# Research Proposal: Global Valencia Peanut Niche Market Development Program

## Principal Investigator

Naveen Puppala
New Mexico State University
2346 State Road 288
Clovis, NM USA 88101
Phone: +1 (575) 985-2292 ext 103
Fax: +1 (575) 985-2419
Email: npuppala@nmsu.edu

## Co-Project Investigator(s)

Curtis Jolly
Auburn University
Department of Agriculture Economics
Auburn, AL 31443
Phone: (334) 844-5583
Email: cjolly@auburn.edu

Carel Ligeon
Phone: (334) 244-3486
Email: clgeon@aum.edu

Amade Muitia
Mozambique Agriculture Research Institute
Northeast Zonal Center
Nampula Research Station
Av. FPLM, Via Corrane, Km7
Nampula, Mozambique

## Geographical Locations

Kenya, Uganda

## Submitted

09/30/2007

## Focus

Domain - Production Values Region - Global

## Background

i) Screening for end-of-season drought tolerance: The Valencia peanut crop grown in New Mexico matures in about 120 days that this crop duration can be partitioned into three distinct phases of crop development: early crop

establishment (up to 40 days after sowing; DAS); flowering, pegging and early pod initiation and development (41 to 80 days from DAS); and pod/seed development and maturity (81 to 120 days from DAS). It is the end-of-season drought which coincides with the pod/seed development and maturity that causes substantial yield reduction in peanut.

It is proposed to evaluate Valencia peanut core collection for physiological traits associated with drought tolerance under end-of-season drought conditions at Agricultural Experiment Station, University of Arizona, Yuma, Arizona. Yuma is an ideal location for drought screening work as it is located in a desert with an annual rainfall of 75 mm. Most of the rainfall occurs mainly during late autumn to early spring with an average temperature of 29 to 33o C during July to September (Holbrook et al. 1994). It is proposed to withhold irrigation from 80 DAS until the crop is harvested to impose end-of-season drought treatment; however, control plots will receive the full irrigation for comparisons with drought stressed plots. Low cost movable rainout shelter will be developed to avoid rains when plots exposed to end-of-season drought treatment (Chauhan et al. 1997). Observations on physiological traits (in addition to plot yield) such as specific leaf area (SLA), specific leaf nitrogen (SLN), SPAD chlorophyll meter reading (SCMR), transpiration efficiency (TE), and harvest index (HI) will be recorded both from irrigated control and drought stressed plots. SCMR provides a rapid assessment of SLA and SLN as surrogate trait to measure TE in peanut (Nageswara Rao et al. 2001; Sheshshayee et al. 2006).

(ii) Screening for resistance to pod rot: A sick plot for pod rot complex diseases has been developed at South Research Facility (SRF), Clovis, New Mexico. Initially, all the accessions will be evaluated for resistance to pod rot complex diseases under field conditions, and only the promising accessions will be screened for pod rot resistance (Godoy et al. 1984) in micro plots at SRF. In field and micro plot evaluations, the test entries will be scored on 0 to 10 scale, where 0 is no decay and 10 is total pod decay (Pattee et al. 1974). Inoculum of each pathogen will be produced using cornmeal-sand medium (Garren 1970).

(iii) Evaluating core collection for agronomic and seed quality traits: The core collection will also be evaluated in replicated trials for agronomic and seed quality traits at Brownfield (West Texas). Observations on maturity (based on thermal degree days), pod yield, shelling percentage, percent sound mature kernels, 100-seed weight, and seed quality traits (oil and protein contents, fatty acid composition, and total sugars) will be recorded. Seed quality testing will be done at J. Leek Associates, commercial lab located in Brownfield, Texas or at the Texas Agricultural Experiment Station (TAES) Lubbock Center. (iv) Genotyping core collection (using SSRs and high throughput assay) to dissect genetic structure and marker-trait association: DNA will be extracted using the Qiagen DNeasy Plant mini kit (Qiagen Inc., Valencia CA) following the manufacturer's protocol. Molecular profiling will be done using Beckman CEQ8000 Analyzer (Fullerton, CA) at the TAES or an ABI 3100 (Applied

Biosystems, Foster City, CA) at the USDA-ARS Cropping Systems lab, and amplification products will be scanned using software provided by the manufacturer (CEQ STR [Short Tandem Repeats]) module or ABI Genescan 3.1.2 and scored as binary data. We have devised a scoring protocol to identify true bands from spurious peaks caused by stuttering of Taq polymerage (Kottapalli et al. 2007).

Statistical analysis will be done following DARwin 5.0 version (Perrier et al. 2003) for dissecting the genetic structure, PowerMarker 3.0 for determining gene diversity, rare alleles, common alleles and group specific alleles (Liu and Muse 2005; http://www.powermarker.net). Genetic distances will be calculated using Chord's distance formula (Cavalli-Sforza and Edwards 1967) as it generates correct tree topologies regardless of the microsatellite mutation model (Takezaki and Nei 1996) and the matrix thus generated will be used to construct dendrogram based on Neighbor Joining (NJ) clustering procedure implemented in NTSYSpc2.2 (Rohlf 1998; Exeter Software, Setauket, NY). Bootstrapping will be carried out by performing 1,000 bootstrap resamplings using PowerMarker and Winboot (Yap and Nelson 1996). All analysis will be performed using the MXCOMP module of the NTSYSpc software 2.2.

The extent of linkage disequilibrium will be estimated using TASSEL (http://www.maizegenetics.net/index.php?page=bioinformatics/index.html). Population structure will be determined using STRUCTURE (Pritchard et al., 2000) (http://pritch.bsd.uchicago.edu/structure.html), relative kinship (Loiselle et al., 1995) using SPAGeDi (http://www.ulb.ac.be/sciences/ecoevol/spagedi.html). Association mapping will be performed using the mixed Q+K model (Yu and Buckler, 2006) method in TASSEL (http://www.maizegenetics.net/index.php?page=bioinformatics/index.html) or SAS (SAS Institute, Cary, NC).

(v) Transcriptional profiling and statistical analysis of microarray data: The laboratory of collaborator Dr. Payton is currently developing oligonucleotide microarrays for expression profiling in peanut. Baseline expression profiling studies on one line identified as stress tolerant and one line identified as stress-sensitive will directly compare: a) control (full irrigation) to drought-stressed leaf and root tissue and b) stressed leaf and root tissue from tolerant and sensitive lines (outlined in Figure 1) (Yang and Speed 2002). We will use a modified dye-swap experiment for each comparison (Kerr et al. 2002). Briefly, for each pair-wise comparison (control vs. stress or tolerant vs. sensitive) 6 RNA extracts will be obtained from independent, pooled tissue samples. The RNA extracts from each tissue pool will be used as template to prepare six corresponding cDNA populations for each stress state (three cy3 and three cy5). Six hybridizations will be carried out for each specific comparison for a total of 36 hybridizations (18 for leaf and 18 for root).

## Technical Review

Our goals are that we will investigate the physiological and genetic factors that control stress and use this information to develop strategies for the genetic improvement of crop varieties. Improved drought tolerance of crop plants is vital to the sustainability of agriculture in USA and host countries. The principal aim of this collaborative project is to identify genetically diverse germplasm with beneficial traits, dissect the physiological and molecular basis of those traits, and establish marker/trait association.

Major traits of significance in peanut include yield, drought tolerance, and diseases (Dwivedi et al. 2003). Pod rot is a serious constraint to Valencia peanut production worldwide. Additionally, it also reduces the quality of edible grade peanut. Valencia peanuts with distinct bright color (<25% discoloration) are paid an additional premium price of US $ 50/ton by the New Mexico processors while dark colored peanuts with black hull and pod rot diseases receive trade penalty (i.e. low price and possibly rejected).

Peanut is widely used in the food and confectionery industry due to its high nutritive value. Dwivedi and Nigam (2003) detected variation in pod and seed characteristics and sweetness among 33 Valencia germplasm. Similarly, variations are reported when a set of Valencia peanut germplasm lines assessed for differences in molecular profiles using SSRs. For example, Krishna et al. (2004) classified 48 Valencia peanut germplasm into four distinct groups. These limited studies demonstrate that it is possible to identify diverse germplasm from subsps fastigiata var. fastigiata using morphological, agronomical and marker profiling data. In peanut, there are over 500 SSRs available in the public domain that will be used in this project (Hopkins et al. 1999; He et al. 2003, 2005; Ferguson et al. 2004; Moretzsohn et al. 2005). Of these, 170 markers have been mapped into 11 linkage groups covering 1,230.89 cM peanut genome (AA genome) (Moretzsohn et al. 2005). More recently, the BB genome based genetic map has also been constructed in peanut (DJ Bertioli, Univerisidade Catolica de Brasil, Brazil, pers. comm.). Conventional genetic linkage mapping approaches for polygenic traits are confounded by epistasis (adaptation and phenology traits influencing the target trait) and genotype by environment interaction (reducing the accuracy of phenotype data) that erodes the precision and power of QTL (quantitative trait loci) detection. In addition, linkage mapping has two other major constraints, particularly affecting practical applications: (i) marker-trait associations determined in genetic populations must be validated in target breeding populations before routine application can be considered which is time consuming and often introduces a major level of redundancy into the process, and, (ii) marker-trait associations identified in this way are based on genetic distance in the mapping population and tight linkage (and thus power of selection) may be eroded or lost entirely when the marker is applied to breeding populations with very different recombination patterns between the target loci and marker. Association mapping (AM), also known as linkage disequilibrium (LD) mapping, is a method that relies on linkage disequilibrium to study the relationship between phenotypic variation and genetic polymorphism (Flint-Garcia et al. 2003). LD refers to non-random association

between two markers, or two genes or between a gene and a marker locus. Mutation, population structure, epistasis, population perturbations like migration, inbreeding and selection all influence LD, and some of these can lead to spurious associations (Jannink and Walsh 2002); and thus an understanding of the structure of the population and extent of linkage disequilibrium are necessary to avoid false associations. AM deals with unrelated individuals or members of a family with varying levels of phenotypic expression that are evaluated to detect and measure the degree of association between molecular markers and traits of interest. The principal advantage of this procedure lies in its ability to capture informative data stored in unrelated individuals who have undergone several rounds of gene shuffling over multiple generations. Significantly, it can be used on material offering better overall relevance to breeding programs and thus reduce the level of redundancy between marker identification and marker validation steps. AM can be investigated using candidate genes as well from randomly chosen molecular markers that are evenly distributed across genome. In recent years, several studies conducted in plants have detected DNA markers associated with disease resistance and quantitative traits (Thornsberry et al. 2001; Sun et al. 2001; Ivandic et al. 2003; Gebhardt et al. 2004; Kraakman et al. 2004;), demonstrating that it is a viable alternative to classical QTL analyses. In addition, many of the associated markers were located in chromosome regions previously identified as harboring QTL for yield and yield components, providing good validation that AM of diverse germplasm is a viable alternative to classical QTL analyses based on crosses between inbred lines (genetic populations), especially for complex traits (Szalma et al. 2005; Breseghello and Sorrells 2006; Kraakman et al. 2006).

Research examining peanut physiological response to developing moisture stress and the recovery response to supplemental irrigation could have profound impacts on increasing water- use efficiency in production. Information gained from identifying signals of moisture stress or quantifying actual seasonal water use would improve current peanut irrigation scheduling methods dramatically. In addition, examining the variation in the molecular response to drought response among current peanut breeding lines would be crucial to developing more U.S. water- use efficient and drought tolerant varieties. This project evolved based on our shared interest in the improvement of peanut for breeding and we feel that our respective strengths in agronomy, whole-plant physiology, plant breeding and molecular biology will allow us to successfully meet our goals.

## Problem Statement

Valencia peanuts constitutes the niche crop and are known for their taste and having three or more seeds per pod. Valencia peanuts has the highest score for sweet attribute of any line tested (Pattee et al., 2001). The area under Valencia is very less compared to other market types mainly due to low yield potential and non-existence of drought and disease resistant Valencia peanuts. In U.S. the Valencia peanuts are grown as in-shell market predominantly in eastern New Mexico and west Texas. Average yield in this

region is roughly 3300 lbs. per acre, with an annual production of 123 million pounds (lbs.) and estimated value of $ 37 million in income to growers. However, the area under peanut has declined in recent years. Higher yield, drought tolerance, and reduced diseases would improve the competitiveness of peanut vs. corn production, especially as corn requires significantly more water than peanut. This is of increasing importance for the sustainability of agriculture in the region with the depletion of underground water supplies, and the elevated price of corn as a result of the demand for ethanol. This project proposes to address two important aspects:

a. Development of drought and disease resistant Valencia peanuts
b. Study the socio economic impact assessment through collaborators in the host countries and in partnership with Auburn University, USA, ICRISAT, India and Khon Kaen University, Thailand

## Vision and Approach

### Goals

Our long term goal is to identify diverse germplasm with beneficial traits from Arachis hypogaea L. subsp. fastigiata var. fastigiata (Valencia) and markers associated with agronomic and seed quality traits, resistance to pod rot, and tolerance to drought that breeders can use in enhancing the genetic potential of Valencia peanuts.

The specific goals of this project are: 1. To evaluate the Valencia peanut core collection for agronomic/seed quality traits, pod rot resistance, and drought tolerance, 2. To study the genetic structure of the Valencia peanut core collection using SSRs in a high-throughput assay, 3. To study variation in transcript abundance among drought tolerant and susceptible lines using oligonucleotide microarrays, 4. To make an initial attempt at identifying allelic variation (markers) associated with beneficial traits for use in breeding programs.

### Objectives

a. Evaluate accessions for agronomic/seed quality traits, resistance to pod rot and tolerance to drought
b. Study the genetic structure of accessions using SSRs and high throughput assay
c. Study variation in transcript abundance among drought tolerant and susceptible lines using oligonucleotide microarrays
d. Identify allelic variation (markers) associated with beneficial traits in Valencia peanut germplasm
e. Screening of Valencia core collection for Rossette and Leaf Spot
f. Evaluating Valencia variety Red Beauty with fungicide Abound at two locations and its economic impact on yield and disease.
g. To support and train two graduate students from Makerere University in Plant

h. Breeding and Agronomy.

## Research Approach

i) Screening for end-of-season drought tolerance: The Valencia peanut crop grown in New Mexico matures in about 120 days that this crop duration can be partitioned into three distinct phases of crop development: early crop establishment (up to 40 days after sowing; DAS); flowering, pegging and early pod initiation and development (41 to 80 days from DAS); and pod/seed development and maturity (81 to 120 days from DAS). It is the end-of-season drought which coincides with the pod/seed development and maturity that causes substantial yield reduction in peanut.

It is proposed to evaluate Valencia peanut core collection for physiological traits associated with drought tolerance under end-of-season drought conditions at Agricultural Experiment Station, University of Arizona, Yuma, Arizona. Yuma is an ideal location for drought screening work as it is located in a desert with an annual rainfall of 75 mm. Most of the rainfall occurs mainly during late autumn to early spring with an average temperature of 29 to 33o C during July to September (Holbrook et al. 1994). It is proposed to withhold irrigation from 80 DAS until the crop is harvested to impose end-of-season drought treatment; however, control plots will receive the full irrigation for comparisons with drought stressed plots. Low cost movable rainout shelter will be developed to avoid rains when plots exposed to end-of-season drought treatment (Chauhan et al. 1997). Observations on physiological traits (in addition to plot yield) such as specific leaf area (SLA), specific leaf nitrogen (SLN), SPAD chlorophyll meter reading (SCMR), transpiration efficiency (TE), and harvest index (HI) will be recorded both from irrigated control and drought stressed plots. SCMR provides a rapid assessment of SLA and SLN as surrogate trait to measure TE in peanut (Nageswara Rao et al. 2001; Sheshshayee et al. 2006).

(ii) Screening for resistance to pod rot: A sick plot for pod rot complex diseases has been developed at South Research Facility (SRF), Clovis, New Mexico. Initially, all the accessions will be evaluated for resistance to pod rot complex diseases under field conditions, and only the promising accessions will be screened for pod rot resistance (Godoy et al. 1984) in micro plots at SRF. In field and micro plot evaluations, the test entries will be scored on 0 to 10 scale, where 0 is no decay and 10 is total pod decay (Pattee et al. 1974). Inoculum of each pathogen will be produced using cornmeal-sand medium (Garren 1970).

(iii) Evaluating core collection for agronomic and seed quality traits: The core collection will also be evaluated in replicated trials for agronomic and seed quality traits at Brownfield (West Texas). Observations on maturity (based on thermal degree days), pod yield, shelling percentage, percent sound mature kernels, 100-seed weight, and seed quality traits (oil and protein contents, fatty acid composition, and total sugars) will be recorded. Seed quality testing will be done at J. Leek Associates, commercial lab located in Brownfield, Texas

or at the Texas Agricultural Experiment Station (TAES) Lubbock Center. (iv) Genotyping core collection (using SSRs and high throughput assay) to dissect genetic structure and marker-trait association: DNA will be extracted using the Qiagen DNeasy Plant mini kit (Qiagen Inc., Valencia CA) following the manufacturer's protocol. Molecular profiling will be done using Beckman CEQ8000 Analyzer (Fullerton, CA) at the TAES or an ABI 3100 (Applied Biosystems, Foster City, CA) at the USDA-ARS Cropping Systems lab, and amplification products will be scanned using software provided by the manufacturer (CEQ STR [Short Tandem Repeats]) module or ABI Genescan 3.1.2 and scored as binary data. We have devised a scoring protocol to identify true bands from spurious peaks caused by stuttering of Taq polymerage (Kottapalli et al. 2007).

Statistical analysis will be done following DARwin 5.0 version (Perrier et al. 2003) for dissecting the genetic structure, PowerMarker 3.0 for determining gene diversity, rare alleles, common alleles and group specific alleles (Liu and Muse 2005; http://www.powermarker.net). Genetic distances will be calculated using Chord's distance formula (Cavalli-Sforza and Edwards 1967) as it generates correct tree topologies regardless of the microsatellite mutation model (Takezaki and Nei 1996) and the matrix thus generated will be used to construct dendrogram based on Neighbor Joining (NJ) clustering procedure implemented in NTSYSpc2.2 (Rohlf 1998; Exeter Software, Setauket, NY). Bootstrapping will be carried out by performing 1,000 bootstrap resamplings using PowerMarker and Winboot (Yap and Nelson 1996). All analysis will be performed using the MXCOMP module of the NTSYSpc software 2.2.

The extent of linkage disequilibrium will be estimated using TASSEL (http://www.maizegenetics.net/index.php?page=bioinformatics/index.html). Population structure will be determined using STRUCTURE (Pritchard et al., 2000) (http://pritch.bsd.uchicago.edu/structure.html), relative kinship (Loiselle et al., 1995) using SPAGeDi (http://www.ulb.ac.be/sciences/ecoevol/spagedi.html). Association mapping will be performed using the mixed Q+K model (Yu and Buckler, 2006) method in TASSEL (http://www.maizegenetics.net/index.php?page=bioinformatics/index.html) or SAS (SAS Institute, Cary, NC).

(v) Transcriptional profiling and statistical analysis of microarray data: The laboratory of collaborator Dr. Payton is currently developing oligonucleotide microarrays for expression profiling in peanut. Baseline expression profiling studies on one line identified as stress tolerant and one line identified as stress-sensitive will directly compare: a) control (full irrigation) to drought-stressed leaf and root tissue and b) stressed leaf and root tissue from tolerant and sensitive lines (outlined in Figure 1) (Yang and Speed 2002). We will use a modified dye-swap experiment for each comparison (Kerr et al. 2002). Briefly, for each pair-wise comparison (control vs. stress or tolerant vs. sensitive) 6 RNA extracts will be obtained from independent, pooled tissue samples. The RNA extracts from each tissue pool will be used as template to

prepare six corresponding cDNA populations for each stress state (three cy3 and three cy5). Six hybridizations will be carried out for each specific comparison for a total of 36 hybridizations (18 for leaf and 18 for root).

### Training & Capacity Development Approach

We plan to develop training workshops inviting farmers, extension agents, NGOS and other technicians for demonstration purpose of preliminary yield data and show the performance of varieties in comparison to the local checks.

## Intended Benefits & Impact Responsiveness

### Development Benefits

Host country will be benefited from increased yields in peanut production for growers through drought and disease resistant varieties. Input cost will be reduced by not applying fungicides. Smallholder farmers will be benefitted by growing improved cultivars. Consumers will benefit from both nutritionally and financially by having a high quality disease and drought resistant sweet tasting Valencia peanuts.

### US Benefits

Drought and disease has a significant impact on peanut production in the USA especially eastern New Mexico and west Texas. We hope to have a Valencia peanut that has both drought and disease resistant which uses less water and produces more peanuts for unit of water applied. As a result Valencia peanuts that are sold in-shell has bright pod color and less discoloration resulting in good quality and quantity peanuts and be able to compete with world market. This will eventually result in economic stability and reduced import of peanuts from Argentina and other countries at cheaper rate that are extremely sensitive to fluctuation in world market.

### Potential Impacts

The Valencia peanuts for in-shell market are predominantly grown in eastern New Mexico, and west Texas. Of recent, water has become a scarce resource in this region. The area under peanut cultivation has therefore declined thus affecting the monopoly of this region to produce peanuts for in-shell market types. Development and cultivation of high yielding Valencia peanut cultivars that mature early, uses less water to produce more (high water use efficiency), tolerant to drought, and resistant to pod rot and aflatoxin will benefit peanut producers in this part of the US and the host countries. The overall goal of this proposal is to develop a core collection specific to Valencia peanut that will be evaluated for various morpho-agronomic traits, physiological traits related to drought, and resistance to pod rot and aflatoxin contamination to identify promising germplasm for enhancing the genetic potential of Valencia peanuts. Association mapping on this core collection will enable us to identify markers associated with beneficial traits, as discussed above, for use in breeding programs. Over 500 publicly available SSRs will be

initially screened to identify polymorphic markers on a set of diverse germplasm and then the full core collection will be genotyped using 100 polymorphic SSRs and high throughput assay (ABI3700) to dissect the genetic structure of this core collection. The New Mexico Peanut Research Board, National Peanut Research Board provides limited funds to support ongoing peanut breeding program at NMSU; however, additional support from USAID through University of Georgia through this proposed project will help us identify critical germplasm and the marker technology for effective use in development of high-yielding Valencia peanut cultivars in US and in the host countries.

### Equipment

It is early to predict what equipment a host country may require at this point. But depending on the mission interest and the host country that is selected we may be able to identify the equipment need while writing a full proposal.

### Project Timeline

1. To evaluate the Valencia peanut core collection for agronomic/seed quality traits, pod-rot resistance, and drought tolerance. This objective is a long-term project involving multi-location and year study to evaluate the germplasm along with control and will take full five years of the project.
2. To study the genetic structure of the Valencia peanut core collection using SSRs in a high-throughput assay. This will be started in year 1 and anticipated to be completed by year 2.
3. To study variation in transcript abundance among drought tolerant and susceptible lines using oligonucleotide microarrays. This objective will be started in year 3 and will be completed in year 5.
4. To make an initial attempt at identifying allelic variation (markers) associated with beneficial traits for use in breeding programs. This objective will be started in year 3 and will be completed in year 5.
5. Economic analysis of peanut production and its efficiency in host countries. This objective is a long-term project and will take at least five years to complete the project. We need to collect the data from the growers, generate crop budget report for peanuts and compare it to crop budgets of other major crops within the host country. Report on peanut production system and constraints and cite how to improve production under existing conditions. List products available in the market and market opportunities.

## USAID Mandate Responsiveness

### MDGs

Poverty/Hunger: Improved Health: Raised Rural Incomes: Sustainable Development

### Foreign Assistance Framework

Governance: Human Capacity: Economic Structure: Persistent Dire Poverty:

Global Issues (HIV and Infectious Diseases, climate change, biodiversity)

IEHA

Science and Tech Applications: Increased demand for peanuts: Market Access: Increased Trade

USAID Focal Areas

Greater incomes: Greater value and market demand: Public Health: Food Security: Sustainable Value Chain: Improved Human Capacity