Utilization of SNP, SSR, and Biochemical Data to Evaluate Genetic and Phenotypic Diversity in the U.S. Peanut Germplasm Collection

Barkley NA, Wang ML, Pittman RN
USDA-ARS Plant Genetic Resources Conservation Unit, Griffin, GA 30223

Abstract

Peanuts (*Arachis hypogaea* L.) are nutritious because their seeds typically contain high amounts of oil, protein, phytochemicals such as resveratrol, and antioxidants such as tocopherol and folic acid; therefore, they are an important oil seed crop worldwide. The genetic diversity and population structure of the U.S. cultivated mini-core collection was assessed by genotyping 94 accessions with 81 SSR markers and two functional SNP markers derived from the fatty acid desaturase 2 (*ahFAD2*) gene. The SNP markers were developed to detect wild type and mutant alleles in both *ahFAD2A* and *ahFAD2B*, which are known to affect oleic acid (C18:1) and linoleic acid (C18:2) accumulation. Further, biochemical data such as total oil content, fatty acid composition, flavonoids, resveratrol, and morphological traits were also collected from the mini core accessions. The molecular markers and phenotypic trait data were employed to assess allelic variation, phylogenetic relationships, population structure, and association analysis. Population structure analysis identified four major subpopulations that were related to four botanical varieties. Genetic variation was revealed in the mini core and molecular markers which associated with phenotypic traits were identified. The information obtained will be useful for enhancing breeding efficiency and improving seed quality in peanuts.

Introduction

Germplasm repositories preserve plant materials (seeds, buds, etc) so that agricultural and horticulture plants will not be lost due to development and population growth. These genetic resources can be maintained over decades or centuries with limited regeneration dependent on the species being conserved. Because peanuts contain a large portion of oil in their seed (typically contain ~50% oil in the seed), their shelf life is time-limited. It is critically important to preserve and regenerate plant germplasm so that genetic improvements can be made to thwart abiotic and biotic stresses that occur in the environment. These improvements can occur through traditional breeding or mutagenesis to create new combinations of alleles to improve yield, disease resistance, or drought. Readily available germplasm allows scientists to evaluate traits to select appropriate parents in breeding programs to improve crop plants.

The USDA-ARS Plant Genetic Resources Conservation Unit in Griffin, GA houses a large collection of cultivated and wild *Arachis* accessions. This collection consists of 9,917 *Arachis* accessions including both cultivated and wild germplasm. Wild and cultivated peanuts have significant variation in pod shape, pod reticulation, seed color and seed pattern (Figure 1), as well as, other important agronomic traits such as disease resistance and yield. In order to enhance the utilization of peanut germplasm, a core (831 accessions) and mini core (112 accessions) collection were established to represent maximum genetic diversity with minimal redundancy to help researchers mine important traits from a manageable sample set (Holbrook et al, 1993; Holbrook and Dong, 2005). These collections have been utilized to identify important traits such as leaf spot and tomato spotted wilt resistance (Holbrook and Dong, 2005). The purpose of this study was to evaluate the genetic diversity and population structure of the cultivated mini core collection and associate phenotypic traits to specific molecular markers. This was accomplished by employing SSR and SNP markers along with measuring biochemical and morphological traits in the mini core collection.

Materials and Methods

All DNA samples were extracted by following the instructions from an Omega Bio-Tek E.Z.N.A. Plant DNA kit (Norcross, GA.) Seed slices (~100 mg) were placed in a 2 mL centrifuge tube along with two 3 mm tungsten carbide beads and 600 µl of P1 buffer from the extraction kit. Tissue was pulverized in a Retsch Mixer Mill 301 (Leeds, UK) at 30 Hz for three minutes. All DNA samples were loaded into a 1% agarose gel along with a Low DNA Mass™ Ladder from Invitrogen (Carlsbad, CA) to evaluate quality and quantity of each extraction. Extracts were quantified with a Nanodrop 2000C (Termo Fisher Scientific, Waltham MA) and subsequently diluted to 10 ng/µl for PCR. A previously developed SNP assays to detect G448A in *ahFAD2A* and 442insA in *ahFAD2B* were employed on all samples (Barkley et al. 2010; Barkley et al. 2011). Mutant alleles were detected with probes labeled with 6-Fam; whereas, wild type alleles were detected with Vic labeled probes. PCR reactions were performed in an ABI StepOne machine using 48-well optical plates and seals. SSR markers were chosen from previously published studies. PCR products were separated on an ABI 377 and Mega gel system. Total oil content was measured on a Maran Puls nuclear magnetic resonance machine. Fatty acids were separated on a gas chromatograph (Agilent 7890A), and flavanoids were quantified by using high performance liquid chromatography (Agilent 1100). Further detail on all of the methods can be obtained from Wang et al., 2011.

Results and Discussion

The genetic diversity and population structure of the mini core collection was assessed by employing SSR and SNP markers. The SSR markers revealed a total of 664 alleles with an average of 8.1 alleles/locus. The mean polymorphism content (PIC) and genetic diversity was 0.53 and 0.59, respectively. Functional SNP markers were employed to detect G448A in *ahFAD2A* and 442insA in *ahFAD2B* which are the key SNPs that control the accumulation of oleic and linoleic acid (Figure 2). Three genotypes for *ahFAD2A* were revealed with a total of 52 wild type (G/G), 39 mutants (A/A) and 3 heterozygotes (G/A). Only the wild type allele was detected in *ahFAD2B* in the mini core. The SNP and SSR data was utilized to evaluate population structure and construct a phylogeny. The structure analysis revealed 4 subpopulations with relatively low levels of admixture which may be due to the nature of peanut being a self pollinated species. The subpopulations correlated well with four botanical varieties (*fastigiata*, *hypogaea*, *peruviana*, & *vulgaris*) of cultivated peanut (Figure 3). Biochemical and morphological data such as total oil, fatty acid composition, flavanoids, seed and pod characteristics were also collected from the mini core. The markers, biochemical, and morphological data were tested for marker trait associations. The SNP marker *ahFAD2A* was significantly associated with oleic acid (P value = 1.75×10^{-5}) and linoleic acid (P value = 4.58×10^{-5}). Further, a SSR marker (Ah041) was also significantly associated with linoleic acid. None of the other markers associated with the morphological or biochemical data that were collected. Future work will include evaluating more markers and traits to find significant associations. Molecular markers linked to important agronomic traits can be used in marker assisted selection for improving cultivated peanuts.

Literature Cited

Barkley NA, Chamberlin KD, Wang ML, Pittman RN. 2010. Development of a real-time PCR genotyping assay to identify high oleic acid peanuts (*Arachis hypogaea* L.) Mol Breed DOI 10.1007/s11032-009-9338-2.
Barkley NA, Wang ML, Pittman RN. 2011. A real-time PCR genotyping assay to detect FAD2A SNPs in peanuts (Arachis hypogaea L.). Elec J Biotech DOI: 10.2225/vol14-issue1-fulltext-12.
Holbrook CC, Anderson WF, Pittman RN (1993). Selection of a core collection from the US germplasm collection of peanut. Crop Sci 33:859-861.
Holbrook CC, Dong WB (2005). Development and evaluation of a mini core collection for the US peanut germplasm collection. Crop Sci 45:1541-1544.
Wang ML, Sukumaran S, Barkley NA, Chen Z, Chen CY, Guo B., Pittman RN, Stalker HT, Holbrook CC, Pederson GA, Yu J. Population structure and marker trait association analysis of the US peanut mini core collection. Theor Appl Genet DOI: 10.1007/s00122-011-1668-7.

Figure 1: A selection of the diversity in morphological traits of seeds and pods found in the U.S. peanut germplasm collection.

Figure 2: SNP genotyping utilizing a real-time PCR to detect *ahFAD2* alleles (top panel). The bottom panel shows the percentage of oleic acid detected for each accession along with the *ahFAD2* genotype.

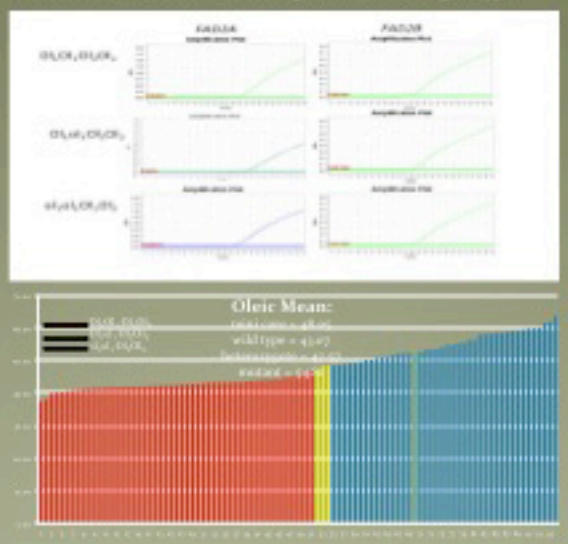

Figure 3: a) STRUCTURE analysis of peanut mini core collection b) Association analysis of marker data and phenotypic traits.