Molecular characterization of groundnut (Arachis hypogaea L.) composite collection

Project PI: HD Upadhyaya, ICRISAT
Project Team: ICRISAT – R Bhattacharjee, DA Hoisington, S Chandra, RK Varshney and S Singh
EMBRAPA – JFM Valls, MC Moretzsohn, S Leal-Bertioli and Patricia Guimaraes
UCB – David Bertoili

**Groundnut (Arachis hypogaea L.)**

- Cultivated groundnut is a tetraploid (2n = 40) and highly self-pollinated crop.
- Primary center of origin is the Chaco region between southern Bolivia and northwestern Argentina.
- Important oilseed crop grown in 109 countries on 25.2 million ha area (FAOSTAT, 2005).
- Over two-thirds of global production occurs in seasonally rainfall regions.

**Groundnut composite collection**

- The groundnut composite collection was developed considering the phenotypic diversity present in the entire collection held at ICRISAT genebank and EMBRAPA, Brazil (Upadhyaya et al., 2005).
- The composite collection consists of accessions from ICRISAT comprising of mini-core (Upadhyaya et al., 2002), comparator mini-core and trait-based accessions along with 52 accessions of 14 wild Arachis species, and accessions from EMBRAPA, Brazil.

**Genetic diversity in composite collection**

**Plant material**

At ICRISAT, DNA extracted from 916 accessions following a high-throughput procedure and quantified to a working concentration of 5 ng/µl.

**Selection of SSR markers**

- At ICRISAT, 20 SSR markers, available in public domain (Ferguson et al., 2004), initially selected and pre-screened on 184 mini core accessions from which 10 polymorphic markers identified.
- Eleven SSR markers from EMBRAPA (Moretzsohn et al. 2005) also included to fingerprint the composite collection.

**Molecular characterization**

- The PCR components of all 21 SSR markers optimized following Taguchi method (Taguchi, 1986) described in Cobb and Clarkson (1994).
- Fluorescent-based multiplex genotyping system used to generate five multiplexes of four markers each.
- Capillary electrophoresis with an automated system (ABI 3700) to separate the amplified PCR products.
- SSR fragment sizes called to two decimal places using Genotyper v3.7 software.
- Phylogenetic tree, constructed at 0.76 to 0.89 bootstrap confidence level (Table 2).
- Further analysis of data in progress to fully understand the genetic diversity and population structure of the composite collection.
- Future plan

**Results**

- Statistical analysis detected a total of 491 alleles, ranging from 5 (7H6) to 46 (SDS) with a mean of 23.4 alleles per locus (Table 1).
- The mean Polymorphic Information Content (PIC) value was 0.796 (ranging from 0.483 to 0.923) (Table 1).
- The regional distribution (Table 2) showed that number of alleles per locus ranged from 3.3 in accessions from Oceania to 22.1 in accessions from South America.

**Future plan**

- Further analysis of data in progress to fully understand the genetic diversity and population structure of the composite collection.
- Results from genotypic data will be used to identify a reference set of 300 diverse accessions for future use.
- To ascertain the quality and position of SSR markers, these will be checked on 15-20 plants in each of four F_r populations, whose parents have been included in the composite collection.

**References**


**Data analysis**

- DAReWm 5.0 Structure program (Perrier et al., 2003) used to determine population structure of the composite collection.
- Sixty accessions with high missing values excluded from data analysis.
- Principal coordinate analysis done considering taxonomical classification of Arachis, i.e. at the level of two subspecies (hypogaea and fastigiate) and six botanical varieties (hypogaea, fastigiate, vulganes, peruviana, aquetioniana and hirsuta), and the wild species.
- Cervus 2.0 software used to calculate allele frequencies and PIC values.

**Plant material**

All ICRISAT, DNA extracted from 916 accessions following a high-throughput procedure and quantified to a working concentration of 5 ng/µl.

**Selection of SSR markers**

- At ICRISAT, 20 SSR markers, available in public domain (Ferguson et al., 2004), initially selected and pre-screened on 184 mini core accessions from which 10 polymorphic markers identified.
- Eleven SSR markers from EMBRAPA (Moretzsohn et al. 2005) also included to fingerprint the composite collection.

**Molecular characterization**

- The PCR components of all 21 SSR markers optimized following Taguchi method (Taguchi, 1986) described in Cobb and Clarkson (1994).
- Fluorescent-based multiplex genotyping system used to generate five multiplexes of four markers each.
- Capillary electrophoresis with an automated system (ABI 3700) to separate the amplified PCR products.
- SSR fragment sizes called to two decimal places using Genotyper v3.7 software.
- Phylogenetic tree, constructed at substespecies level, revealed that both hypogaea and fastigiate accessions formed distinct clusters (Fig. 2). Further, some fastigiate accessions grouped with hypogaea types mainly due to common geographical origin.
- The accessions belonging to 14 wild Arachis species grouped together in a separate cluster, but close to hypogaea accessions (Fig. 2).

**Future plan**

- Further analysis of data in progress to fully understand the genetic diversity and population structure of the composite collection.
- Results from genotypic data will be used to identify a reference set of 300 diverse accessions for future use.
- To ascertain the quality and position of SSR markers, these will be checked on 15-20 plants in each of four F_r populations, whose parents have been included in the composite collection.

**References**

Taguchi G. 1998. Introduction to Quality Engineering, Asian Productivity Organization, American Supplier Institute, Dearborn, MI.