

Revised 3/13/6

**United States  
Department of  
Agriculture**

Research, Education &  
Economics

Agricultural Research  
Service

National Program Staff

March 2006

# **National Strategic Plan for the Peanut Genome Initiative**

**2004-2008**

**Version 2.4**



## **TABLE OF CONTENTS**

<b>EXECUTIVE SUMMARY</b>	<b>3</b>
<b>INTRODUCTION</b>	<b>5</b>
<b>RESEARCH PRIORITIES FOR PEANUT GENOMICS</b>	
<b>-GENETIC TOOLS &amp; BREEDING METHODS</b>	<b>6</b>
<b>-PLANT TRANSFORMATION TECHNOLOGY</b>	<b>8</b>
<b>-GENOME SEQUENCING AND GENE DISCOVERY</b>	<b>12</b>
<b>-FUNCTIONAL GENOMICS &amp; PROTEOMICS</b>	<b>14</b>
<b>-BIOINFORMATICS</b>	<b>15</b>
<b>-IMMUNOLOGY OF PEANUT PROTEINS IN MODEL SYSTEMS</b>	<b>17</b>
<b>APPENDICES</b>	
<b>-PROCESS</b>	<b>18</b>
<b>-PEANUT GENOME INITIATIVE ADDRESS LIST</b>	<b>19</b>

## EXECUTIVE SUMMARY

On March 23, 2004, twenty-six scientists with expert knowledge of the critical fields of peanut genomics participated in a workshop hosted by the Peanut Foundation/American Peanut Council in Atlanta GA. These scientists reviewed the current status of peanut genomic research and reached consensus on the following strategic plan framework for outlining research priorities and significant near-term milestones that represent ‘quantum leaps’ in the advancement of the science of peanut genomics.

### Strategic Plan Framework

The structure of this plan—five strategic goals and five research objectives—provides both a program and a scientific focus to ensure that the peanut genomics community attains planned results in an effective and timely manner. The strategic goals span the programmatic range of the field of plant genomics. The research objectives address initiatives that seek to improve core capabilities in this scientific arena.

Strategic Goals	Research Objectives
1. Improve the utility of genetic tools for peanut genomics research	Develop useful gene markers & genetic maps for peanut
2. Improve the efficacy of technology for gene manipulation in genomes	Develop useful transformation methods for functional genomic research in peanut
3. Develop a framework for assembling the peanut genetic blueprint	Identify and locate abundant and rarely expressed genes, using genetic and physical approaches to integrate diverse data types
4. Improve knowledge of gene identification and regulation	Provide baseline data and tools that facilitate determination of the biological basis for gene function
5. Provide bioinformatics management of peanut biological information resources.	Establish a state-of-art interactive system and mechanisms for assuring public distribution of high quality data and information
6. Determine the allergenic potential of peanut proteins	Provide quantitative model systems to evaluate allergy and threshold levels of potential allergens

To achieve these strategic goals and research objectives, this plan emphasizes achievements that hinge on teamwork throughout the peanut genomics community. For that reason, all actions and results will be attained in a manner that is both inclusive and open to public scrutiny. As part of this plan, the ability of the research community to carry out and advance peanut genomics in the U.S. public and private sectors will be evaluated in reference to the following performance measures and expected research accomplishments in peanut genomics over the next four years.

<b>Performance Plan for Application of Peanut Genomics Research: 2004-2008</b>				
<b>ST P</b>	<b>Performance Measure</b>	<b>Baseline 2004</b>	<b>Target in 2006</b>	<b>Target in 2008</b>
1.1	Develop a molecular map of the peanut genome	Two maps developed: A diploid map between A-genome species; and a tetraploid map	Develop mapping populations for A- & B-genomes; populations for inbred line selection in <i>A. hypogaea</i>	Develop saturated genetic maps of allergen genes in the A- and B-genomes
1.2	Develop a comprehensive set of genetic markers for peanut	Very few SNPs exist for peanut	Develop 100 new SNP markers for peanut	Develop 1000 new SNP markers for peanut
1.3	Develop DNA resource from the core-collection accessions	Screened core peanut germplasm collection for allergen-null genes	Extract and preserve DNA from sample plants	Establish a repository and distribution system for DNA samples
1.4	Association of phenotypic traits with molecular markers	RFLP, RAPD, AFLPs for a few pest, nematode, morphology, disease traits	Develop mapping populations for quality and yield-limiting traits	Position markers associated with variation in phenotypic traits on genetic maps
1.5	Develop genomic comparisons among legume species for allergens	Few comparisons among peanut species with molecular tools	Identify sequence tag sites in peanut for cross-species markers	Develop reconstituted genetic maps of the peanut genome with CMAP
1.6	Develop DNA marker assisted selection projects for enhancement of peanut germplasm	No MAS programs in place	Associations of DNA markers among genotypic and phenotypic variation for specific traits	Demonstrate the utility of MAS selection in peanut for quality and agronomic traits
2.1	Improve transformation efficiency and utility for peanut	Biolistic protocols are common. Agrobacterium-mediated protocols not repeatable in general application	Establish Agrobacterium-mediated protocols for high through-put transformation in peanut	Develop effective alternative protocols for high through-put transformation of peanut
2.2	Improve transformation technology for peanut genomics	Ability to generate & evaluate transgenics expressing 10 constructs among entire community of transformers	Develop & test new gene promoters, selectable markers & terminators in peanut	Establish high through-put peanut community transformation facilities with expanded capacity
2.3	Develop transgenic screens to understand gene function	Impractical to introduce BAC clones into peanut. Tagging systems developed in theory	Confirm utility of RNAi methods for targeted knockouts and modification of IgE binding sites.	Evaluate heterologous Ac/Ds and/or tDNA tagging systems
2.4	Develop Biotech Risk Assessment & Mitigation Strategies	No known commercial transgenic peanuts. Little data on gene flow or contamination problems	Identify major safety concerns and assess experimental approaches to mitigate allergens	Develop management and/or technological approaches for reducing risk of allergen mitigation
3.1	Advance discovery and knowledge of abundantly expressed genes	5000 expressed sequence tags (ESTs)	50,000 ESTs submitted to GenBank & publically available	150,000 ESTs submitted to GenBank plus chromatograms
3.2	Advance discovery & knowledge of rarely-expressed genes, & regulatory sequences that determine gene expression	Virtually no data. Sequences of various allergen gene isoforms complete	Evaluate methodologies in standard genotypes that represent A, B, and AB genomes	Apply methods. Collect transcriptome sequence information from genotypes that represent diploid A and B genomes
3.3	Integration of diverse tool sets to visualize organizational structure of the peanut genome	A 6X BAC library exists for one Florunner component line	Further characterization of the Florunner BAC library and identification of gene-rich BACs	Florunner BACs anchored to detailed genetic map of the peanut genome. Initiate physical maps of diploid genomes
4.1	Proteomics research to characterize seed protein composition and function relevant to potential allergens	Virtually no proteomics research in peanut	Generate high-density proteome map of mature peanut seed with peptides that represent abundant seed proteins. Characterize allergen protein isoforms	Western blot proteomics to detect allergenic seed proteins. Generate a reference proteomic map for leaves, to target allergen proteins
4.2	Application of reverse genetics to explore gene function	Initial development of a TILLING population for mutations in genes encoding potential allergens	Establish facility for high-throughput TILLING in peanut. Develop M2 TILLING populations. DNA libraries.	Advance TILLING populations to M3 families. Large-scale screening for mutations in allergen genes
5.1	Establish collaborative relations with other legume research communities	No peanut presence in the trait and plant ontology efforts	Assure inclusion of peanut concepts in the plant and trait ontology efforts	Assure implementation of peanut specific genes and traits in controlled vocabularies for all plant bioinformatics resources
5.2	Establish a peanut bioinformatics oversight committee	Establish a peanut bioinformatics oversight committee. Designate developer of a Peanut Information Resource	Establish a curator from the peanut research community who assures data and information content of the Peanut Information Resource	Initiate plans to integrate the Peanut Information Resource into LIS
5.3	Develop a peanut information resource (ArachisDB) as a component of the Legume Information System	There is no single Peanut Information Resource. Peanut genetic & genomic information is stored among diverse resources. A web enabled community bulletin board for peanuts was built and hosted as part of LIS	Map linkage & physical data in LIS. Port metadata into CMAP for comparison of genetic maps between peanut and other legumes. CMAP modified for seamless operation with LIS. Establish automated linkage of sequence-based markers to EST & genomic data in LIS	All publically available sequences (EST and genomic), expression array and protein information from peanuts will be integrated with other legume information and fully supported with comparative analysis tools in a web-enabled environment
6.1	Develop reliable non-rodent animal models for bioassay and mechanism testing	The pig immune system is more similar to humans than mice. Data from a swine model of food allergy is directly relevant to human response	Test gene expression profiles associated with inflammation and allergy with 100-gene RT-PCR arrays to profile pig immunological response to <i>Trichuris suis</i>	Develop pure-line swine populations selected for hypersensitivity to peanut proteins, and validate method with biomarkers and monoclonal antibodies to peanut proteins
6.2	Determine the threshold & kinetics of absorption of sensitizing proteins and allergen peptides in humans	International allergen threshold protocol has been established for allergen detection in foods. There are no credible estimates of the maximum levels of peanut proteins that can elicit allergenic reaction in humans	Access kinetics of allergen absorption into the bloodstream of non-sensitive humans. Collect serum IgE from peanut sensitive individuals.	Screen genetically modified peptides and proteins from TILLING and RNAi-allergen gene deletion libraries. Estimates of the maximum threshold dose that elicits reactivity to peanut proteins.

## INTRODUCTION

Cultivated peanut (*Arachis hypogaea* L.) is grown on 25.5 million hectares with a total global production of about 35 million tons; ranking among the top five oilseeds grown in the world. Peanuts originated in South America. Today, peanut is produced primarily in Asia (66.8%), Africa (24.6%), and North America (8.6%). India, China and the U.S. consume about 65% of world peanut production. Per capita, Americans consume more than 3 kg of peanuts and peanut products per year. During the 1990's, the U.S. peanut crop had an average value of over \$1 billion and ranked as the second-most important seed legume, after soybean. Peanut offers numerous human health benefits. The oil contains relatively high levels of oleic acid which confers high oxidative stability and reduces the need for hydrogenation. Hence, refined peanut oil does not contribute to the 'trans-isomer' content of foods. In addition to no trans-isomers, the low saturated fat content of peanut oil has been shown to lower LDL-cholesterol levels in the blood. Endogenous nutraceuticals such as resveratrol also may improve cardiovascular health. Peanut also is a good source of folic acid, essential vitamins and minerals. However, peanut is one of the primary food allergens and human sensitivity to peanut is growing at an increasing rate. Adverse allergenic reaction to peanut can be severe and at present little is known of the potential to either modify peanut to mitigate the allergenic risk or to improve prevention for individuals at risk. Genomics and gene manipulation has an important role to play in altering peanuts to lessen the rate of adverse reactions.

Given its economic and nutritional importance in the U.S. and abroad, peanut is virtually unexplored at the genomic level because the peanut genome is complicated (contains duplicate sets of chromosomes) and is large (2800 Mb or about the size of the human genome). Slow progress in overcoming these factors has impeded the development of intrinsic low-cost and environmentally benign solutions to many challenges that increase the cost and risk of peanut production. The application of genomic tools and information in breeding programs would greatly facilitate the genetic enhancement of cultivated peanut. The most urgent needs include the development of large numbers of user-friendly genetic mapping tools; sequencing of substantial populations of expressed sequences from diverse tissues and genotypes; the assembly of a genetically-anchored physical map and its alignment to the emerging genomic sequences of related legume crops; and the characterization of gene-rich regions to discover genes that influence peanut quality, nutritional value and productivity.

Recognizing the emerging need to address the longer-term aspects of this issue, the peanut genetics community participated in a July 2001 meeting at Hunt Valley, MD to draft a strategic plan for the future of legume genomics. The *U.S. Legume Crops Genomics White Paper* from that historic meeting provided a road-map that was based on consensus in six critical areas of genomic research that were needed to make progress across all legume species. These areas included (i) genome sequencing of strategic legume species, (ii) physical map development and refinement, (iii) functional analysis, (iv) development of DNA markers for comparative mapping and breeding, (v) characterization and utilization of legume biodiversity, and (vi) development of legume data resources. That document set a solid foundation for launch of a new era in legume genetics, the characterization and interpretation of the legume genomes.

In addition, the *U.S. Legume Crops Genomics White Paper* provided a framework for the peanut genetics community to form a coalition with other legume crop geneticists in the creation of the U.S. Legume Crops Genomics Initiative (USLCGI). This initiative represents a landmark collaborative effort among the best minds involved in genomic research on soybean, peanut, common- and dry beans, alfalfa, peas and lentils, and model legumes. Because of dynamic advances in genomic technology, USLCGI partners have agreed to develop strategic plans for genomic research relative to their crop. In that regard, the peanut community has rejoined to assess the current status of peanut genomics, define the future direction of this science, and to draft strategic performance plan for peanut genomics research. Therefore, this document outlines a consensus on priority research approaches and targets for the peanut genomics community. It also establishes benchmarks for the evaluation of performance of the peanut genomics community and other organizations with interests in collaborative genomics research.

## RESEARCH STRATEGIES FOR PEANUT GENOMICS

### GENOMIC TOOLS AND BREEDING METHODS

*Arachis hypogaea* is believed to be a species of relatively recent origin as compared to other crop species. Molecular data indicates that the domesticated peanut had a single event origin and that introgression from related species has been extremely limited. The two genomes can be differentiated by analyzing DNA markers, and the species *A. duranensis* and *A. ipaensis* are believed to be the ancestral diploids. Although a large amount of morphological variation is evident among accessions of cultivated peanuts, little molecular variation has been observed within this species. However, large amounts of variation have been observed among *Arachis* species. Two restriction fragment length polymorphism (RFLP) maps have been produced, one of a cross between two diploid species and the second between *A. hypogaea* and TxAG-6, a synthetic polyploidy derived from three diploid species. Molecular markers have been associated with only a few peanut diseases and insects. There is critical need to utilize DNA marker systems either from peanut or other legume crops that can enhance the knowledge base of peanut and facilitate crop improvement.

#### STP 1.1 Develop a molecular map of the peanut genome

Developing molecular maps for peanut is critical for identifying linkage relationships, comparative mapping among legumes, and better utilizing the genetic resources within the cultivated and related species. Although molecular variation has been observed among accessions of the cultivated peanut, extremely low levels of polymorphism are observed between individual genotypes. Thus, mapping the tetraploid peanut is extremely difficult. An alternative is to map closely related diploid species and then use this information for constructing a map of *A. hypogaea*. During the early 1990s a diploid map was developed for two A-genome species and 117 markers were associated into 11 linkage groups. Some remnant DNA may still exist from the analyzed plants, but the F<sub>2</sub> population no longer exists and further mapping with this material is no longer possible. A second map was produced between *A. hypogaea* and a complex allotetraploid derived from three species (two A-genome and one B-genome) and about 375 markers were associated into a map. However, mapping is complicated and interpretation of data is difficult when such a large number of species is in the pedigree.

Unfortunately, the existing molecular maps of peanut do not have sufficient numbers of DNA-markers to be highly useful for genetic studies. New populations will be required to adequately map the peanut genomes. To avoid complications associated with high levels of sterility, mapping within the A- and within the B-genomes at the diploid level will allow easier interpretation of data. Using progenitor species of *A. hypogaea* will also allow diploid maps to more easily be transferred to the tetraploid species.

*Baseline for 2004: A diploid map between A-genome species, and a tetraploid map with A. hypogaea and multiple diploid species have been created with several hundred markers*

*Goals for 2006: Develop populations for creating separate diploid maps with A- and B-genome species; develop populations for inbred line selection with different A. hypogaea varieties*

*Goals for 2008: Develop saturated genetic maps of the A- and B-genomes*

#### STP 1.2 Develop a comprehensive set of genetic markers for peanut

Early attempts to identify polymorphism among varieties and landraces using restriction fragment length polymorphism and randomly amplified polymorphic DNA (RAPD) markers failed. Simple-sequence repeats (SSRs) were only slightly more successful; but recent reports indicate that SSR markers have potential use for genetic studies of the cultivated peanut. The peanut molecular biology community needs to investigate additional marker systems to identify polymorphism in *A. hypogaea*. For example, single nucleotide polymorphisms (SNP) frequently occur within plant genomes, their mutation rate is low, and analyses can be accomplished by high throughput genotyping. On the other hand, there are large numbers of DNA marker polymorphism in the wild species of *Arachis*. Although there is evidence for chromosome structural differences in peanut, most linkage studies indicate that the genomes of different species in section *Arachis* are mostly collinear. Thus, data obtained from wild species will likely be applicable to *A. hypogaea*.

*Baseline for 2004: Very few SNPs exist for peanut*

*Goals for 2006: Develop 100 additional SNP's for peanut that are informative among elite genotypes.*

*Goals for 2008: Develop 1000 additional SNP's for peanut that are informative among elite genotypes.*

### **STP 1.3 Preserve DNA from core-collection accessions as a genetic resource**

The cultivated peanut collection contains more than 8,000 accessions in the U.S. and 14,000 at the International Crops Research Institute for the Semi-Arid Tropics. Evaluation of this large group of materials on a timely basis is not possible, so a core collection was developed as a subset to represent the range of variation within the entire collection. Information about the core collection has been assembled such that clusters of accessions can be analyzed to increase the probability of finding a desired trait. The peanut core collection has been used to identify areas where additional plant collections may be warranted to increase genetic variation, to identify accessions for resistance to leaf spots, nematodes, aflatoxin and several other diseases of peanut, and to identify genetic variation for oil content and fatty acid composition. Preserving DNA from core accessions will allow more efficient use of time and facilities to answer pertinent questions in molecular biology.

*Baseline for 2004: None available*

*Goals for 2006: Extract and preserve DNA from sample plants*

*Goals for 2008: Establish a repository and distribution system for DNA samples*

### **STP 1.4 Association of phenotypic traits with molecular markers**

Developing a system for marker-assisted selection for specific traits requires identification of germplasm with contrasting phenotypes, accurate evaluation techniques for phenotyping populations, identification of markers closely associated with the loci of interest, and technologies to facilitate rapid and cost effective screening of large populations. The high levels of variation within and among closely related *Arachis* species lead to their potential use for gene identification, marker assisted selection, and introgression to the cultivated species. Homologies between the genomes of *A. hypogaea* and related species have been estimated. For example, genes from *A. cardenasii* (an A-genome species) have been introgressed into 10 linkage groups of *A. hypogaea*. Hybrids from these crosses have been used to identify RAPDs and sequence characterized amplified regions (SCARs) to map genes conferring resistance to the peanut root-knot nematode. RAPDs have also been linked to several components of leaf spot resistance, to *Clindrocladium* black rot resistance, and to several insect pests. AFLP markers in other hybrids have been linked to tomato spotted wilt virus resistance. Although linkages of resistance genes to different molecular markers may prove useful for selecting breeding lines with desirable traits, there have been limited successes in peanut for utilizing these materials for cultivar development.

Examples of traits that have been associated with DNA markers in *A. hypogaea* include early and late leafspot, nematodes, 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 and drought tolerance that are difficult to measure in large segregating generations and substantially influenced by genotype-by-environment interactions may also benefit from marker-assisted selection. Many of the most difficult traits to acquire in a selection program are multigenic, and for several peanut pathogens there exists multiple components of resistance which must be combined into a single genotype. Crop improvement will require identifying appropriate sources of resistance and then developing segregating populations for inheritance studies and marker identification.

*Baseline for 2004: RFLP, RAPD, AFLP, SSR markers are associated with a few morphological traits, insect resistances, nematodes and peanut diseases*

*Goals for 2006: Develop mapping populations for quality or yield limiting traits*

*Goals for 2008: Position markers associated with variation in phenotypic traits on genetic maps*

### **STP 1.5 Develop genomic comparisons among legume species**

EST resources provide a means to conduct comparative functional genomic analyses across genera and plant families such as *Glycine*, *Medicago* and *Phaseolus*. This information will improve efficiency and solve complex problems in peanut. Studies can then be conducted to investigate features that distinguish the major legume lineages from one another. Studies to distinguish legumes from other well-studied plant groups, including the *Brassicales*, *Malvales* (cotton), and monocots then will be possible. Targeted sequences will be tagged to investigate traits unique to peanut, ones similar to other genera of legumes to compare species for gene discovery, and to investigate the origin of polyploidy and gene expression.

*Baseline for 2004: Few preliminary studies have been completed to compare peanut species with molecular tools*

*Goals for 2006: Test recognition of sequence-tag-sites in peanut with markers from other legume species*

*Goals for 2008: Develop reconstituted maps of the peanut genome with bioinformatics resources in LIS*

### **STP 1.6 Germplasm enhancement for quality and agronomic traits with molecular markers.**

A MAS system for selection for specific traits requires identification of germplasm with contrasting phenotypes, identification of markers closely associated with QTL (quantitative trait loci), and technologies to facilitate rapid/cost effective screening of large populations. Linkages of resistance genes to different molecular markers have demonstrated the value of selecting breeding lines with desirable traits. Further progress in improving the efficiency of peanut cultivar development is limited by the lack of more complete coverage of the gene-space in the peanut genome with appropriate molecular markers.

*Baseline for 2004: No active MAS projects in practice*

*Goals for 2006: Association of DNA markers with phenotypic and genotypic variation for specific traits*

*Goals for 2008: Demonstrate the utility of MAS methods in peanut for quality and agronomic traits*

## **PLANT TRANSFORMATION TECHNOLOGY**

Successful protocols for peanut transformation by both microprojectile bombardment and *Agrobacterium*-mediated gene transfer were developed over a decade ago. Since then, the bulk of peanut transformation work has been carried out by a small number of laboratories, primarily using microprojectile bombardment of embryogenic callus cultures. While this procedure is reliable, it is highly laborious, and nearly a year elapses between the initiation of the procedure and development of flowers on transgenic plants. However, through-put is severely limited by available resources. It is unlikely that all U.S. laboratories capable of peanut transformation, working in concert, could analyze more than about 10 genes per year.

Microprojectile bombardment has been used for two major reasons. First, ownership of intellectual property ownership of microprojectile technology is more clearly defined than for *Agrobacterium* based procedures. Secondly, naked DNA transfer procedures, such as microprojectile bombardment, can deliver selectable markers on separate plasmids from those carrying the genes of interest. This may eliminate selectable markers (typically antibiotic resistances) from the final transgenic product. This could be an important factor in mitigation of perceived environmental and health risks posed by the use of transgenic food crops, and could therefore facilitate the acceptance of transgenic peanuts in the marketplace.

It is essential to develop greater transformation capacity in order to study the many candidate sequences (genes and regulatory elements) that we anticipate a large-scale peanut genomics program will produce. This is most likely to come about with the development of improved microprojectile-based protocols. However, study of genetic elements does not involve the production of commercial products; and because *Agrobacterium*-based procedures appear to have the potential to generate transgenic events more efficiently than naked DNA procedures, it is also essential to direct resources toward the development of efficient *Agrobacterium* based transformation protocols for peanut.

It will also be important to develop transgenic methods for analyzing gene function. This could include development of methods for transforming peanut with large DNA sequences (e.g., from BAC libraries), and the development of tagging strategies that could facilitate the cloning of important genes by simultaneously interrupting normal function and by marking genes with identifiable molecular sequences. At present transfer of large DNA elements must be carried out by either microprojectile bombardment or microinjection.

### **STP 2.1 Improve transformation efficiency and utility for peanut**

*Agrobacterium*-based transformation has certain advantages compared with “naked DNA” transfer procedures, such as microprojectile bombardment. Chief among these are lower average transgene copy

number in recipient plants, and higher transformation efficiency when protocols are optimized. The issue of efficiency is especially important. Bombardment requires the direct introduction of DNA-bearing particles into cells to accomplish gene transfer. This approach is inherently limited because only a small proportion of treated cells actually receive transforming particles. *Agrobacterium*-based methods take advantage of the ability of the bacteria to infect much larger numbers of cells over a protracted infection period. If an *Agrobacterium*-based transformation technology could be employed reproducibly in peanut, it is likely that its efficiency could be higher than that currently achieved with the gene gun. For this reason, re-evaluation of *Agrobacterium*-based methods may be a critical step in the development of large-scale transformation efforts. Development of these protocols has lagged due to limited funding.

*Agrobacterium*-based transformation methods used for transformation of, e.g., soy and lupine, along with published accounts of *Agro.*-based transformation of peanut, will be re-evaluated. New protocols will be developed for peanut, taking into account all that has been learned regarding maximization of biolistic transformation efficiency, such as reduction of time in culture, use of appropriate selectable marker systems, etc. Emphasis will be on development of an appropriate test system, so the choice of recipient genotypes will be based on their performance in the transformation system, rather than their suitability for specific agronomic purposes. This is an important caveat, because at present virtually all peanut transformation is carried out to achieve specific agronomic goals, and thus choice of recipient variety is dictated by proposed application, regardless of its performance in the transformation system.

High-throughput analysis and evaluation capacities will be developed by laboratories involved in large-scale peanut transformation, regardless of the technology used to accomplish gene transfer. The principal resources required for this activity, in addition to appropriately trained personnel, will be greatly increased growth chamber capacity and greenhouse space, and certain specialized analytical instrumentation, at least some of which is often available in transformation laboratories.

*Baseline for 2004: Biolistic transformation protocols commonly used. Agrobacterium-mediated transformation is reported in literature, but not repeatable in most laboratories.*

*Goals for 2006: Establish Agrobacterium-based protocols for use in high through-put transformation in peanut. Scale-up current microprojectile-based protocols.*

*Goals for 2008: Provide the peanut research community with the capacity to achieve, analyze and evaluate a baseline of at least 100 transgenic lines per year.*

### **STP 2.2 Improve transformation technology for functional genomics**

In the decade since the first reports of successful peanut transformation, modest improvements have been made to gene-gun mediated transformation protocols that reduce the time required to recover fertile transgenic plants, and increase transformation efficiency. It seems unlikely that very large (i.e., order-of-

magnitude) increases in efficiency of these protocols will be accomplished in the near term. Therefore, short-term increases in transformation capacity will be accomplished most effectively by developing two or more high-throughput peanut transformation facilities capable of carrying out numerous simultaneous transformations, employing the most efficient microprojectile-based protocols available. These laboratories should be located at public institutions where peanut variety development is being carried out, in order to maximize information exchange among genomics researchers, growers and breeders.

To insure maximum productivity, steps must be taken to organize collaboration among these laboratories in order to eliminate duplication of effort, and to insure that modifications in transformation protocols are made available to all labs as quickly as possible. Further, production of transgenic peanuts carrying genomic elements of interest would ideally be made available as a service to the community of peanut genomic scientists. Regulatory approvals and appropriate agreements among collaborators will be needed to ensure efficient transfer of putative transgenic materials among collaborating institutions.

With relatively modest investment, it would be possible to increase the capacity of public-sector laboratories that are already capable of peanut transformation could be increased by a factor of ten within four years. This increased capacity is expected to result from both greater resource allocation, continued improvement of existing transformation protocols, and close coordination of the work carried out in all labs involved.

*Baseline for 2004: Ability to generate and evaluate transgenics expressing 10 constructs among entire community of peanut transformers*

*Goals for 2006: Expand capacity to test new gene promoters, selectable markers and terminators in peanut and other legume genomics programs.*

*Goals for 2008: Establish high through-put community peanut transformation facilities with expanded capacity to produce and distribute 100 transgenic lines per year to collaborators..*

### **STP 2.3 Develop transgenic screens to understand gene function.**

Transformation of peanut is an integral part of the study of its genome. Two areas of special interest are its use in the analysis of large (e.g., multi-genic) segments of DNA, and its ability to simultaneously inactivate and tag genetic sequences. These are broadly termed 'forward' (introducing new DNA) and 'reverse' (impeding endogenous gene expression) genetics.

Using current transformation protocols, it is not possible to introduce DNA sequences of more than a few kilobase pairs (< 10,000 bp) into cells. The ability to introduce far larger DNA molecules (>100,000 bp) such as bacterial artificial chromosomes (BACs) would be an important tool for the analysis of large genomic sequences, and could also have practical application in metabolic engineering of the species. We suggest a goal of this program should be the development of a transformation system capable of transferring these large sequences into peanut. At present it is unclear what technologies will be required, but experiments involving other species suggest that certain naked DNA technologies, including microinjection and, perhaps, a modified version of microprojectile bombardment, may be appropriate.

Transformation can be used to suppress the expression of targeted genes by post transcriptional silencing (RNAi), thus eliminating or reducing certain proteins. This approach also can be used to examine the collateral consequences of removing these proteins. In parallel, 'forward' genetic protocols should be developed to introduce highly expressed genes, i.e. proteins in which *IgE*-binding sites are removed.

Advances in the development of this technology will have their greatest utility if BAC libraries of peanut genomic sequences are prepared and thereafter become generally available to the peanut research community. Downstream analysis of large numbers of transgenic plants is also required in order to study the effects of newly-introduced genetic sequences. This is seen as an integral part of the transformation process because the laboratory generating a transgenic plant is typically best equipped to differentiate between the effects of expression (or knock-out) of a gene, and secondary effects resulting from the transformation process using both DNA array (transcription) and proteomic (translation) approaches.

Another goal of this program is to develop for use in peanut a 'reverse' genetic approach using either a transposon-based tagging system, such as the *Ac/Ds* system from maize, or an *Agrobacterium*-based

tDNA tagging system. In both systems, genes of interest are simultaneously interrupted and marked by the incoming DNA. Integration (or transposition) of the tagging DNA into a gene of interest disrupts its function, resulting in a null or altered phenotype which can be detected either by inspection or biochemical analysis. The tagged sequence can then be isolated, cloned and analyzed using standard molecular techniques.

It is currently feasible to deliver a transposon-based tagging system into peanut by microprojectile bombardment of embryogenic calli. Based on success of Ac/Ds tagging in a wide array of unrelated species, it seems likely that such a system could be successfully adapted for peanut. Development of an efficient tDNA tagging system is, of course, dependent upon the development of efficient, highly repeatable *Agrobacterium*-based transformation protocols.

*Baseline for 2004: RNAi and other gene suppression techniques largely untried for peanut. Impractical to introduce BAC clones into peanut. Tagging systems developed in theory.*

*Goals for 2006: Test use of RNAi and other gene suppression techniques to prevent synthesis of primary allergens. Test the ability of peanut to produce modified proteins lacking IgE-binding sites. Develop BAC clones and to initiate tDNA gene interruption.*

*Goals for 2008: Large scale tests of targeted gene suppression with analysis of collateral consequences. Optimize proteins production in peanut with the goal of providing maximum accumulation of alerted proteins such as those lacking IgE binding sites. Characterize and sequence selected BAC clones and to initiate working with transformation teams to introduce specific BACs that will test the capacity to undertake metabolic engineering. Evaluate heterologous AC/Ds and/or tDNA tagging systems.*

#### **STP 2.4 Develop biotech risk assessment and mitigation strategies.**

There are certain risks, both real and perceived, associated with the use of any transgenic germplasm in agriculture. Some include: unanticipated movement of transgenes from the modified crop into either weedy relatives or other, non-transformed varieties of the same crop; escape of transgenic plants that exhibit increased fitness, possibly allowing them to become troublesome weeds in the same or other crops; movement of transgenes, especially those conferring antibiotic resistance, into microflora that could subsequently infect humans, livestock or wild animal populations; and unexpected effects of the transgenic crop on the environment, by deleterious alteration of soil micorflora, insect populations, etc.

Further, a second kind of danger is posed by the possibility that the expression of transgenes might become unstable. For example, a transgene conferring resistance against an important insect pest could cease to be expressed, leaving the crop vulnerable to predation, resulting in significant financial loss.

Finally, peanut has recently been engineered to express a pharmaceutical protein. Experience with both maize and soybean suggests that the use of food crops for pharmaceutical production is inherently risky, because the transgenic crop could be mixed with the normal crop. This could expose consumers to foreign products, with potentially disastrous result. This would appear to be especially true with peanut because seeds are borne on underground structures, and some seeds typically remain in the field after harvest. Thus, it is appropriate to address these biosafety issues as a part of the overall effort in peanut genomics.

*Baseline for 2004: No known commercial deployment of transgenic peanuts. All risk associated with experimental releases. Little data exists regarding the probability of escape, gene movement, etc.*

*Goals for 2006: Identify major safety concerns, and derive experimental approaches to assess their importance.*

*Goals for 2008: Development management and/ or technological approaches for mitigation of risks.*

## GENOME SEQUENCING and GENE DISCOVERY

The nuclear genome of the cultivated peanut is approximately  $3 \times 10^9$  bp, similar to that of human. The peanut genome is anticipated to contain perhaps 50,000 genes that collectively determine its growth and development, interactions with the biotic and abiotic environment, and properties as a staple food crop. A detailed analysis of most of these genes, as well as their regulatory elements (promoters and enhancers) is essential to gaining understanding of the unique features that differentiate peanut from other organisms, and to realizing the full genetic potential of peanut as a healthful and profitable food crop. Polyploidy is a key consideration in crafting research approaches for cultivated peanut because two or more copies of many genes are present, resulting from the joining of two genomes from ancestors with divergent evolutionary histories.

### STP 3.1: Advance discovery and knowledge of regularly-expressed genes in peanut.

'EST' (expressed-sequence tag) sequencing, the random sequencing of clones from cDNA libraries, is the most efficient approach for gaining an initial picture of genes in the peanut genome. Currently, about 5,000 peanut ESTs are thought to exist in the public domain. EST sequencing from high-quality cDNA libraries representing a sampling of 10-20 diverse tissues and developmental stages, to depths of 5000-10,000 ESTs per library, is needed to obtain a first glimpse into the peanut gene set.

Sound justification exists both for concentration of effort on a few especially-important genotypes, and for sampling diversity associated with large numbers of genotypes that better represent needs of specific locales. To balance the respective merits of these approaches, we suggest community emphasis on two representatives each of the A, B, and tetraploid genomes. In each case, the two genotypes should be genetically cross compatible to enable production of mapping populations. One of the A genome genotypes should represent *A. duranensis* (i.e. accession 10038), the diploid that most closely resembles the A genome donor to cultivated peanut. One B genome genotype should represent *A. ipaensis* (i.e. acc. 30076), which most closely resembles the B genome of cultivated peanut. One of the component lines of the cultivar Florunner has been used in constructing BAC libraries and would be ideal for broader community use in view of its widespread historical importance. However, it is recognized that it is growing more difficult to obtain large quantities of seed from the original stocks of Florunner, may necessitate occasional deviations from this preference. Possible alternatives to Florunner for some experiments are being explored by the community. In any case, the very low level of DNA polymorphism in peanut (especially within botanical varieties) suggest that information from other elite peanut genotypes will often be highly transferrable, and that research should not be hindered by the lack of a specific genotype as a gold standard. A (to-be-identified) genotype representing Spanish peanut types would be a logical complement at the tetraploid level.

In all cases, it is highly desirable that original chromatograms remain accessible to the community, as these are an essential complement to the data types archived in Genbank that permit robust methods to be applied to Genbank data to extract additional information such as SNPs from raw sequence.

*Goals for 2004: 5,000 ESTs emphasizing seed traits to provide a basis for allergenicity objectives. The initial goal should be to assure that all genes encoding allergen isoforms are identified and available as complete unigenes.*

*Goals for 2006: To expand coverage to 50,000 ESTs submitted to Genbank, together with public availability of chromatograms. Providing a comprehensive seed set and coverage of other important other characteristics such as disease resistance and stress response. To produce seed-specific oligo based arrays (500-1000 sequences) for community use.*

*Goals for 2008: To attain deep coverage of 150,000 ESTs submitted to Genbank, together with public availability of chromatograms that provides a genetic basis analyzing peanuts as a crop and to support a wide range of genetic modification. To use the larger scale EST to produce oligo arrays that encompass a wide variety of agronomic traits*

**STP 3.2. Advance discovery and knowledge of rarely-expressed genes in peanut, together with regulatory sequences that determine the levels and patterns of gene expression.**

While EST sequencing is a first step in gene discovery, only a fraction of the genes found in an organism are expressed in any single source tissue. Even by studying cDNA libraries from multiple tissues (as described above), diminishing returns typically accrue after about  $10^5$  sequences. In particular, many genes expressed only rarely or at low levels are likely to be missed, such as the genes that are responsible for the initiation of major developmental ‘cascades’ that have widespread effects on plant growth and development. Finally, EST sequencing yields little if any information on regulatory sequences or other important low-copy elements.

Several methodologies now exist that may advance discovery of rarely-expressed genes, and also gene regulatory sequences. Modifications of EST sequencing, based on ‘subtraction’ of known sequences, or ‘normalization’ of EST libraries may warrant exploration.

Two approaches have recently emerged to sift probable genes out of genomic DNA (which is largely comprised of repetitive DNA). Methyl filtration, based upon degrees of differential methylation of expressed versus non-expressed sequences, reduces the abundance of repetitive DNA in plant genomic DNA libraries. However, this approach is subject to variable relations between methylation and gene expression across genes, taxa, and physiological states. Cot-based cloning and sequencing involves the fractionation of a genome into ‘components’ based on the degree of sequence repetition, providing access to the entire genome (including regulatory and repetitive elements), and with flexibility to a wide range of permutations based on the biology of the system and the goals of the investigator.

*Goals for 2004: virtually no data*

*Goals for 2006: Exploratory – evaluate methodologies in a sampling of prioritized genotypes, noting that diploid genotypes may be preferable by virtue of inherently-lower complexity*

*Goals for 2008: Begin to apply subset of methodologies*

*[Goals for 2010: Capture sequence complexity of the transcriptome, preferably for one A-genome diploid and also one B-genome diploid.]*

**Step 3.3 Integration of diverse tool sets into a picture of the peanut genomic landscape.**

In lieu of comprehensive study of the entire peanut genome sequence, interim steps that confer incremental benefits include a ‘gene map’, which plots the locations of ESTs (or clusters thereof), on scaffolds of BAC (bacterial artificial chromosome) clones that have been aligned to one another and also to the peanut chromosomes. Such a gene map provides many of the benefits associated with a completed genomic sequence, but at much less cost. Further, a gene map provides reference points that will help in ultimate assembly of a peanut genomic sequence.

In developing information on the structural organization of the peanut genome, it is important to recognize the complexities that result from the polyploidy of cultivated peanut. The combination of two similar ancestral genomes into the nucleus of modern peanut increases the complexity of the peanut genome exponentially, ie. not just doubling, but perhaps quadrupling the time and cost associated with ‘assembling’ the peanut genome. Thus, analysis of diploid progenitors may considerably reduce the cost and accelerate progress toward a complete picture of the genetic potential of peanut. In particular, good-quality BAC libraries for diploids *A. duranensis* and *ipaensis* will be important (for reasons addressed above). One of these should be BIBAC-based, to facilitate many potential functional genomic studies.

The genomic sequences of *Medicago* and Lotus are likely to precede such an effort in peanut, and a comprehensive set of soybean genes obtained from gene-rich regions may also be known. Hence, it is logical to investigate the degree to which comparative approaches might be used to prioritize peanut genomic regions for a future sequencing pipeline. It is highly likely that a comprehensive set of soybean genes obtained from gene-rich regions will also be known. Together the Lotus, *Medicago* and Soybean efforts will provide considerable resources for comparative genomics of peanut.

*Baseline 2004: EST availability has been described above. A 6-genome-equivalent BAC library exists for one Florunner component line.*

*Goals for 2006: Additional characterization of the tetraploid BAC library toward identification of gene-rich BACs*

*Goals for 2008: Anchoring with a much more detailed genetic map than already exists; also have available diploid BAC libraries that could be anchored at the same time.*

*[Goals for 2010: Sequence-ready genetically anchored physical maps for the tetraploid and component diploids.]*

## **FUNCTIONAL GENOMICS & PROTEOMICS**

Proteins are the structural and enzymatic products encoded by genes. Although gene expression on a genomics scale can be examined at the mRNA level using micro-arrays, proteins can contain post-translational modifications that affect their abundance and activity. High-resolution protein separation on two-dimensional gels followed by mass spectrometry of excised protein spots allows the rapid profiling of gene expression at the protein level. In peanuts, little genomic information is yet available, i.e., no large-scale EST projects have been initiated, and it is premature to focus on a large proteomics effort. However, the DNA sequence information that is anticipated from peanut genomics projects would provide the information that is necessary to define gene function using a reverse genetics approach.

### **STP 4.1 Proteomics research to characterize seed protein composition and function**

One area of peanut research that could greatly benefit from proteomics is the extensive characterization of seed proteins. Peanut seeds are consumed by humans and many of the seed proteins have been shown to generate an allergic response in humans. The extent of genetic variability among seed protein profiles has not been explored with high-resolution methods.

*Baseline for 2004 – Preliminary proteomic information on number, type, accumulation, & identity of peanut proteins that elicit allergenic response in sensitive humans.*

*Goals for 2006 – Generate a high-density proteomic map of developing & mature peanut seed that includes at least 1000 peptides from the most abundant seed proteins. Initiate fingerprint patterns for allergenic proteins. Characterize changes in protein composition & allergenicity in seed of peanuts with genetically modified storage protein gene families.*

*Goals for 2008 – Generate a reference proteomics map for leaves. Expand capacity to include a wide range of proteins associated with agronomic characteristics or plant responses to biotic & abiotic stress.*

### **STP 4.2 Application of reverse genetics to explore gene function**

EST sequencing projects will provide an abundance of DNA sequence information from expressed genes, some of whose functions can be predicted based on sequence similarity. The prediction of function would not, however, fully describe the role of a particular gene in peanut growth and development. Gene function could be definitively determined by the creation of a series of mutant alleles with evaluation of resultant phenotypes. Since cultivated peanut is a tetraploid, and gene function may not simply be additive with respect to the combination of two putative progenitors of *A. hypogaea* into one species, it would be most informative to create mutants in cultivated peanut. An effective genome-based approach to exploring the function of genes is called TILLING (Targeting Induced Local Lesions in Genomes). This method relies on high-throughput screening at the DNA level for mutations in genes. A TILLING population for peanut should be developed as a public resource.

*Baseline 2004 – Funding for a pilot project on TILLING has been provided in order to support the testing of mutagenesis protocols.*

*Goals for 2006 – Establish a facility for high-throughput TILLING in peanut and initiate the development of a large TILLING population.*

*Goals for 2008 – Advance the TILLING population to M3 families and initiate large-scale screening for mutations in specific genes.*

## **BIOINFORMATICS**

*Bioinformatics* consists of data management (acquisition, storage, integration and dissemination) and data interpretation (data analysis, visualization and biological modeling). The scope of bioinformation being developed for model and crop legumes includes DNA sequences, RNA expression levels, protein interactions, map positions and QTL. However, until robust databases are developed for peanut, bioinformation must be leveraged from other plant species, especially legumes, to advance knowledge about peanut. This may require the ability to define gene functions with bioinformation from related legume species such as soybean, *M. truncatula* or *L. japonicus*. Further, it may require the ability to find gene products that can be associated with biochemical pathways or networks, allowing researchers to discover the molecular basis for phenotypic traits from related species.

The Legume Information System (LIS), a Congressionally mandated joint collaboration between USDA-ARS and the National Center for Genomic Research (NCGR), was conceived to be a comparative legume resource, and a first step towards leveraging model plants to gain insights into crop species. LIS is a data management system that includes an annotation pipeline of data analysis tools and legume (soybean, peanut, alfalfa, pea, dry beans) data bases developed by multiple groups. LIS is populated initially with data from *G. max*, *M. truncatula* and *Lotus japonicus*. A major bioinformatics goal is to develop a robust means of comparative transcript analysis, initially between *G. max*, *M. truncatula* and *L. japonicus*, and eventually unigenes from *Arabidopsis thaliana* as a non-legume for comparative purposes. The results of comparative transcript analysis should be immediately useful for comparative genomic analysis for peanut-specific gene expression.

Since inception, the development of LIS has been guided by the legume research communities. Future advances also require collaborative relations with other plant bioinformation groups to ensure representation in plant community defined ontologies and controlled vocabularies. Recognizing the power of comparative genomics among species to identify candidate genes, unique genes, and evolutionary relationships among genes for crop improvement, a Peanut Information Resource (ArachisDB) should be developed and should become a component of LIS.

### **STP 5.1 Establish collaborative relationships with other legume research communities.**

At the ontological level peanut researchers have much in common with other legume researchers. From a bioinformatics perspective, the utility and relevance of peanut information resources will be affected by the ability to exchange and compare bioinformation with other plant resources. However, collaborative relations must be developed with other legume information resources to ensure that peanut genetic bioinformation is represented in plant community defined ontologies and controlled vocabularies.

*Baseline for 2004: No peanut presence in the trait and plant ontology efforts.*

*Goals for 2006: Assure inclusion of peanut concepts in the Plant and Trait Ontology Efforts.*

*Goals for 2008: Assure implementation of peanut specific genes and traits in controlled vocabularies for all plant biology information resources.*

### **STP 5.2 Establish a Bioinformatics Oversight Committee**

A Steering Committee of peanut researchers and bioinformaticists is needed to guide the development of a Peanut Information Resource and assure relevant information is incorporated into LIS. This committee will solicit input on the perceived needs of the peanut research community, which will directly influence the system's design and user interface development. Communication will be facilitated by periodic convening of user panels in workshops and by providing a forum for online comments.

*Baseline for 2004: Established the peanut bioinformatics oversight committee. Designation of a group to develop the Peanut Information Resource.*

*Goals for 2005: Establish a curator from the peanut research community who will be responsible for assuring data and information content of the Peanut Information Resource.*

*Goals for 2006: Initiate plans to integrate the Peanut Information Resource into LIS.*

### **STP 5.3 Assure Peanut Genomic Information becomes part of the Legume Information System**

With advances in the Peanut Genome Initiative, a distinct peanut information resource will be needed to facilitate the storage and use of bioinformation for peanut. The Peanut Information Resource eventually will be technically integrated with LIS.

The success of a biological information resource is assured if the research community served exercises a sense of ownership in the system. The potential for high quality peanut information and subsequent comparative genomics approaches will be greatly enhanced if peanut researchers have ownership and control over information resource that are distinctly *Arachis*. Subsequent or concomitant incorporation of peanut data into LIS does not preclude such ownership, but greatly enhances the utility of the peanut information resource.

*Baseline for 2004: There is no singular Peanut Information Resource, peanut genetic and genomic information is stored among diverse resources. A web enabled community bulletin board for peanuts was built and hosted as part of LIS.*

*Goals for 2005: Begin design of ArachisDB and expansion of the LIS design to accommodate peanut information. Utilize the NCGR automated annotation pipeline for annotation of ESTs, construction of contigs and deposition of information into ArachisDB and LIS. Additionally LIS will import peanut Map data (linkage and physical) and associated metadata (authors, affiliations, literature etc.) into the relational CMAP database and visualization software to enable comparison of genetic maps between peanut and other legumes. CMAP has been modified to interoperate seamlessly with LIS. Established automated linkage of sequence-based markers to EST and genomic data housed in LIS.*

*Goals for 2006: All publicly available sequence(EST and genomic), expression array and protein information from peanuts will be integrated with other legume information and fully supported with comparative analysis tools in a web-enabled environment.*

*Goals for 2008: Anchored physical map with associated annotations included in LIS. Tilling and functional genomic information from peanut included in LIS. Linkages between LIS and Immunological Response Databases established.*

## **IMMUNOLOGY OF PEANUT PROTEINS IN MODEL SYSTEMS**

Hypersensitivity to dietary peanut proteins has not been critically evaluated. The expression of clinical disease follows a pattern of peanut sensitization and transitory hypersensitivity to subsequent exposure, but the generalized intestinal response is common to several different forms of food hypersensitivity. Some investigations suggest that neonatal immune development with the production of anti-peanut protein IgG/IgM antibodies and subsequent complement-mediated epithelial cell injury is responsible, while others indicate hypersensitivity based on IgE-mediated immediate intestinal anaphylaxis. It is essential to define dietary peanut intolerance in order to develop appropriate control strategies.

### **STP 6.1: Develop reliable non-rodent animal models for bioassay and mechanism testing**

A survey of 55 immunological variables in pig, human and mice, showed that the pig immune system is more similar to humans than mice for >80% of the variables compared. Thus, data derived from a swine model of food allergy is more directly relevant to human than data derived from rodents. A practical pig model to study peanut food allergy will provide procedures that alleviate or reduce clinical disease, and provide a means to evaluate food intolerance to other legumes and extrapolate to human food allergy.

*Baseline for 2004: The pig immune system is more similar to humans than mice. Data from a swine model of food allergy is directly relevant to human*

*Goals for 2006: Test gene expression profiles associated with inflammation and allergy with 100-gene RT-PCR arrays to profile pig immunological response to *Trichuris suis**

*Goals for 2008: Develop pure-line swine populations selected for hypersensitivity to peanut proteins, and validate method with biomarkers and monoclonal antibodies to peanut proteins*

**STP 6.2: Use threshold estimates and absorption kinetics to confirm the allergenic potential of peanut proteins.** International allergen threshold protocol has been established for allergen detection in foods. However, additional investigations are needed to develop greater confidence levels in estimates of the maximum levels of peanut proteins that can elicit allergenic reaction in humans. Such information will guide genomic strategies for developing peanut cultivars with reduced allergenic properties.

*Baseline for 2004: International allergen threshold protocol has been established for allergen detection in foods. There are no credible estimates of the maximum levels of peanut proteins that can elicit allergenic reaction in humans*

*Goals for 2006: Assess kinetics of allergen absorption into the bloodstream of non-sensitive humans. Collect serum IgE from peanut sensitive individuals.*

*Goals for 2008: Screen genetically modified peptides and proteins from TILLING and RNAi-allergen gene deletion libraries. Estimates of the maximum threshold dose that elicits reactivity to peanut proteins.*

**APPENDIX - Process**

On March 23, 2004, twenty-six scientists with expert knowledge of the critical fields of peanut genomics participated in a workshop hosted by the Peanut Foundation/American Peanut Council in Atlanta GA. These scientists reviewed the current status of peanut genomic research and reached consensus on the most important research priorities for the next five years. The resultant Strategic Plan for Peanut Genomics 2003 to 2007 provides a framework for outlining those research priorities and significant near-term milestones that represent 'quantum leaps' in the advancement of the science of peanut genomics.

The workshop was planned by: Howard Valentine, Director, Research & Technology, Peanut Foundation/American Peanut Council, H. Thomas Stalker, Head, Crop Science Department, North Carolina State University; and Richard F. Wilson, National Program Leader, USDA, Agricultural Research Service. Dr. Wilson facilitated the workshop. We wish to acknowledge the assistance of H. Thomas Stalker, Arthur K. Weissinger, Andrew H. Paterson, Peggy Ozias-Akins and William Beavis in the preparation of the initial draft of the strategic plan.

## Peanut Genome Initiative Steering Committee Contact Information

Eliot M. Herman  
 USDA-ARS  
 Donald Danforth Plant Science Center  
 975 N. Warson Rd., St. Louis MO 63132  
 314-587-1292  
[eherman@danforthcenter.org](mailto:eherman@danforthcenter.org)

Peggy Ozias-Akins  
 NESPAL, University of Georgia-Tifton  
 P.O. Box 748, Tifton GA 31793  
 229-391-2594  
[ozias@tifton.uga.edu](mailto:ozias@tifton.uga.edu)

H. Thomas Stalker (Chairperson)  
 Dept. of Crop Science  
 North Carolina State University  
 100 Derieux Place, Raleigh, NC 27695-7620  
 919-515-2647  
[tom\\_stalker@ncsu.edu](mailto:tom_stalker@ncsu.edu)

Howard Valentine (ex officio member)  
 Peanut Foundation / American Peanut Council  
 Alexandria, VA 22314  
 1500 King Street, Suite 301  
 706-579-1755  
[pnuttech@alltel.net](mailto:pnuttech@alltel.net)

N.C. Nielsen, USDA-ARS  
 2-333 Lilly Hall of Life Sciences  
 Agronomy Dept., Purdue University  
 W. Lafayette, IN 47907-1150  
 765 494-8057  
[nnielsen@dept.agry.purdue.edu](mailto:nnielsen@dept.agry.purdue.edu)

C. Corley Holbrook  
 USDA-ARS,  
 MOORE HWY, TIFTON, GA, 31793  
 Phone: (229) 386-3176; Fax: (229) 386-3437  
 Email: [holbrook@tifton.usda.gov](mailto:holbrook@tifton.usda.gov)

C. Victor Nwosu  
 Masterfoods USA  
 800 High St., Hackettstown, NJ 07840  
 Phone: 908-850-7545; Fax: 908-850-2697  
 Email: [victor.nwosu@effem.com](mailto:victor.nwosu@effem.com)

Bill Beavis  
 National Center for Genomic Research  
 2935 Rodeo Park Drive East  
 Santa Fe, NM 87505  
 800-450-4854  
[wdb@ncgr.org](mailto:wdb@ncgr.org)

Pat Donahue, Planters Division  
 Kraft Foods, Inc.  
 40906 10<sup>th</sup> Str., PO Box 215  
 Greenfield, CA 93927  
 831.674.3131  
[Pat.Donahue@kraft.com](mailto:Pat.Donahue@kraft.com)

Andrew H. Paterson  
 228 Plant Genome Mapping Laboratory  
 University of Georgia, 111 Riverbend Rd.  
 Athens, GA 30602-6810  
 706.583-0162  
 Email: [paterson@uga.edu](mailto:paterson@uga.edu)

Richard F Wilson (ex officio member)  
 USDA-REE-ARS-NPS-CPPVS  
 Room 4-2214, GWCC  
 5601 Sunnyside Ave.  
 Beltsville, MD 20750-5139  
 301-504-4670  
 Email: [r fw@ars.usda.gov](mailto:r fw@ars.usda.gov)

Mark D Burow  
 Lubbock R&G Base  
 Texas Agricultural Experiment Station  
 806 746 6101  
[mburow@tamu.edu](mailto:mburow@tamu.edu)

Sandra M. Allina  
 1304 WILLIAMS HALL, Crop Science Dept  
 NC State University  
 100 Derieux Place, Raleigh, NC 27695-7620  
 919-513-4647  
[sallina@unity.ncsu.edu](mailto:sallina@unity.ncsu.edu)

Wesley Burks, M.D.  
 Department of Pediatric Allergy-Immunology  
 Room 304, Jones Building  
 Box 3530 Medical Center, Durham, NC 27710  
 919 681 2949  
 Email: [burks010@mc.duke.edu](mailto:burks010@mc.duke.edu)

Gretchen Flanley (ex officio member)  
 USDA-ARS-OIRP  
 Room 4-1124B GWCC  
 5601 SUNNYSIDE AVENUE  
 BELTSVILLE, MD, 20705  
 301-504-4547  
[gef@ars.usda.gov](mailto:gef@ars.usda.gov)

Soheila J Maleki  
 USDA-ARS-SRRC, Room 2130  
 1100 ROBERT E. LEE BLVD  
 NEW ORLEANS, LA, 70124-4305  
 504-286-4590 ext. 590  
[sjmaleki@src.ars.usda.gov](mailto:sjmaleki@src.ars.usda.gov)