Published online July 30, 2007

RESEARCH

Molecular Characterization of the U.S. Peanut Mini Core Collection Using Microsatellite Markers

Kameswara Rao Kottapalli, Mark D. Burow, Gloria Burow, John Burke, and Naveen Puppala*

ABSTRACT

Peanut (*Arachis hypogaea* L.) is the second-most important legume crop in the United States. A limitation to increased peanut productivity is that peanut improvement is hampered by relatively low genetic variability in the germplasm commonly used by breeding programs. To facilitate accessibility to diverse germplasm sources for breeding applications, a core subset of the USDA peanut germplasm was previously established that later was further refined and developed into a mini core collection consisting of 112 accessions. This report details an extensive characterization of genetic diversity and relationships in the U.S. peanut mini core using microsatellite or simple sequence repeat (SSR) markers. Seventy-two peanut accessions from the U.S. peanut mini core were genotyped with 73 SSR markers; all but six produced reliable, polymorphic bands. Moderate levels of genetic variation were found with genetic distances (D) values among accessions ranging from 0.088 to 0.254. Distinct groupings of the accessions based on subspecies classification and on botanical (market) type groupings were established. Twelve of the markers, mapped previously to the A genome, were found to be sufficient to identify both subspecies and botanical types and gave a clustering pattern very similar to the entire 67 SSR marker set. The genetic variation observed within U.S. peanut mini core can be utilized for selection of diverse parents for breeding and development of mapping populations.

K.R. Kottapalli and N. Puppala, Agric. Sci. Cent. at Clovis, New Mexico State Univ., 2346 SR 288 Clovis, NM 88101; M.D. Burow, Texas Agric. Exp. Stn., Texas A&M Univ., Lubbock, TX 79403, and Texas Tech Univ., Dep. of Plant and Soil Science, Lubbock, TX 79409; G. Burow and J. Burke, USDA-ARS, Plant Stress and Germplasm Development Unit, Cropping Systems Research Lab, 3810 4th Street, Lubbock, TX 79415. This research was supported by a USDA grant (2004-34186-14533) through the Southwest Consortium for Plant Genetics and Water Resources and New Mexico State University Agric. Exp. Stn. Mention of trademark or proprietary product does not constitute a guarantee or warranty of a product by the USDA and does not imply its approval to the exclusion of other products that may also be suitable. Received 21 June 2006. *Corresponding author (npuppala@nmsu.edu).

Abbreviations: AFLP, amplified fragment length polymorphism; SSR, simple sequence repeat.

Genetic variability is key to progress in crop improvement programs. However, the cultivated tetraploid peanut (*Arachis hypogaea* L.) has limited genetic variability, which has been explained as the result of a genetic bottleneck resulting from a single domestication event approximately 3500 yr ago (Simpson et al., 2002). Recent studies by Milla et al. (2005) using amplified fragment length polymorphism (AFLP) verified previous findings by Kochert et al. (1996) that limited molecular polymorphism was found in the cultigen. In addition to paucity of genetic variation in the species, commercial peanut breeding programs further reduced genetic variation due to the predominance of few common elite parents, for example 'Dixie Giant' and 'Small White Spanish-1' (Knauft and Gorbet, 1989), in the ancestry of many varieties. The consequence is a reduced level of exploitable

Published in Crop Sci. 47:1718–1727 (2007).
doi: 10.2135/cropsci2006.06.0407

677 S. Segoe Rd., Madison, WI 53711 USA

genetic variation and, therefore, a limited response to selection. For example, the peanut crop is vulnerable to numerous pests and diseases due to its narrow genetic base. To sustain progress in peanut breeding and minimize genetic vulnerability, it is essential to broaden the genetic base of the elite U.S. germplasm used by breeders. However, commercial breeders are hesitant to use wild or exotic relatives directly to increase the genetic diversity of elite germplasm because it takes considerable resources to develop adapted, high- performing materials from crosses between elite and wild or exotic genotypes. While it is ideal that breeders can tap the available germplasm in collections such as the Agricultural Research Services (ARS) Germplasm Resource Information Network (GRIN) database (http://www.ars-grin.gov), detailed passport data on the features and variation generally available in the large collection are limited. For peanut it may be the ultimate objective to characterize the available collection but this is quite laborious and time consuming and it will take some time for these types of data to become available. A pragmatic solution to this problem is the use of a characterized subset of cultivated germplasm accessions, referred to as a core collection, for transfer of agronomically important and complex traits. A core collection is a subset of accessions from the entire collection that capture most of the available genetic diversity of the species (Brown, 1989).

Core collections have been developed for many crops, including peanut. Holbrook et al. (1993) developed a core collection of 831 accessions from a set of 7432 accessions in the U.S. peanut germplasm collection based on six morphological variables. Subsequently, a core consisting of 1704 accessions (14 morphological descriptors on 14 310 accessions) was developed using the ICRISAT peanut collection (Upadhyaya et al., 2003), and a core of 582 accessions was established by random selection from the almost 6000 accessions of the Chinese peanut germplasm collection (Holbrook et al., 2004). However, when the size of collection is too large and a core collection (10% of entire collection) becomes unmanageable, Upadhyaya and Ortiz (2001) suggested a strategy of selecting a mini core (core of core) collection (10% of the core or 1% of the entire collection). Using this strategy, mini cores of 184 accessions from the ICRISAT core collection (Upadhyaya et al., 2002) and of 112 accessions from the U.S. core collection (Holbrook and Dong, 2005) were developed. The peanut core and mini core collections have been evaluated for various traits, and new sources of variation are reported for tolerance to drought (Upadhyaya, 2005); early maturity (Upadhyaya et al., 2006); high yield and other agronomic traits (Upadhyaya et al., 2005); resistances to peanut root-knot nematode [*Meloidogyne arenaria* (Neal)] (Holbrook et al., 2000), *Tomato spotted wilt virus* (Anderson et al., 1996), cylindrocladium black rot [caused by *Cylindrocladium crotalariae* (Loos) Bell and Sobers] and early leafspot (caused by *Cercospora arachidicola* Hori) (Isleib et al., 1995), rhizoctonia limb rot (caused by *Rhizoctonia solani* Kühn) (Franke et al., 1999), sclerotinia blight [caused by *Sclerotinia minor* Jagger and *S. sclerotiorum* (Lib.) de Bary], pepper spot [caused by *Leptosphaerulina crassiasca* (Sechet) Jackson and Bell] (Damicone et al., 2003); and reduced preharvest aflatoxin contamination (Holbrook et al., 1998). The use of these diverse sources can help in bringing in much needed diversity to broaden the genetic base of cultivars.

Germplasm descriptions may be based on morphological descriptors, agronomic descriptors, and/or molecular markers. Morphological descriptors have been used widely for classification of species. Use of discrete, character state data has the advantage of being less subject to environmental variation than continuously variable traits. For classification within a species, some discrete character state descriptors are available, but classification of large numbers of accessions by these has some practical difficulties. In addition, most of the traits useful for improvement are quantitative in nature. However, these are influenced strongly by environment, and repeated screening often leads to conflicting results. Additionally, the cost of large scale experimentation to generate such data is prohibitive. Recently, germplasm characterization based on molecular markers has gained importance due to the speed and quality of data generated. The smaller core or mini core germplasm accessions were characterized initially using DNA markers such as random amplified polymorphic DNA (RAPDs) in common bean (*Phaseolus vulgaris* L.) (Skroch et al., 1998) and potato (*Solanum tuberosum* L.) (Ghislain et al., 1999). AFLP markers have been used for studying the variation in core subsets of oats (Fu et al., 2005) and in combination with SSRs and isozymes in cassava (*Manihot esculenta* Crantz) (Chavarriaga-Aguirre et al., 1999). Microsatellite (SSR) markers were utilized to reveal genetic identities, diversity, and relationships in apple (*Malus* spp.) core collections (Hokanson et al., 1998). SSR and RFLP markers were used to characterize a diverse subset of rice cultivars for use in the development of the U.S. rice (*Oryza sativa* L.) core collection (Xu et al., 2004). Single nucleotide polymorphism (SNP) markers associated with cooking qualities of rice were used for characterization of a core subset of rice germplasm collections maintained at the USDA-ARS National Small Grains Collection (McClung et al., 2004).

Among various DNA markers, microsatellites are well known for their potentially high information content and versatility as molecular tools. They are also amenable to high throughput genotyping and have proven to be highly versatile and useful markers for germplasm characterization. In peanuts, SSRs detected genetic diversity among 48 valencia genotypes (Krishna et al., 2004) and also differentiated botanical types (He et al., 2005). Two other

studies (Ferguson et al., 2004a; Moretzsohn et al., 2004) used SSRs to determine the genetic diversity of cultivated peanut and its wild relatives. However, the mini core collections developed recently in peanuts have been characterized only using morphological descriptors and few agronomic traits (Upadhyaya, 2005). The specific goals of the present study are: (i) to characterize the U.S. peanut mini core collection utilizing simple sequence repeat (SSR) DNA markers, (ii) to assess genetic affinity of peanut botanical and market types based on mapped markers, and (iii) to determine whether the molecular diversity within the collection is correlated to geographical origin of the accessions.

MATERIALS AND METHODS

Plant Materials

The core subset of the U.S. peanut mini core collection consists of 112 accessions (Holbrook and Dong, 2005), and the 99 available accessions were graciously provided by C. Holbrook, USDA, ARS-Coastal Plain Experiment Station, Tifton, GA. These were planted at the Texas Tech University experimental farm in Lubbock, TX, and in the greenhouse of the New Mexico State University Agricultural Sciences Center at Clovis, NM. Of these 99, 27 accessions appeared to segregate for multiple morphological characteristics in the field and were excluded from the analysis, leaving 72 accessions representing the two subspecies and four botanical varieties of peanut grown in the United States. The segregating accessions were excluded because of uncertainty whether the segregation represented contamination of seed lots or natural variation within these accessions.

Information about the accessions used (accession identity as designated by PI number, core collection number, botanical type, proposed market type, and country of origin) is listed in Table 1. Information on market types are based on plant, pod, and seed characteristics observed in the field and after harvest. These took into account growth habit, pod and seed size and shape, and number of seeds per pod.

Genomic DNA Extraction and SSR Genotyping

DNA was extracted using the Qiagen DNeasy Plant minikit (Qiagen Inc., Valencia, CA) following the manufacturer's protocol, except that the quantity of starting material was reduced to 75 to 100 mg and the final purified DNA was eluted in 120 μL of elution buffer. The 73 SSR primers employed to analyze the accessions are given in Table 2. A modified M13-tagged forward and normal reverse primer was used for each marker. Briefly, a 20mer M13 oligo (GAC GTT GTA AAA CGA CGG CC) was concatenated to the 5′ end of each forward primer, generating an M13-tagged primer. To facilitate detection, a 20mer M13 oligo labeled with one of three fluorescent dyes, 6-HEX, FAM, or NED, was added to the polymerase chain reaction (PCR) mix to label SSR products of each marker. PCR was performed in a volume of 5 or 10 μL, containing 1× PCR buffer (10 m*M* Tris-HCl pH 8.3, 50 m*M* KCl, 1.5 m*M* $MgCl_2$), 0.2 m*M* of each dNTP, 0.25 units Hotstart *Taq* DNA polymerase (Qiagen Inc.), 5 pmol of each M13-tagged forward and normal reverse primer, 0.02 pmol of fluorescently labeled 20mer M13 primer (labeled with either HEX, NED, or FAM), and 20 ng of template genomic DNA. Amplifications were performed using a PTC-225 (MJ Research, Waltham, MA) Peltier thermal cycler with the following conditions for the PG series of primers: 94°C initial denaturation for 17 min (1 cycle); 35 cycles of 94°C for 45 s, 60°C for 60 s, 72°C for 30 s; and a final extension at 72°C for 10 min. The PCR conditions used for primer sets were as follows: for the AH series primers, according to Hopkins et al. (1999); for the PGP and PGS series primers, Ferguson et al. (2004a, 2004b); for PM series primers, He et al. (2003); and for the RKN series primers, Burow et al. (1996).

The 73 primer sets were tested first using DNA of four accessions for optimizing the PCR conditions and concentrations of the reaction mix, and products were visualized on 1.5% agarose gels. All the primers amplified PCR products. These primers were used for genotyping the 72 available, homogeneous accessions of the U.S. peanut mini core. After initial fragment analysis, 18 primers were used to reamplify the missing genotypes and test the repeatability of patterns. During the process of reamplification, one sample with only sterile water was used as template (negative control) in the PCR and was used to detect and eliminate artifacts. Repeatability of banding pattern was totally consistent for all primers tested, whereby detection of major band sizes was observed at 100% (data not presented).

Markers were amplified separately using the set of 72 accessions and were size separated by capillary electrophoresis using an ABI Prism 3100 DNA Analyzer (Applied Biosystems, Foster City, CA) at the USDA-ARS Plant Stress and Germplasm Development Unit (Lubbock, TX). Multiplexing of three differentially labeled PCR products per well was performed to increase efficiency. The multiplexing protocol enabled the study of the entire set of accessions with 73 primers in 3 to 4 wk.

SSRs were analyzed with Genescan 3.1.2 software (Applied Biosystems) and scored using Gene Mapper V3.0 software (Applied Biosystems) as binary data.

Data Analysis

Pairwise comparisons of genetic distances were calculated using Nei's formula (Nei, 1972). Genetic distances were also 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). A genetic distance matrix was established and subsequently used to construct a dendrogram based on the Neighbor Joining (NJ) clustering procedure implemented in NTSYSpc2.2 (Rohlf, 1998; Exeter Software, Setauket, NY) and PowerMarker3.0 (Liu and Muse 2005; http://www.powermarker.net). Bootstrapping was performed by performing 1000 bootstrap resamplings using PowerMarker and Winboot (Yap and Nelson, 1996). A goodness of fit test was performed for the cophenetic matrix of the NJ tree to the genetic distance matrix by the Mantel test (Mantel, 1967). Similarly, a test was made on the goodness of fit of the cophenetic matrix of the NJ tree from the 12 mapped SSRs to the genetic distance matrix based on the entire set of 67 SSRs. All procedures were performed using the MXCOMP module of the NTSYSpc software 2.2. PowerMarker was used to calculate the average number of bands per SSR primer pair,

Table 1. List of cultivated accessions comprising the US peanut mini core collection used in the study.

Core collection no.	PI no.	Country of origin	Subspecies	Market type†
008	295730	India	*hypogaea*	Virginia
012	493329	Argentina	*fastigiata*	Valencia
016	493356	Argentina	*hypogaea*	Virginia
033	493547	Argentina	*fastigiata*	Valencia
038	493581	Argentina	*fastigiata*	Valencia
041	493631	Argentina	*fastigiata*	Valencia
047	493693	Argentina	*hypogaea*	Virginia
050	493717	Argentina	*fastigiata*	Valencia
068	493880	Argentina	*fastigiata*	Valencia
082	494034	Argentina	*fastigiata*	Spanish
087	475863	Bolivia	*fastigiata*	Valencia
097	497395	Bolivia	*hypogaea*	Virginia
112	497517	Brazil	*fastigiata*	Valencia
115	496401	Burkina Faso	*hypogaea*	Virginia
119	496448	Burkina Faso	*hypogaea*	Virginia
132	497639	Ecuador	*fastigiata*	Valencia
149	502040	Peru	*fastigiata*	Spanish
166	494795	Zambia	*hypogaea*	Runner
189	339960	Argentina	*fastigiata*	Valencia
208	274193	Bolivia	*hypogaea*	Virginia
223	290620	Argentina	*hypogaea*	Virginia
227	290566	India	*hypogaea*	Runner
230	290594	India	*hypogaea*	Runner
233	290536	India	*hypogaea*	Virginia
249	343384	Israel	*fastigiata*	Intermediate
266	200441	Japan	*fastigiata*	Spanish
270	196635	Madagascar	*hypogaea*	Runner
277	259851	Malawi	*hypogaea*	Virginia
294	372271	Nigeria	*hypogaea*	Virginia
296	399581	Nigeria	*hypogaea*	Virginia
310	337406	Paraguay	*hypogaea*	Runner
334	159786	Senegal	*hypogaea*	Virginia
338	268696	South Africa	*fastigiata*	Spanish
367	268868	Sudan	*hypogaea*	Virginia
384	166107	Uruguay	*fastigiata*	Valencia
388	162655	Uruguay	*fastigiata*	Spanish
406	152146	Uruguay	*fastigiata*	Spanish
408	262038	Brazil	*fastigiata*	Valencia
431	337293	Brazil	*fastigiata*	Valencia
477	268806	Zambia	*fastigiata*	Spanish
506	259658	Cuba	*hypogaea*	Runner
516	288146	India	*hypogaea*	Virginia
529	319768	Israel	*hypogaea*	Virginia
540	295250	Israel	*hypogaea*	Virginia
542	370331	Israel	*hypogaea*	Virginia
546	259836	Malawi	*fastigiata*	Spanish
548	325943	Venezuela	*fastigiata*	Valencia
559	158854	China	*fastigiata*	Valencia
580	268586	Zambia	*hypogaea*	Valencia
610	475931	Bolivia	*hypogaea*	Virginia
631	408743	Brazil	*fastigiata*	Intermediate
643	433347	China	*fastigiata*	Spanish
650	478819	India	*fastigiata*	Valencia
673	481795	Mozambique	*fastigiata*	Spanish
678	476636	Nigeria	*hypogaea*	Virginia
698	372305	Nigeria	*hypogaea*	Virginia
703	476432	Nigeria	*fastigiata*	Intermediate
728	292950	South Africa	*hypogaea*	Runner
731	162857	Sudan	*hypogaea*	Virginia
740	407667	Thailand	*fastigiata*	Spanish
755	482189	Zimbabwe	*fastigiata*	Spanish
760	471952	Zimbabwe	*fastigiata*	Spanish
763	442768	Zimbabwe	*hypogaea*	Virginia
775	402120	Zimbabwe	*fastigiata*	Spanish
781	471954	Zimbabwe	*fastigiata*	Valencia
787	429420	Zimbabwe	*fastigiata*	Valencia
802	196622	Côte d'Ivoire	*hypogaea*	Virginia
805	355268	Mexico	*hypogaea*	Virginia
808	337399	Morocco	*fastigiata*	Spanish
812	323268	Pakistan	*hypogaea*	Virginia
Valencia C‡	565461	USA	*fastigiata*	Valencia
S-3870‡	Registered	Bulgaria	*fastigiata*	Valencia

†Data on growth habit, botanical and market type were obtained from the field study

‡Valencia-C and S-3870 are not part of mini core collection but used to represent diverse valencia genotypes

gene diversity, and polymorphism information content (PIC) values. PIC is defined as: PIC = $1 - \Sigma p_i^2$, where p_i is the frequency of the ith allele.

RESULTS

Microsatellite Polymorphism

The 72 available, phenotypically uniform accessions of the core subset of the U.S. peanut mini core collection were amplified using 73 SSR primer pairs. Two primer pairs, which did not amplify most of the accessions, were eliminated from the analysis. Of the 71 remaining primer pairs, four produced monomorphic patterns and were not considered further. The remaining 67 primers amplified 528 polymorphic bands with an average of 7.88 bands per primer pair (Table 3). The number of bands scored ranged from 2 to 28 per primer pair. Notably, some of the primers generated genotype-specific bands, which can be used as molecular identity data for specific genotypes. The PIC for all the genotyped bands was estimated using PowerMarker and ranged from 0.063 to 0.918. The gene diversity for all the 528 bands ranged from 0.027 to 0.50, which is relatively high due to the presence of genotype-specific bands.

In the subspecies *fastigiata*, 462 polymorphic bands (87.5% polymorphism) were found with an average number of 6.9 bands per primer pair (Table 3). The average gene diversity and PIC values were similar to the entire mini core collection values. There were 66 *fastigiata*-specific bands accounting for 12.5% of the total amplified bands and 14.3% of the polymorphic bands. The number of group-specific bands per genotype was only 1.69. Among the *hypogaea* subspecies, 390 polymorphic bands were observed (73.9% polymorphism).

Table 2. List of primers used in the study and corresponding number of bands resolved, polymorphism information content (PIC), and source of primer (reference).

Reference	Primer	No. of bands observed	PIC Value
Hopkins et al. 1999	AH 41	5	0.529
	AH 193	11	0.800
	AH 229	12	0.808
	AH 522	2	0.388
Ferguson et al. 2004a	PGP 01B09	5	0.712
	PGP 02A06	4	0.520
	PGP 03B06	3	0.456
	PGP 03B10	3	0.485
	PGP 03C02	6	0.678
	PGP 03D09	3	0.259
	PGP 04D04	5	0.585
	PGP 04G02	12	0.803
	PGP 06B08	28	0.888
	PGP 07B03	18	0.846
	PGP 07B09	4	0.295
	PGP 07H06	22	0.836
	PGP 07H09	3	0.413
Ferguson et al. 2004b	PGS 09F01	13	0.608
	PGS 09G05	16	0.879
	PGS 09G12B	4	0.457
	PGS 10C12	8	0.751
	PGS 10H01	10	0.843
	PGS 11E11	12	0.712
	PGS 11H01	13	0.866
	PGS 12A07	6	0.727
	PGS 12B06	3	0.384
	PGS 12E10	4	0.520
	PGS 12F07	4	0.608
	PGS 13A07	14	0.695
	PGS 13C08	12	0.642
	PGS 13E06A	7	0.606
	PGS 13E11	5	0.618
	PGS 14A10	2	0.081
	PGS 14C11	8	0.834
	PGS 14D01	3	0.233
	PGS 14E10	6	0.439
	PGS 14F05	9	0.786
	PGS 15B01	3	0.489
	PGS 15E11	10	0.872
	PGS 15F12	7	0.774
	PGS 16C06	2	0.209
	PGS 17H05	7	0.636
	PGS 18A05	4	0.411
	PGS 18C05	17	0.720
	PGS 19B12	3	0.231
	PGS 19C03	8	0.677
	PGS 19E09	4	0.695
	PGS 19G05	2	0.284
He et al. 2003	PM 3	20	0.918
	PM 15	10	0.790
	PM 32	3	0.387
	PM 35	10	0.857
	PM 36	13	0.831
	PM 42	5	0.632
	PM 45	4	0.482
	PM 50	7	0.709
	PM 53	2	0.140
	PM 65	2	0.063
	PM 137	7	0.710
	PM 145	7	0.680
	PM 183	8	0.677
	PM 188	12	0.757
	PM 200	15	0.819
	PM 201	8	0.742
	PM 204	14	0.606
	PM 210	9	0.661
	PM 238	5	0.763

The average number of bands per primer pair, average gene diversity, and PIC values (5.8, 0.16 and 0.13 respectively) were lower than the values of the entire mini core collection. However, the number of group-specific bands (138) and their percentage of the total amplified bands (26%) was twice the number observed for *fastigiata* group.

Analysis of Genetic Relationship in the U.S. Peanut Mini core

The NJ tree based on Nei's genetic distances is shown in Fig. 1. The genotypes were grouped into 2 major clusters, which corresponded to the subspecies groupings; subspecies *fastigiata* and *hypogaea* (Fig. 1). The group of accessions from two clusters *hypogaea* and *fastigiata* were separated by a genetic distance of 0.26. But the genotypes within the *fastigiata* group (largest group) were separated by an average genetic distance of 0.23, indicating low to moderate genetic diversity in the entire collection. However, the estimates of genetic distances observed with tree have clearly distinguished 94% of the mini core collection accessions into the major botanical types. Groupings of the accessions based on market types were also observed and the grouping was in agreement with morphological descriptors with a mixture of ~5% in both subspecies (Table 1 and Fig. 1). Both virginia and runner market types grouped together while valencia and spanish market types showed 88 and 60% (15 out of 17 and 12 out of 20) true types, respectively. The correlation between genetic distance and NJ tree cophenetic matrix was highly significant with an *r* value of 0.78, which indicated that the cluster generated by NJ was a good fit to the matrix.

Clustering and Grouping Using Mapped SSR Markers

Twelve SSR primer pairs that were mapped previously onto seven linkage groups (Moretzsohn et al., 2005) were used for clustering (Table 4). Another mapped marker,

Ah-75, was also analyzed in the present study but it resulted in monomorphic bands. The mapped SSR loci amplified from 2 to 20 bands per primer pair, with an average of 9.16 bands per primer pair. The gene diversity ranged from 0.03 to 0.50. The data from 12 SSR primer sets (mapped markers AH-193, AH-229, PM-3, PM-32, PM-35, PM-36, PM-42, PM-45, PM-65, PM-188, PM-204, and PM-238) were used to classify the four market types prevalent in the United States. The NJ tree based on Chord's 1967 genetic distance unambiguously differentiated the market types into four major groups with the valencia and spanish market types (var. *fastigiata*) as out groups (Fig. 2). Virginia and runner market types (var. *hypogaea*) grouped together as expected. The cophenetic matrix for the NJ tree based on 12 markers was significantly correlated with the genetic distance matrix derived from the 67 primer sets. The *r* value for this comparison was 0.72, which indicates that the 12 mapped markers that were distributed among seven different linkage group are very promising candidates for studies of genetic diversity.

Phylogeographic Pattern of Genetic Diversity

Geographic diversity of the accessions of the mini core collection was examined within marker-derived groupings of subspecies *hypogaea* and *fastigiata*. Two major clades were observed in *hypogaea* subspecies with larger number of accessions from South America and Africa in clade I and II respectively (Fig. 1). Three major clades were formed in *fastigiata* group that contained genotypes from multiple countries of origin (Table 1 and Fig. 1). The Valencia clade of *fastigiata* subspecies contained high percentage of accessions from South America and Spanish clade I mostly contained five accessions from South America and four from Africa. Spanish clade II had large number of accessions from Africa. Seven accessions representing valencia market type were from Argentina.

DISCUSSION

Differentiation of Genotypes

This study presents a genotypic characterization of a peanut mini core subset germplasm collection. In contrast to reports generated using RFLP and AFLP markers (Kochert et al., 1996; Milla et al., 2005), SSR markers were able to identify genetic variability among accessions of the cultivated species *A. hypogaea*, with a moderate level of molecular variation being observed in this study. In contrast to other studies using SSR markers, where as wide a genetic diversity as possible was used to identify polymorphisms (e.g., Ferguson et al., 2004a), the current study used a data set deemed to be more representative of the variability within the cultigen. The molecular data generated in the present study can be utilized by peanut breeders for selection of parents and genotyping of mapping populations.

Table 3. Summary statistics for all the accessions of *Arachis hypogaea* and each subspecies.

Summary statistics	All	var. *fastigiata*	var. *hypogaea*
Sample size	72	39	33
Total polymorphic bands	528	462	390
Average no. of bands per primer set	7.88	6.90	5.82
Group-specific bands		66	138
Group-specific bands per accession		1.69	4.18
Group-specific alleles, %		12.5	26.1
Average gene diversity	0.18	0.18	0.16
Average PIC value†	0.15	0.15	0.13

†PIC, polymorphism information content.

Molecular patterns generated for the mini core can also be compared with the existing database as well as with the patterns of newly reported germplasm to avoid genotype redundancy (Khadari et al., 2003). This in turn will enable efficient management of peanut germplasm.

Several genotype-specific bands were generated with more than 20 bands identified for some primer pairs. The average number of bands per primer pair was as expected for peanut SSRs (Krishna et al., 2004; He et al., 2005; Ferguson et al., 2004a). Between the two subspecies, *fastigiata* had more polymorphic bands but half the number of group-specific bands. The reason for such a deviation may be attributed to the origin of the accessions within the subspecies.

Differentiation of Subspecies

The core collection was mainly compiled using the two major subspecies, *fastigiata* and *hypogaea*, including only the four market types cultivated in the United States. The clustering based on Chord's distance and primer sets corresponding to 12 mapped markers could consistently

Table 4. Linkage group location, size range and number of alleles as well as gene diversity detected for 72 mini core collection accessions based on 12 mapped simple sequence repeat markers (Moretzsohn et al., 2005)

Primer locus	Linkage group location	Size range of alleles (bp)	No. of alleles	Gene diversity (range)
Ah-193	LG1	122–472	11	0.03–0.45
PM32	LG 2	118–124	3	0.08–0.39
PM238	LG 3	163–179	5	0.22–0.49
PM3	LG 3	171–252	20	0.04–0.45
PM42	LG 3	182–222	5	0.03–0.49
Ah-229	LG 4	193–293	12	0.03–0.45
PM35	LG 5	142–162	10	0.11–0.48
PM36	LG 5	119–246	12	0.07–0.47
PM45	LG 5	116–122	4	0.03–0.46
PM65	LG 5	117–245	2	0.06–0.21
PM204	LG 7	128–264	12	0.03–0.22
PM188	LG 8	109–133	12	0.29–0.50

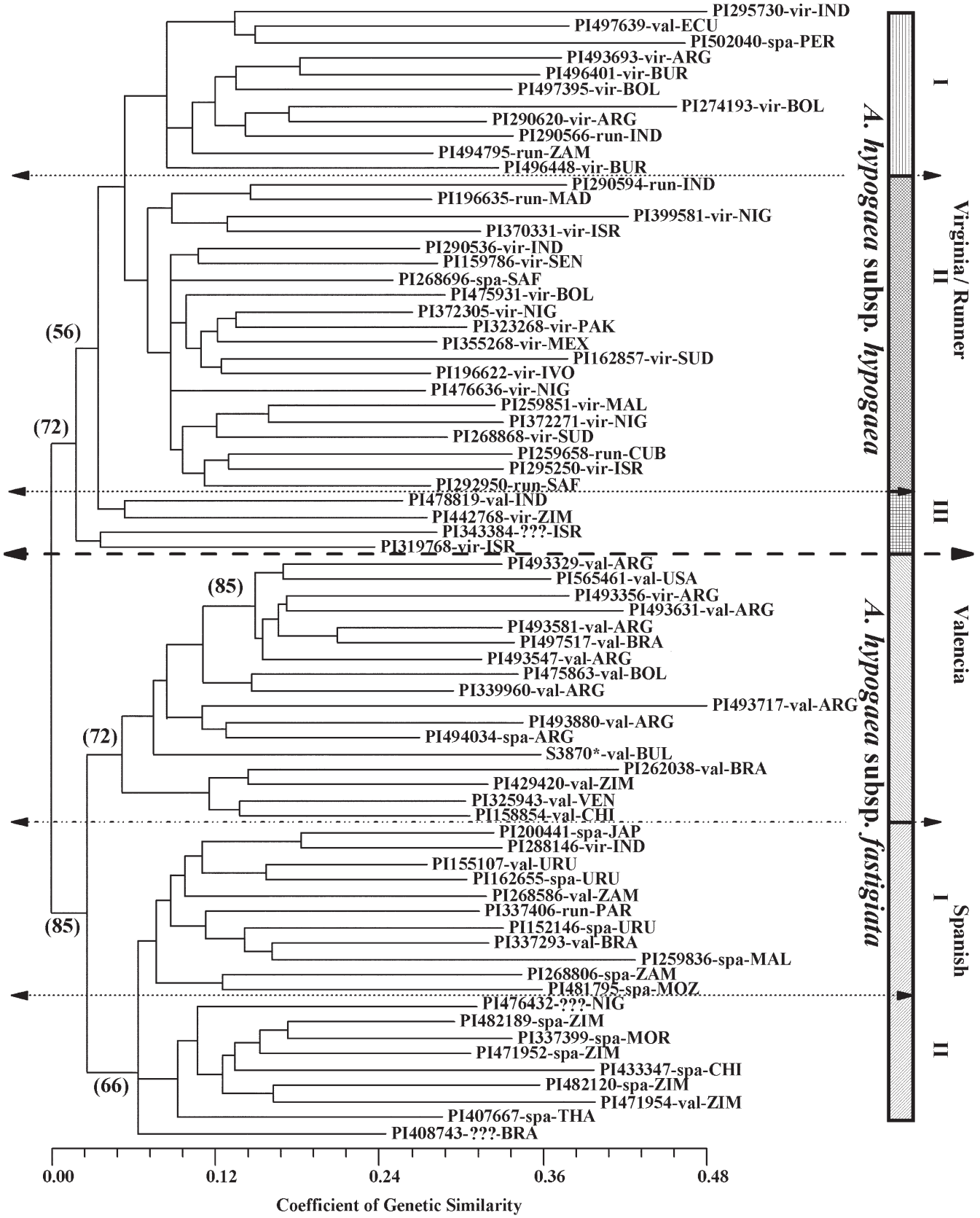

Figure 1. Neighbor-joining dendrogram of 72 *A. hypogaea* accessions from the U.S. peanut mini core revealed by using similarity coefficients based on Nei's genetic distance. The letters after each accession number refer to market types as: val, subsp. *fastigiata*/*fastigiata*/valencia; spa, subsp. *fastigiata*/vulgaris/spanish; vir, subsp. *hypogaea*/*hypogaea*/virginia; run, subsp *hypogaea*/*hypogaea*/runner; ???, intermediate market type. Country of origin is abbreviated as the first three letters of the country after the subspecies (for e.g., IND = India). The dashed line indicates groupings of accessions into the two major subspecies and dotted lines indicate grouping under each subspecies into market types. The numbers inside the parentheses indicate bootstrap values for groupings with >50% bootstrap support based on 1000 bootstrap samples.

differentiate the two major botanical types. The primer PM-188 located on linkage group 8 was also used by He et al. (2005) and could distinguish *fastigiata* and *hypogaea*.

The distinction between subspecies accords with the major morphological differences distinguishing them. Among these are erect vs. spreading growth habit, flowers present vs. absent on the mainstem, and alternate vs. sequential branching. Other important differences that are typical are early vs. late maturity, small light-pigmented vs. large dark-pigmented leaves, and absence or presence of seed dormancy. Previous studies using AFLP markers failed to group the subspecies of cultivated peanut (Herselman, 2003). This may be due to the nature of AFLP markers used in the study. The map position of these markers is unknown, but it is possible that they clustered into a small portion of the genome (as is common with AFLP markers), thus limiting the genomic region that was scanned. With the recent development of SSRs and knowing their positions on the genetic map, it is now possible to scan the genome for variability. This underscores the need for further development of genetic maps of cultivated peanut using SSR markers or PCR-based markers, which would enhance their application in marker-assisted breeding.

Differentiation of Market Classes

We report using mapped SSR markers to distinguish cultivated U.S. market types of peanut. Ferguson et al. (2004b) was able to differentiate six botanical types of peanut (including var. *peruviana*, *hirsuta*, and *aequatoriana* that are not included in the core subset of the U.S. core collection) using 10 SSR primer pairs. He et al. (2005) also reported a separate set of six SSR primer sets providing markers specific to botanical varieties and different from those used in the present study. As suggested, these markers may be linked to genes involved in the expression of morphological traits. In our study, all 12 primer pairs that correspond to mapped loci on seven linkage groups could differentiate the major market types. Both virginia and runner market types were grouped in the same cluster, this could be due to low molecular variation within the *hypogaea* subspecies. This work dealt with distinguishing runner and virginia types based on mapped markers. However, with the larger number of 67 SSR primer sets, only Valencia and Spanish market types were grouped separately with few mixtures in each group. The instance of clustering of a valencia or spanish market type in the *hypogaea* subspecies and virginia or runner in *fastigiata* subspecies of the tree based on 67 primer sets could be due to the inefficiency of some primers either to amplify or fail to amplify a similar DNA fragment in both the genotypes. A similar result was obtained by Garris et al. (2005) while grouping rice genotypes into two major *indica* and *japonica* subspecies.

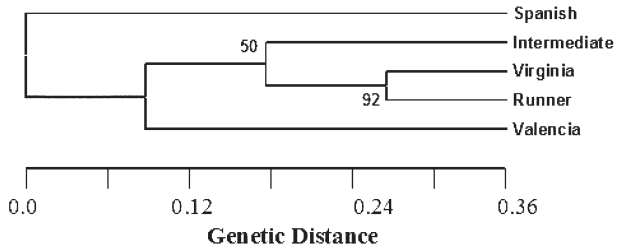

Figure 2. Dendrogram of four market types of the peanut accessions from the U.S. mini core based on 12 mapped simple sequence repeat (SSR) markers used in the study. The numbers inside the parentheses indicate bootstrap values for groupings with >50% bootstrap support based on 1000 bootstrap samples.

Geographic Distribution of Accessions

From the phylogeographic point of view, the mini core collection had genotypes from 19 countries in the *fastigiata* group and 18 countries in the *hypogaea* group (Table 1). Although there is overrepresentation from Argentina (South America) in the *fastigiata* group, this is not unusual because valencia peanut is abundant in the region and Argentina is a possible center of diversity of the market type (Ferguson et al., 2004b). Among hypogaea's, runner market types were underrepresented in the mini core collection. However, generally there is a consistency of selection of genotypes from diverse origins. Ferguson et al. (2004b) also noted that there was greater differentiation among subspecies than among continents, that is, the differences among subspecies is greater than the geographic differences.

In this study it was shown that moderate levels of genetic variation could be detected effectively in peanut using SSR markers. The grouping at the accession level indicated a clear distinction between subspecies and among the accepted major botanical types grown in the US. This is an extensive molecular study of the U.S. peanut mini core and has provided useful information toward parental selections and specific SSR marker that can be used for varietal identification.

Acknowledgments

The authors wish to thank Dr. Corley Holbrook for providing germplasm lines for this research and Halee Hughes, Jacob Sanchez, and Jamie Ayers for their technical assistance. We also thank Dr. S.L. Dwivedi for critical review of an earlier version of this manuscript.

References

Anderson, W.F., C.C. Holbrook, and A.K. Culbreath. 1996. Screening the core collection for resistance to tomato spotted wilt virus. Peanut Sci. 23:57–61.

Brown, A.H.D. 1989. Core collections: A practical approach to genetic resource management. Genome 31:818–824.

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

Cavalli-Sforza, L., and A.W. Edwards. 1967. Phylogenetic analysis: Models and estimation procedures. Am. J. Hum. Genet. 19:233–257.

Chavarriaga-Aguirre, P., M.M. Maya, J. Tohme, M.C. Duque, C. Iglesias, M.W. Bonierbale, S. Kresovich, and G. Kochert. 1999. Using microsatellites, isozymes and AFLPs to evaluate genetic diversity and redundancy in the cassava collection and to assess the usefulness of DNA-based markers to maintain germplasm collections. Mol. Breed. 5:263–273.

Damicone, J.P., K.E. Jackson, K.E. Dashiell, H.A. Melouk, and C.C. Holbrook. 2003. Reaction of the peanut core to sclerotinia blight and pepperspot [Online]. Available at www.apres.okstate.edu/old%20proceedings/APRES%202003%20Proceedings%20Vol%2035.pdf (verified 7 June 2007). Am. Peanut Res. and Educ. Soc., Stillwater, OK.

Ferguson, M.E., P.J. Bramel, and S. Chandra. 2004b. Gene diversity among botanical varieties on peanut (*Arachis hypogaea* L.). Crop Sci. 44:1847–1854.

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

Franke, M.D., T.B. Brenneman, and C.C. Holbrook. 1999. Identification of resistance to Rhizoctonia limb rot in a core collection of peanut germplasm. Plant Dis. 83:944–948.

Fu, Y.B., G.W. Peterson, D. Williams, K.W. Richards, and J.M. Fetch. 2005. Patterns of AFLP variation in a core subset of cultivated hexaploid oat germplasm. Theor. Appl. Genet. 530:530–539.

Garris, A.J., T.H. Tai, J. Coburn, S. Kresovich, and S. McCouch. 2005. Genetic structure and diversity in *Oryza sativa* L. Genetics 169:1631–1638.

Ghislain, M., D. Zhang, D. Fazardo, Z. Huaman, and R.J. Hijmans. 1999. Marker-assisted sampling of the cultivated Andean potato *Solanum phureja* collection using RAPD markers. Genet. Resour. Crop Evol. 46:547–555.

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

He, G., R. Meng, M. Newman, G. Gao, R. Pittman, and C.S. Prakash. 2003. Microsatellites as DNA markers in cultivated peanut (*A. hypogaea* L.) BMC Plant Biol. 3:3. Available at http://www.Biomedcentral.com/1471-2229/3/3.

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

Hokanson, S.C., A.K. Szewc-McFadden, W.F. Lamboy, and J.R. McFerson. 1998. Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a *Malus x domestica* Borkh. core subset collection. Theor. Appl. Genet. 97:671–683.

Holbrook, C.C., W.F. Anderson, and R.N. Pitmann. 1993. Selection of core collection from US germplasm collection of peanut. Crop Sci. 33:859–861.

Holbrook, C.C., and W. Dong. 2005. Development and evaluation of a mini core collection for the U.S. peanut germplasm collection. Crop Sci. 45:1540–1544.

Holbrook, C.C., H. Jiang, B. Liao, X. Liang, and B. Guo. 2004. Selection and evaluation of a core collection for peanut germplasm in China. Abstract 5603 *In* 2004 Agronomy abstracts [CD-ROM]. ASA, Madison, WI.

Holbrook, C.C., M.G. Stephenson, and A.W. Johnson. 2000. Level and geographical distribution of resistance to *Meloidogyne arenaria* in the U.S. peanut germplasm collection. Crop Sci. 40:1168–1171.

Holbrook, C.C., D.M. Wilson, and M.E. Matheron. 1998. Sources of resistance to pre-harvest aflatoxin contamination in peanut. Proc. Amer. Peanut Res. Educ. Soc. 30:54 (Abstr.).

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

Isleib, T.G., M.K. Beute, P.W. Rice, and J.E. Hollowell. 1995. Screening the peanut core collection for resistance to Cylindrocladium black rot and early leaf spot. Proc. Amer. Peanut Res. Educ. Soc. 27:25 (abstr).

Khadari, B., C. Breton, N. Moutier, J.P. Roger, G. Besnard, A. Bervillé, and F. Dosba. 2003. The use of molecular markers for germplasm management in a French olive collection. Theor. Appl. Genet. 106:521–529.

Knauft, D.A., and D.W. Gorbet. 1989. Genetic diversity among peanut cultivars. Crop Sci. 29:1417–1422.

Kochert, G., H.T. Stalker, M. Ginenes, L. Galgaro, and K. Moore. 1996. RFLP and cytogenetic evidence for the progenitor species of allotetraploid cultivated peanut (*Arachis hypogaea* L.). Am. J. Bot. 83:1282–1291.

Krishna, G.K., J. Zhang, M. Burow, R.N. Pittman, S.G. Delikostadinov, Y. Lu, and N. Puppala. 2004. Genetic diversity analysis in Valencia peanut (*Arachis hypogaea* L.) using microsatellites. Cell. Mol. Bio. Lett. 9:685–697.

Liu, K., and S.V. Muse. 2005. Powermarker: Integrated analysis environment for genetic marker data. Bioinformatics 21:2128–2129.

Mantel, N.A. 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27:209–220.

McClung, A.M., M. Chen, H.E. Bockelman, R.J. Bryant, W. Yan, and R.G. Fjellstrom. 2004. Characterization of a core collection of rice germplasm and elite breeding lines in the US with genetic markers associated with cooking quality. Poster 127. Proc. of the 2nd International Rice Functional Genomics Conf., Tucson, AZ. 15–17 Nov. 2004.

Milla, S.R., T.G. Isleib, and H.T. Stalker. 2005. Taxonomic relationships among *Arachis* sect. *Arachis* species as revealed by AFLP markers. Genome 48:1–11.

Moretzsohn, M.C., M.S. Hopkins, S.E. Mitchell, S. Kresovich, J.F.M. Valls, and M.E. Ferreira. 2004. Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis of hypervaraible regions of the genome. BMC Plant Biol. 4:11 http://www.biomedcentral.com/1471-2229/4/11.

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

Nei, M. 1972. Genetic distance between populations. Am. Nat. 106:283–292.

Rohlf, F.J. 1998. NTSYSpc. Numerical taxonomy and multivariate analysis system. Applied Biostatistics, Inc., New York.

Simpson, C.E., A. Krapovickas, and J.M. Valls. 2002. History of *Arachis* including evidence of *A. hypogaea* progenitors. Peanut Sci. 28:79–81.

Skroch, P.W., J. Nienhuis, S. Beebe, J. Tohme, and F. Pedraza. 1998. Comparison of Mexican common bean (*Phaseolus vulgaris* L.) core and reserve germplasm collections. Crop Sci. 38:488–496.

Takezaki, N., and M. Nei. 1996. Genetic distances and reconstruction

of phylogenetic trees from microsatellite DNA. Genetics 144:389–399.

Upadhyaya, H.D. 2005. Variability for drought resistance related traits in the mini core collection of peanut. Crop Sci. 45:1432–1440.

Upadhyaya, H.D., P.J. Bramel, R. Ortiz, and S. Singh. 2002. Developing a mini core of peanut for utilization of genetic resources. Crop Sci. 42:2150–2156.

Upadhyaya, H.D., B.P. Mallikarjuna Swamy, P.V. Kenchana Goudar, B.Y. Kullaiswamy, and S. Singh. 2005. Identification of diverse groundnut germplasm through multienvironment evaluation of a core collection for Asia. Field Crops Res. 93:293–299.

Upadhyaya, H.D., and R. Ortiz. 2001. A mini core subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. Theor. Appl. Genet. 102:1292–1298.

Upadhyaya, H.D., R. Ortiz, P.J. Bramel, and S. Singh. 2003. Development of a groundnut core collection using taxonomical, geographical and morphological descriptors. Genet. Resour. Crop Evol. 50:139–148.

Upadhyaya, H.D., L.J. Reddy, C.L.L. Gowda, and S. Singh. 2006. Identification of diverse groundnut germplasm sources of early maturity in a core collection. 2006. Field Crops Res. 97:261–267.

Xu, Y., H. Beachell, and S.R. McCouch. 2004. A marker-based approach to broadening the genetic base of rice in the USA. Crop Sci. 44:1947–1959.

Yap, I.V., and R.J. Nelson. 1996. Winboot: A program for performing bootstrap analysis of binary data [Online]. Available at www.irri.org/science/software/winboot.asp. IRRI, Manila, Philippines.