon, S.S., A.V. Rake, and J.P. Miksche. 1980. Reassociation kinetics and cytophotometric characterization of peanut (*Arachis hypogaea* L.) DNA. Plant Physiol. 65:1121-1127.

Egnin, M., A. Mora, and C.S. Prakash. 1998. Factors enhancing *Agrobacterium tumefaciens*-mediated gene transfer in peanut (*Arachis hypogaea* L.). *In Vitro* Cell Dev. Biol. Plant. 34:310-318.

Ferguson, M.E., M.D. Burow, S.R. Schulze, P.J. Bramel, A.H. Paterson, S. Kresovich, and S. Mitchell. 2004. Microsatellite identification and characterization in peanut (*A. hypogaea* L.). Theor. Appl. Genet. 108:1064-1070.

Garcia, G.M., H.T. Stalker, and G.A. Kochert. 1995. Introgression analysis of an interspecific hybrid population in peanuts (*Arachis hypogaea* L.) using RFLP and RAPD markers. Genome 38:166-176.

Garcia, G.M., H.T. Stalker, E. Shroeder, and G. Kochert. 1996. Identification of RAPD, SCAR and RFLP markers tightly linked to nematode resistance genes introgressed from *Arachis cardenasii* to *A. hypogaea*. Genome 39:836-845.

Garcia, G.M., H.T. Stalker, E. Schroeder, J.H. Lyerly, and G. Kochert. 2005. A RAPD-based linkage map of peanut based on a backcross population between the two diploid species *Arachis stenosperma* and *A. cardenasii*. Peanut Sci. 32:1-8.

Gepts, P., W.D. Beavis, E.C. Brummer, R.C. Shoemaker, H.T. Stalker, N.F. Weeden, and N.D. Young. 2005. Legumes as a model plant family. Genomics for food and feed report of the cross-legume advances through genomics conference. Plant Phys. 137:1228-1235.

Guo, B.Z., J. Yu, R.D. Lee, C.C. Holbrook, and R.E. Lynch. 2003. Application of differential display RT-PCR and EST/microarray technologies to the analysis of gene expression in response to drought stress and elimination of aflatoxin contamination in corn and peanut. J. Toxicology-Toxin Reviews 22:291-316.

Guo, B.Z., X. Chen, P. Dang, B.T. Scully, X. Liang, C.C. Holbrook, J. Yu, and A.K. Culbreath. 2008. Peanut gene expression profiling in developing seeds at different reproduction stages during *Aspergillus parasiticus* infection. BMC Developmental Biology 8:12.

Guo, B., M. Luo, P. Dang, G. He, and C.C. Holbrook. 2004. Peanut expressed sequence tag (EST) project and the marker development for cultivated peanut (*Arachis hypogaea*). Proc. International Crop Sci. Congress 2004: http://www.cropscience.org.au/icsc2004/poster/3/1/424_guob.htm

Halward, T., T. Stalker, E. LaRue, and G. Kochert. 1992. Use of single-primer DNA amplifications in genetic studies of peanut (*Arachis hypogaea* L.). Plant Mol. Biol. 18:315-325.

Halward, T., H.T. Stalker, and G. Kochert. 1993. Development of an RFLP linkage map in diploid peanut species. Theor. Appl. Genet. 87:379-384.

Halward, T., H.T. Stalker, E.A. Laure, and G. Kochert. 1991. Genetic variation detectable with molecular markers among unadapted germplasm resources of cultivated peanut and related wild species. Genome 34:1013-1020.

He, G. and C.S. Prakash. 1997. Identification of polymorphic DNA markers in cultivated peanut (*Arachis hypogaea* L.). Euphytica 97:143-149.

He, G.H., R. Meng, H. Gao, B. Guo, G. Gao, M. Newman, R.N. Pittman, and C.S. Prakash. 2005. Simple sequence repeat markers for botanical varieties of cultivated peanut (*Arachis hypogaea* L.). Euphytica 142:131-136.

He, G.H., R. Meng, M. Newman, G. Gao, R.N. Pittman, and C.S. Prakash. 2003. Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.). BMC Plant Biology 3:3.

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

Herselman, L., R. Thwaites, F.M. Kimmins, B. Courtois, P.J.A. van der Merwe, and S.E. Seal. 2004. Identification and mapping of AFLP markers linked to peanut (*Arachis hypogaea* L.) resistance to the aphid vector of groundnut rosette disease. Theor. Appl. Genet. 109:1426-1433.

Higgins, C., R. Hall, N. Mitter, A. Cruickshank, and R. Dietzgen. 2004. Peanut stripe potyvirus resistance in peanut (*Arachis hypogaea* L.) plants carrying viral coat protein gene sequences. Transgenic Res. 13:59-67.

Hilu, K. and H.T. Stalker. 1995. Genetic relationships between peanut and wild species of *Arachis* section *Arachis* (Fabaceae): Evidence from RAPDs. Plant Syst. Evol. 188:167-178.

Holbrook, C.C. and H.T. Stalker. 2003. Peanut breeding and genetic resources. Plant Breed. Rev. 22:297-356.

Hopkins, M.S., A.M. Casa, T. Wang, S.E. Mitchell, R.E. Dean, G.D. Kochert, and S. Kresovich. 1999. Discovery and characterization of polymorphic simple sequence repeats (SSRs) in peanut. Crop Sci. 39:1243-1247.

Jain, A.K., S.M. Basha, and C.C. Holbrook. 2001. Identification of drought - responsive transcripts in peanut (*Arachis hypogaea* L.). Mol. Biotechnology Genetics 4(2).

Jayashree, B., M. Ferguson, D. Ilut, J. Doyle, and J.H. Crouch. 2005. Analysis of genomic sequences from peanut (*Arachis hypogaea*). Elec. J. Biotech. 8:226-237.

Kochert, G., T. Halward, W.D. Branch, and C.E. Simpson. 1991. RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. Theor. Appl. Genet. 81:565-570.

Kochert, G., H.T. Stalker, M. Gimenes, S.L. Galgaro, C.R. Lopes, and K. Moore. 1996. RFLP and cytological evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). Amer. J. Bot. 83:1282-1291.

Koppelman, S.J., M. Wensing, M. Ertmann, A.C. Knulst, and E.F. Knol. 2004. Relevance of Ara h1, Ara h2 and Ara h3 in peanut-allergic patients, as determined by immoglobulin E Western blotting, basophile-histamine release and intracultaneous testing: Ara h2 is the most important peanut allergen. Clin. Exper. Allergy 34:583-590.

Lanham, P.G., S. Fennell, J.P. Moss, and W. Powell. 1992. Detection of polymorphic loci in *Arachis* germplasm using random amplified polymorphic DNAs. Genome 35:885-889.

Liang, X.Q., M. Luo, C.C. Holbrook, and B.Z. Guo. 2006. Storage protein profiles in Spanish and runner market type peanuts and potential markers. BMC Plant Biol. 6:24.

Lin, Y.R., X. Draye, X. Qian, S. Ren, L. Zhu, J. Tomkins, R.A. Wing, Z. Li, and A.H. Paterson. 2000. Fine-scale mapping and sequence-ready contig assembly in highly-duplicated genomes, using the BAC-RF method. Nucl. Acids Res. 2823: http://www3.oup.co.uk/nar/methods/Volume_28/Issue_07/gnd023_gml.abs.html.

Lopez, Y. and M.D. Burow. 2004. Development and validation of CAPS markers for the high oleate trait in peanuts. Proc. Amer. Peanut Res. Educ. Soc. 36:25-26 (abstr.).

Lopez, Y., H.L. Nadaf, O.D. Smith, J.P. Connell, A.S. Reddy, and A.K. Fritz. 2000. Isolation and characterization of the Δ^{12}-fatty acid desaturase in peanut (*Arachis hypogaea* L.) and search for polymorphisms for the high oleate trait in Spanish market-type lines. Theor. Appl. Genet. 101:1131-1138.

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

Luo, M., P. Dang, B.Z. Guo, G. He, C.C. Holbrook, M.G. Bausher, and R.D. Lee. 2005a. Generation of expressed sequence tags (ESTs) for gene discovery and marker development in cultivated peanut. Crop Sci. 45:346-353.

Luo, M., X.Q. Liang, D. Dang, C.C. Holbrook, M.G. Bausher, R.D. Lee, and B.Z. Guo. 2005b. Microarray-based screening of differentially expressed genes in peanut in response to *Aspergillus parasiticus* infection and drought stress. Plant Sci. 169:695-703.

Magbanua, Z., H. Wilde, J. Roberts, K. Chowdhury, J. Abad, J. Moyer, H. Wetzstein, and W. Parrott. 2000. Field resistance to tomato spotted wilt virus in transgenic peanut (*Arachis hypogaea* L.) expressing an antisense nucleocapsid gene sequence. Mol. Breed. 6:227-236.

Milla, S.R., S.P. Tallury, H.T. Stalker, and T.G. Isleib. 2004. Identification of molecular markers associated with tomato spotted wilt virus in a genetic linkage map of *Arachis kuhlmannii* × *A. diogoi*. Proc Am Peanut Res and Educ Soc 36:27 (abstr.).

Milla, S.R., T.G. Isleib, and S.P. Tallury. 2005. Identification of AFLP markers linked to reduced aflatoxin accumulation in *A. cardenasii*-derived germplasm lines of peanut. Proc. Amer. Peanut Res. and Educ. Soc. 37:90 (abstr.).

Moretzsohn, M.C., L. Leoi, K. Proite, P.M. Guimarães, S.C.M. Leal-Bertioli, M.A. Gimenes, W.S. Martins, J.F.M. Valls, D. Grattapaglia, and D.J. Bertioli. 2005. A micro satellite–based, gene–rich linkage map for the AA genome of *Arachis* (Fabaceae). Theor. Appl. Genet. 111:1060-1071.

Nelson, A., D.M. Samuel, J. Tucker, C. Jackson, and A. Stahlecker-Roberson. 2006. Assessment of genetic diversity and sectional boundaries in tetraploid peanuts (*Arachis*). Peanut Sci. 33:64-67.

Ozias-Akins, P. and R. Gill. 2001. Progress in the development of tissue culture and transformation methods applicable to the production of transgenic peanut. Peanut Sci. 28:123-131.

Ozias-Akins, P., J.A. Schnall, W.F. Anderson, C. Singsit, T.E. Clemente, M.J. Adang, and A.K. Weissinger. 1993. Regeneration of transgenic peanut plants from stably transformed embryogenic callus. Plant Sci. 93:185-194.

Paik-Ro, O.G., R.L. Smith, and D.A. Knauft. 1992. Restriction fragment length polymorphism evaluation of six peanut species within the *Arachis* section. Theor. Appl. Genet. 84:201-208.

Paterson, A.H., H.T. Stalker, M. Gallo-Meagher, M.D. Burow, S.L. Dwivedi, J.H. Crouch, and E.S. Mace. 2004. Genomics and genetic enhancement of peanut, pp. 97-109. *In* R.F. Wilson, H.T. Stalker, and C.E. Brummer (eds.). Genomics for Legume Crops. Amer. Oil Chemical Soc. Press, Champaign, IL.

Raina, S.N. and Y. Mukai. 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. 214:1-4.

Ramos, M.L., G. Fleming, Y. Chu, Y. Akiyama, M. Gallo, and P. Ozias-Akins. 2006. Genomic and evolutionary context for conglutin genes in *Arachis*. Mol. Gen. 275:578-592.

Singh, K.P., S.N. Raina, and A.K. Singh. 1996. Variation in chromosomal DNA associated with the evolution of *Arachis* species. Genome 39:890-897.

Singh, A.K., S.S. Krishnan, M.H. Mengesha, and C.D. Ramaiah. 1991. Polygenetic relationships in section *Arachis* based on seed protein profiles. Theor. Appl. Genet. 82:593-597.

Smartt, J. and H.T. Stalker. 1982. Speciation and cytogenetics in *Arachis*, pp. 21-49. *In* H. Pattee and C. Young (eds.). Peanut Science and Technology. Amer. Peanut Res. Educ. Soc., Yoakum, TX.

Stalker, H.T., T.G. Phillips, J.P. Murphy, and T.M. Jones. 1994. Diversity of isozyme patterns in *Arachis* species. Theor. Appl. Genet. 87:746-755.

Stalker, H.T., J.S. Dhesi, and G.D. Kochert. 1995. Variation within the species *A. duranensis*, a possible progenitor of the cultivated peanut. Genome 38:1201-1212.

Stalker, H.T. and L.G. Mozingo. 2001. Molecular markers of *Arachis* and marker- assisted selection. Peanut Sci. 28:117-123.

Temsch, E.M. and J. Greihuber. 2000. Genome size variation in *Arachis hypogaea* and *A. monticola* re-evaluated. Genome 43:449-451.

Temsch, E.M. and J. Greihuber. 2001. Genome size in *Arachis duranensis*: A critical study. Genome 44:826-830.

Vanden Bosch, K.A. and G. Stacy. 2003. Summaries of legume genomics projects from around the globe. Community resources for crops and models. Plant Physiol 131:840-865.

Wilson, R.F., H.T. Stalker, and C.E. Brummer (eds.). 2004. Genomics for Legume Crops. Amer. Oil Chemical Soc. Press, Champaign, IL. 362 pp.

Yang, H., C. Singsit, A. Wang, D. Gonsalves, and P. Ozias-Akins. 1998. Transgenic peanut plants containing a nucleocapsid protein gene of tomato spotted wilt virus show divergent levels of gene expression. Plant Cell Rep. 17:693-699.

Yüksel, B. and A.H. Paterson. 2005. Construction and characterization of peanut *HindIII* BAC library. Theor. Appl. Genet. 111:630-639.

Yüksel, B., J.C. Estill, S.R. Schulze, and A.H. Paterson. 2005. Organization and evolution of resistance gene analogs in peanut. Mol. Gen. Genomics. 274:248-263.