Biotechnology Advances 30 (2012) 639–651

Contents lists available at SciVerse ScienceDirect

Biotechnology Advances

journal homepage: www.elsevier.com/locate/biotechadv

Research review paper

Advances in *Arachis* genomics for peanut improvement

Manish K. Pandey [a], Emmanuel Monyo [b], Peggy Ozias-Akins [c], Xuanquiang Liang [d], Patricia Guimarães [e], Shyam N. Nigam [a], Hari D. Upadhyaya [a], Pasupuleti Janila [a], Xinyou Zhang [f], Baozhu Guo [g], Douglas R. Cook [h], David J. Bertioli [i], Richard Michelmore [h], Rajeev K. Varshney [a,j,k,*]

[a] *International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502324, India*
[b] *International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Lilongwe 1096, Malawi*
[c] *The University of Georgia, Athens, GA 30602, USA*
[d] *Crops Research Institute, Guandong Academy of Agricultural Sciences, Guangzhou 510640, China*
[e] *National Center for Genetic Resources and Biotechnology, EMBRAPA, Brasília DF 70770900, Brazil*
[f] *Henan Academy of Agricultural Sciences, Zhengzhou 450002, China*
[g] *USDA-ARS, Crop Protection and Management Unit, Tifton, GA 31793, USA*
[h] *University of California, Davis, CA 95616, USA*
[i] *University of Brasília, Brasília DF 70910900, Brazil*
[j] *CGIAR Generation Challenge Programme (GCP), c/o CIMMYT, DF 06600, Mexico*
[k] *The University of Western Australia, Crawley, WA 6009, Australia*

ARTICLE INFO

Article history:
Received 16 August 2011
Received in revised form 24 October 2011
Accepted 1 November 2011
Available online 9 November 2011

Keywords:
Groundnut
Molecular markers
Molecular breeding
Genomic resources
Genetic improvement

ABSTRACT

Peanut genomics is very challenging due to its inherent problem of genetic architecture. Blockage of gene flow from diploid wild relatives to the tetraploid; cultivated peanut, recent polyploidization combined with self pollination, and the narrow genetic base of the primary genepool have resulted in low genetic diversity that has remained a major bottleneck for genetic improvement of peanut. Harnessing the rich source of wild relatives has been negligible due to differences in ploidy level as well as genetic drag and undesirable alleles for low yield. Lack of appropriate genomic resources has severely hampered molecular breeding activities, and this crop remains among the less-studied crops. The last five years, however, have witnessed accelerated development of genomic resources such as development of molecular markers, genetic and physical maps, generation of expressed sequenced tags (ESTs), development of mutant resources, and functional genomics platforms that facilitate the identification of QTLs and discovery of genes associated with tolerance/resistance to abiotic and biotic stresses and agronomic traits. Molecular breeding has been initiated for several traits for development of superior genotypes. The genome or at least gene space sequence is expected to be available in near future and this will further accelerate use of biotechnological approaches for peanut improvement.

Contents

* Corresponding author at: International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502324, India. Tel.: +91 40 30713305; fax: +91 40 30713074.
E-mail address: r.k.varshney@cgiar.org (R.K. Varshney).

0734-9750/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.biotechadv.2011.11.001

1. Introduction

Peanut or groundnut (*Arachis hypogaea* L.) is the fourth-largest oilseed crop in the world and is cultivated in more than 100 countries, with the annual production of 35.5 million tonnes (FAO, 2009). The largest producers of peanut are China and India, followed by the USA. Peanut seeds contain 40–60% oil, 20–40% protein and 10–20% carbohydrate. Peanut has high nutritional value, possessing vitamin E, niacin, calcium, phosphorus, magnesium, zinc, iron, riboflavin, thiamine and potassium. Peanut is mainly used for direct consumption, in the confectionary industry, for vegetable oil in cooking and also as a source for protein feed in the animal industry. These multiple uses of peanut make it an excellent cash crop for domestic as well as international trade.

Peanut belongs to the genus *Arachis* which is divided into 9 intrageneric taxonomic sections based on morphology, geographic distribution and cross compatibility (Krapovickas and Gregory, 1994). Section *Arachis* is the largest section and most diverse and harbors the only peanut species that is widely cultivated for its seeds and pods (*A. hypogaea*). There are also forage species (*A. pintoi*, *A. glabrata* and *A. sylvestris*) (Valls and Simpson, 1994) and ornamental species (*A. repens*) (Stalker and Simpson, 1995). All the sections of *Arachis* species possess the diploid genome except two tetraploids i.e., cultivated peanut (*A. hypogaea*) and one wild species (*A. monticola*) which are cross compatible with each other (Krapovickas and Gregory, 1994). The species in other sections are mostly diploid and have very limited sexual compatibility with cultivated peanut.

The cultivated species, *A. hypogaea,* with a large and tetraploid genome, is very probably derived from a unique cross between the wild diploid species *A. duranensis* (A-genome) and *A. ipaënsis* (B-genome) resulting in a hybrid followed by spontaneous chromosome duplication (Kochert et al., 1996; Seijo et al., 2004). The origin through a single and recent polyploidization event, followed by successive selection resulted in a highly conserved genome (Young et al., 1996). Although cultivated peanut is a tetraploid, genetically it behaves as diploid (Stalker et al., 1991). It has been concluded that the A- and B-genomes contributed nearly equal amounts of DNA to the domesticated peanut (Singh et al., 1996). Apart from A-genome and B-genome species, section *Arachis* has a lone D-genome species *A. glandulifera* (Stalker, 1991). Cultivated peanut is divided into two subspecies, *hypogaea* and *fastigiata* based on growth habit and the presence or absence of flowers on the main axis. Subspecies *hypogaea* is again divided into two botanical varieties (var. *hypogaea* and var. *hirsuta*) and subspecies *fastigiata* into four botanical varieties (var. *fastigiata*, var. *peruviana*, var. *aequatoriana* and var. *vulgaris*) based on inflorescence, pod and seed characteristics.

Cultivated peanut is mainly grown in the semi-arid tropics (SAT) region by resource-poor farmers. As a result, crop productivity has been adversely challenged by several abiotic and biotic stresses. In addition, aflatoxin contamination deteriorates product quality and greatly reduces grain value. The major biotic stress factors include early leaf spot (*Cercospora arachidicola*), late leaf spot (*Phaeoisariopsis personata*), rust (*Puccinia arachidis*), mottle virus (*Peanut mottle virus*), rosette virus (*Groundnut rosette virus*), aphids (*Aphis craccivora*), jassids (*Amrasca devastans*) and thrips (*Frankliniella* spp.). Drought is the major abiotic stress as 70% of the crop is under semi-arid tropics, which is characterized by low and erratic rainfall. Soil moisture during pod filling stages affects the aflatoxin accumulation in seeds. Although several efforts through conventional breeding have been made to enhance crop productivity, the results have not been encouraging. However, the increasing demand for peanut will require higher productivity and conventional breeding is inadequate to address the complex genetic behavior of the majority of desired traits. In terms of genetic improvement, the narrow genetic base of the cultivated peanut genepool, and the tetraploid and complex nature of genome are some of the serious bottlenecks. Only limited genetic diversification has been achieved in the past through interspecific hybridization between cultivated peanut and other species of section *Arachis* due to differences in ploidy levels and the linkage drag. Eliminating the linkage drag involves a lengthy process that also results in dilution of the level of resistance present in wild relatives of *Arachis*.

As demonstrated in several temperate cereal crops (Varshney et al., 2006) and some legume crops (Varshney et al., 2010a, 2010b), molecular breeding has significant advantages over conventional breeding in handling traits which are difficult to manage through conventional phenotypic selection. Apart from handling trait complexity, molecular breeding also facilitates introgression/pyramiding of multiple recessive alleles very efficiently in less time and with more accuracy along with pyramiding of several monogenic traits or several QTLs for a single trait (Ribaut and Hoisington, 1998; Varshney et al., 2005a, 2009a, 2009b; Xu and Crouch, 2008). However, to apply the molecular breeding approaches, availability of genetic variation in germplasm, critical mass of molecular markers, genetic maps, and appropriate phenotyping platforms are required. In the case of cultivated peanut, however, restricted gene flow has hampered the much needed broadening of the genetic base of the species. In addition to low levels of genetic variation, the tetraploid nature of the genome of cultivated peanut also has been responsible for the slow progress in the area of developing genomic resources such as molecular markers and genetic maps. Until about 2005, it was speculated that molecular breeding approaches such as marker-assisted selection (MAS) may not be realized in routine plant breeding programs in the near future. Nevertheless, as a result of concerted efforts of the international peanut community, the past five years have witnessed significant progress in the area of *Arachis* genomics and some efforts have been initiated towards QTL mapping and molecular breeding for resistance/tolerance to biotic/abiotic stresses for peanut improvement (Varshney et al., 2010a). Examples are discussed in this article to demonstrate that peanut has joined the select group of legume crops where appropriate genomic resources have become available and MAS can be implemented in crop breeding.

2. Assembly of genetic resources

Generation of information about germplasm diversity for economically important traits in a given species is one of the most important pre-breeding activities. Judicious exploitation of available genetic diversity is the mainstay for further crop improvement. Because cultivated peanut originated very recently in evolutionary time, and has a narrow genetic variation as compared to other crop species, tapping available genetic diversity is essential for crop improvement.

The genebank at ICRISAT, India holds the largest collection (15,445 accessions from 93 countries) followed by the National Bureau of Plant Genetic Resources (NBPGR) (14,585 accessions), the Directorate of Groundnut Research (DGR) of the Indian Council of Agricultural Research (ICAR) (9024 accessions) in India; the Oil Crops Research Institute (OCRI) of the Chinese Academy of Agricultural Sciences (CAAS) (8083 accessions) and the Crops Research Institute of the Guangdong Academy of Agricultural Sciences (4210 accessions) in China; and the U.S. Department of Agriculture (USDA) (9917 accessions) along with small to medium collections at the Texas A&M University (TAMU) and

the North Carolina State University (NCSU) in the USA; the EMBRAPA-CENARGEN and the Instituto Agronomico de Campinas in Brazil; and the Instituto Nacional de Technologia Agropecuaria (INTA) and the Instituto de Botánica del Nordeste (IBONE) in Argentina (Table 1). Major centers holding wild *Arachis* species include the Texas A&M University (1200 accessions), the USDA (607 accessions); the NCSU (406 accessions) in USA; the EMBRAPA-CENARGEN, Brazil (1220 accessions); the ICRISAT in India (477 accessions) and the IBONE (472 accessions) in Argentina.

Availability of a large number of cultivated accessions poses a question about selection of appropriate lines for their utility in crop breeding. The 'core collection' (i.e. 10% of the entire germplasm collection) presents a manageable and cost-effective entry point into germplasm collections for identifying candidate genotypes with new sources of disease and pest resistance or abiotic stress tolerance (Brown, 1989). Therefore, a core collection comprising of 1704 *A. hypogaea* genotypes was developed at ICRISAT (Upadhyaya et al., 2003), which is similar to the U.S. core collection (831 accessions) (Holbrook et al., 1993). Following the same procedure, the Chinese core collection comprising of 576 accessions was developed (Jiang et al., 2008). However, the size of these core collections is too large for easy exploitation by breeders. Therefore, 'mini-core collections' (i.e. 10% of the core collections and 1% of entire germplasm collection) with 184, 112 and 298 accessions have been developed at ICRISAT (Upadhyaya et al., 2002), USDA/ARS (Holbrook and Dong, 2005), and China (Jiang et al., 2010), respectively.

Apart from the above collections, a composite collection of 1000 accessions also was developed based on available phenotypic characterization, geographic origin and taxonomic data. These accessions were used for genotyping at 20 SSR loci. Based on genetic diversity analysis and after taking the population structure in account, a set of 300 accessions has been identified as a "reference set" (Upadhyaya et al., 2002, 2003).

In addition to germplasm collections that represent naturally-occurring variation, induced mutant collections for tetraploid peanut also are being developed. Over 3400 mutant (M_2) lines generated by ethylmethane sulfonate (EMS) treatment have been used to screen for mutations in six genes using the technique of TILLING (targeting induced local lesions in genomes) (Knoll et al., 2011). TILLING is a reverse genetic technique that requires knowledge of genome sequence (McCallum et al., 2000). In peanut, the homeolog pairs *Ara h 1.01/1.02*, *Ara h 2.01/2.02* (encoding major allergen proteins), and *FAD2A/B* (encoding delta-12-desaturase, the enzyme primarily controlling oleic to linoleic acid ratio in seeds) were screened in the mutagenized population to discover gene knockouts or functional mutations in at least one of the two homeologs for each gene duplicate (Knoll et al., 2011). The TILLING population will be useful for functional genomics studies as well as to recover mutations of potential value for crop improvement.

3. Repertoire of genomic resources

Genomic resources such as molecular markers are powerful tools to characterize and harness the genetic variation present in the germplasm collection. Availability of selected genomic resources in *Arachis* species is listed below:

3.1. Molecular markers

Among all genomic resources, molecular markers have direct use for germplasm characterization, trait mapping and molecular breeding. Several marker systems have been developed during the last three decades. For instance restriction fragment length polymorphisms (RFLPs), random amplified polymorphic DNAs (RAPDs), amplified fragment length polymorphisms (AFLPs) and diversity arrays technology (DArT) markers have proved their utility from time to time (Gupta et al., 2010; Varshney et al., 2006). However, simple sequence repeats (SSRs) or microsatellites and single nucleotide polymorphism (SNP) markers are generally preferred for plant genetics and breeding applications. While SSR markers are multi-allelic, co-dominant and easy to use, the SNP markers are highly amenable to high-throughput genotyping approaches. Development and application of SNP markers, however, is still not routine in crop species and especially not in low-tech laboratories.

In the case of peanut, some efforts were made to use RFLP, RAPD and AFLP markers for diversity analysis (Bravo et al., 2006; Dwivedi et al., 2001; He and Prakash, 2001; Herselman, 2003; Hilu and Stalker, 1995; Kochert et al., 1996; Subramanian et al., 2000), genetic mapping (Burow et al., 2001; Garcia et al., 2005; Halward et al., 1993; Herselman et al., 2004; Leal-Bertioli et al., 2009; Milla, 2003) and trait mapping (Herselman et al., 2004). However, only few hundred SSR markers were available until 2005. The low diversity detected with SSR markers in the cultivated genepool demanded large-scale SSR marker development. As a result, aggressive efforts were made worldwide on development of SSR markers from SSR-enriched libraries, BAC-end sequences, EST sequences and transcript sequences generated

Table 1
List of major peanut germplasm collections.

Research centers	Cultivated genotypes	Wild genotypes	Total genotypes	Sources
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India	14,968	477	15,445	Hari D. Upadhyaya, ICRISAT, India
National Bureau of Plant Genetic Resources (NBPGR), Indian Council of Agricultural Research (ICAR), India	–	–	14,585	http://www.nbpgr.ernet.in/monthly_progress/cons_july11.pdf
Plant Genetic Resource Conservation Unit (PGRCU), Griffin, U. S. Department of Agriculture (USDA), USA	9310	607	9917	Noelle Barkley, PGRCU, USA
Directorate of Groundnut Research (DGR), ICAR, India	8960	64	9024	T. Radhakrishan, DGR, India
Oil Crops Research Institute (OCRI), Chinese Academy of Agricultural Sciences (CAAS), China	7837	246	8083	Jiang Huifang, OCRI, China
Crops Research Institute (CRI), Guangdong Academy of Agricultural Sciences (GAAS), China	4210	–	4210	Xuanquiang Liang, CRI, China
Instituto Nacional de Technologia Agropecuaria (INTA), Argentina	3534	106	3640	Guillermo Seijo, IBONE, Argentina
EMBRAPA and CENARGEN, Brazil	1200	1220	2420	José F. M. Valls, EMBRAPA, Brazil
Instituto Agronomico de Campinas, Brazil	2140	–	2140	David J. Bertioli, University of Brasília, Brazil
Texas A&M University (TAMU), USA	–	1200	1200	Charles E. Simpson, TAMU, USA
North Carolina State University (NCSU), USA	740	406	1146	Tom H. Stalker, NCSU, USA
Instituto de Botánica del Nordeste (IBONE), Argentina	–	472	472	Guillermo Seijo, IBONE, Argentina

by using 454/FLX sequencing technology. In summary, >6000 SSR markers have been generated during the past decade (Table 2). Most recently a set of highly informative SSR markers (199 SSRs with >0.50 PIC) were identified along with polymorphism features of 946 novel SSR markers (Pandey et al., 2011). It is anticipated that these informative markers will be very useful to accelerate molecular genetics and breeding studies in cultivated peanut.

In terms of other marker systems, a DArT platform comprising of ca. 15,000 features has been developed at DArT Pty Ltd (Australia) in collaboration with ICRISAT (India), CIRAD (France) and Catholic University of Brasília and EMBRAPA (Brazil). Use of DArT arrays with a range of genotypes representing diploid (AA, BB) and tetraploid (AABB) genome species showed a very low level of polymorphism in tetraploid genotypes and a moderate level of diversity among accessions from diploid genome species (Kilian, 2008; Varshney et al., 2010a). These results indicate that DArT markers may not be very useful for genetics and breeding applications in cultivated peanut, but they will be very informative in monitoring genome introgression from diploids into cultivated peanut lines.

In context of SNP discovery, although some studies have been undertaken on allele-specific sequencing of several genes in diploid and tetraploid species (Alves et al., 2008), limited success was obtained in detecting sequence variation in cultivated peanut genotypes. A significant effort has been made at the University of Georgia (Steve Knapp and Peggy Ozias-Akins, pers. comm.) where >2000 SNPs have been identified by comparing the 454/FLX transcript sequences of 17 genotypes. Based on these SNPs, one 1536-SNP Illumina GoldenGate SNP array has been developed. In parallel, one 768-SNP Illumina GoldenGate array based on SNPs identified between diploid genotypes for tentative orthologous genes (TOGs) has been developed at the University of California-Davis (Douglas Cook, pers. comm.). While these arrays are very informative for SNP genotyping in diploid species, homoeology between AA- and BB-genomes pose a constraint in using these arrays for tetraploid species. While efforts are underway to develop informative DArT and SNP markers in cultivated peanuts, SSR markers remain the best choice for genetic and breeding studies in cultivated peanut.

3.2. Genetic maps

Although efforts were initiated towards developing genetic maps as early as 1993, genome mapping in peanut has significantly increased since 2005. Initial genetic maps were developed based on mapping populations derived from the diverse parental genotypes. For instance, one F_2 population (87 lines) developed from a cross of *A. stenosperma*×*A. cardenasii* was used to develop the genetic map with 117 RFLP loci on 11 linkage groups with a genome coverage of 1063 cM (Halward et al., 1993) (Table 4). Subsequently, a backcross population (78 BC_1F_1 lines) generated from the cross of TxAG-6, a synthetic amphidiploid line ([*A. batizocoi*×(*A. cardensii*×*A. diogoi*)]^{4x}) and Florunner was used to develop the genetic map with 370 RFLP loci onto 23 linkage groups and 2210 cM genome coverage (Burow et al., 2001). The common markers mapped in both the crosses showed a high degree of co-linearity between the diploid and tetraploid chromosomes (Burow et al., 2001; Jesubatham and Burow, 2006).

In terms of the second generation genetic maps based on SSR markers, an F_2 population obtained from a cross between two diploid species with AA genome (*A. duranenis* and *A. stenosperma*) was used. As a result, the AA-genome map comprised of 170 SSR loci on 11 linkage groups covering 1231 cM was developed (Moretzsohn et al., 2005). This map was saturated further and the current map based on this population has 369 marker loci (Leal-Bertioli et al., 2009). In parallel, a high density AA-genome genetic map (*A. duranenis* and *A. duranenis*) has also been developed with 2319 markers (971 SSRs, 221 single stranded DNA conformation polymorphism (SSCP) markers and 1127 SNPs) mapped on 10 linkage groups (Nagy et al., 2010a). For the BB-genome, a genetic map with 149 SSR loci on 11 linkage groups covering 1294 cM genome coverage has been developed based on a F_2 population (93 lines) derived from the cross between *A. ipäensis* (KG30076) and *A. magna* (KG30097) (Table 4).

Although low genetic diversity in the cultivated peanut gene pool was a serious bottleneck until recently in developing the genetic maps using mapping populations of cultivated peanut lines, availability of large numbers of SSR markers facilitated the development of the first SSR-based genetic map based on a RIL population derived from TAG 24×ICGV 86031 (Varshney et al., 2009c). This map was further saturated with more markers and the current map based on the above mapping population has 191 SSR loci onto 20 linkage groups with 1785 cM genome coverage (Ravi et al., 2011). Subsequently, several other genetic maps for cultivated peanut have become available recently (Table 4). With an objective of estimating the marker order for a maximum number of marker loci based on a single map, a composite map comprising of 175 marker loci has been developed by Hong et al. (2010a). Although several genetic maps have become available for cultivated peanut, SNP markers have not yet been integrated into these maps. Availability of

Table 2
List of *Arachis* SSR markers available in public domain.

Marker series	Markers developed	References	Research Institute/University
Ah, Lec	26	Hopkins et al., 1999	USDA-ARS, USA
pPGPseq, pPGSseq	226	Ferguson et al., 2004	University of Georgia, USA/Cornell University, USA
Ah, Lec, Ap	32	Palmieri et al., 2002, 2005; Gimenes et al., 2007	Universidade Estadual Paulista (UNESP), Brazil
PM	103	He et al., 2003; Luo et al., 2005b	USDA-ARS/Tuskegee University, USA
AC, Ah, gi, RN, TC, Seq	338	Moretzsohn et al., 2004, 2005	EMBRAPA, Brazil/USDA-ARS, USA
S	103	Nelson et al., 2006	University of Western Australia, Australia
LG, Lup	188	Proite et al., 2007	University of Brasília/EMBRAPA, Brazil
RN, RM	123	Wang et al., 2007	Shandong Peanut Research Institute, China
Lup, Dal, Stylo, Ades, Amor, Chaet, IPAHM, ICGM	178	Mace et al., 2007; Cuc et al., 2008; Gautami et al., 2009	International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India
EM	290	Liang et al., 2009b	Guangdong Academy of Agricultural Sciences, China/USDA-ARS, USA
ES	685	Hong et al., 2010b	Guangdong Academy of Agricultural Sciences, China
PM	138	Yuan et al., 2010	Shandong Peanut Research Institute, China/USDA-ARS/Tuskegee University, USA
F, H, PD	94	Song et al., 2010	Shandong Academy of Agricultural Sciences, China
AHS, AhTE	1571	Koilkonda et al., 2011	Kazusa DNA Research Institute (KDRI), Japan
GM613-GM709	97	Nagy et al., 2010a	University of Georgia, USA
GM710-GM2847	2138	Nagy et al., 2010a	University of Georgia, USA
Fl, Ahl	1152	Douglas R Cook, unpublished	University of California, USA

Table 3
List of the main populations used in *Arachis* genomics research.

Population	Lines/types	Segregating traits	Research Institute/University
AA genome			
A. stenosperma × *A. cardenasii*	F_2	–	UGA, NCSU
A. stenosperma × (*A. stenosperma* × *A. cardenasii*)	44 BC_1F_1	–	NCSU
A. kuhlmanni × *A. diogoi*	179 F_2	Tomato spotted wilt virus (TSWV)	NCSU
A. duranensis (PI 475887) × *A. duranensis* (Grif 15036)	94 F_2	–	UGA
A. duranensis × *A. stenosperma*	87 RILs	Late leaf spot resistance, transpiration response to drought stress, various aspects of plant morphology	EMBRAPA, UB
BB genome			
A. ipäensis × *A. magna*	93 RILs	Rust, various aspects of plant morphology	EMBRAPA, UB
A. batizocoi (PI 298639) × *A. batizocoi* (PI 468327)	94 F_2	Resistance to nematode, early leaf spot and late leaf spot	TAMU
AABB genome			
A. hypogaea × *A. cardenasii*	46 $F_{10}C_9$	Late leaf spot resistance	ICRISAT*, NCSU
A. hypogaea cv. IAC-Runner 886 × (*A. ipäensis* × *A. duranensis*)	93 RILs	Rust and late leaf spot resistance, various morphological and domestication traits	EMBRAPA, UB
ICG 12991 × ICGV-SM93541	200 F_2	Aphid vector of groundnut rosette disease	ARC-GCI
TAG 24 × ICGV 86031	318 RILs	Drought related traits	ICRISAT†
ICGS 44 × ICGS 76	188 RILs	Drought related traits	ICRISAT†
ICGS 76 × CSMG 84-1	176 RIL	Drought related traits	ICRISAT†
TAG 24 × GPBD 4	266 RILs	Late leaf spot and rust resistance	UAS-D
TG 26 × GPBD 4	146 RILs	Late leaf spot and rust resistance	UAS-D
Tamrun OL01 × BSS 56	88 RILs	Yield parameter and oil content	Texas Tech University
Yueyou 13 × Zhen Zhuhei	142 RILs	Protein content	GAAS
Yueyou 13 × FU 95-5	84 RILs	Oil content	GAAS
Yueyou 13 × J 11	136 RILs	Resistance to *Aspergillus flavus* and aflatoxin contamination	GAAS
CG7 × ICGV-SM 94584	111 F_5	Groundnut rosette disease	ICRISAT*
JL24 × ICGV-SM 94584	219 F_5	Groundnut rosette disease	ICRISAT*
CG7 × ICGV-SM 90704	338 F_4	Groundnut rosette disease	ICRISAT*
Chalimbana × ICGV-SM 90704	597 F_4	Groundnut rosette disease	ICRISAT*
JL24 × ICGV-SM 90704	151 F_4	Groundnut rosette disease	ICRISAT*
ICGV 93437 × ICGV 94114	107 F_5	Leaf rust	ICRISAT*
ICGV 93437 × ICGV 95342	466 F_5	Leaf rust	ICRISAT*
ICGV 93437 × ICGV-SM 95714	105 F_5	Early leaf spot	ICRISAT*
ROBUT 33–1 × ICGV-SM 95714	186 F_5	Early leaf spot	ICRISAT*
Tifrunner × Bailey High O/L	400 F_5	Oil quality, early and late leaf spot	USDA-ARS, UGA, NCSU
Tifrunner × C76-16, Florida-07 × C76-16	400 F_5	Drought tolerance and reduced PAC	USDA-ARS, UGA, NCSU
Tifrunner × NC 3033, Florida-07 × NC 3033	400 F_5	*Cylindrocladium* black rot (CBR) disease	USDA-ARS, UGA, NCSU
Tifrunner × SPT 06-06, Florida-07 × SPT 06-06	400 F_5	Early and late leaf spot	USDA-ARS, UGA, NCSU
Florida-07 × Bailey High O/L	400 F_5	Oil quality, white mold disease	USDA-ARS, UGA, NCSU
Tifrunner × Olin	550 F_3	Oil quality, maturity	USDA-ARS, UGA, UFL
Tifrunner × NM Valencia A	225 F_3	Tomato spotted wilt virus (TSWV), maturity	USDA-ARS, UGA, UFL
Tifrunner × Florunner	700 F_3	Tomato spotted wilt virus (TSWV)	USDA-ARS, UGA, UFL
Florida-07 × Olin	450 F_3	Sclerotinia	USDA-ARS, UGA, UFL
Florida-07 × NM Valencia A	270 F_3	Oil quality, tomato spotted wilt virus (TSWV), *Sclerotium rolfsii*	USDA-ARS, UGA, UFL
Florida-07 × Florunner	460 F_3	Oil quality, tomato spotted wilt virus (TSWV), *Sclerotium rolfsii*	USDA-ARS, UGA, UFL
Florida-07 × SSD6, Tifrunner × SSD6	66-400 F_3	Early and late leaf spot	USDA-ARS, UGA
PI 158839 (554CC) × Tifguard	400 F_5	Nematode resistance, drought tolerance	USDA-ARS, UGA
Gregory × Tifguard	78 RILs	Nematode resistance, late leaf spot, seed traits	USDA-ARS, UGA
SunOleic 97R × NC94022	354 RILs	Tomato spotted wilt virus (TSWV), oil quality	USDA-ARS
Tifrunner × GT-C20	246 RILs	Tomato spotted wilt virus (TSWV), early and late leaf spot, maturity	USDA-ARS
Yueyou 13 × Zhen Zhuhei, Zhen Zhuhei × Yueyou 13	156 F_2	Dark purple testa	GAAS
A. hypogaea × (*A. batizocoi* × (*A. cardenasii* × *A. diogoi*))	78 BC_1F_1	Wild introgression	UGA
A. hypogaea cv. Fleur11 × (*A. ipäensis* × *A. duranensis*)	59 BC_2	Wild introgression	CIRAD, ISRA CERAAS

UGA: University of Georgia, USA; NCSU: North Carolina State University, USA; EMBRAPA: Brazilian Agricultural Research Corporation, Brazil; UB: University of Brasília, Brazil; TAMU: Taxas A & M University, USA; ARC-GCI: Agric. Res. Council-Grain Crops Institute, South Africa; ICRISAT: International Crops Research Institute for the Semi-Arid Tropics, India† & Malawi*; UAS-D: University of Agricultural Sciences-Dharwad, India; GAAS: Guangdong Academy of Agricultural Sciences, China; USDA-ARS: U. S. Department of Agriculture-Agricultural Research Station, Tifton, USA; UFL: University of Florida, USA; CIRAD: Centre de Coopération Internationale en Recherche Agronomique pour le Développement, France; ISRA-CERAAS: Institut Sénégalais d Recherches Agricoles-Centre d Etude Régional pour Ľamélioration de Ľadaptation ála Sécheresse, Senegal.

Table 4
Details of some major genetic maps constructed in *Arachis* species.

Population	Population size	Marker loci mapped	References
A-genome genetic maps			
A. stenosperma × *A. cardenassi*	F_2	117 RFLPs	Halward et al., 1993
A. kuhlmanni × *A. diogoi*	179 F_2	102 AFLPs	Milla, 2003
A. stenosperma × (*A. stenosperma* × *A. cardenassi*)	44 BC_1F_1	167 RAPDs	Garcia et al., 2005
A. duranensis × *A. stenosperma*	F_2	170 SSRs	Moretzsohn et al., 2005
A. duranensis × *A. stenosperma*	93 F_2	369 markers (SSR, AFLP, SNP, RFLP, SCAR)	Leal-Bertioli et al., 2009
A duranensis × *A. duranensis*	94 F_2	2319 markers (1127 SNPs, 971 SSRs, 221 SSCPs)	Nagy et al., 2010a
B-genome genetic maps			
A. ipäensis × *A. magna*	93 F_2	149 SSRs	Gobbi et al., 2006, Moretzsohn et al. 2009
A. batizocoi PI 298639 × *A. batizocoi* PI 468327	94 F_2	449 SSRs	Guo et al., 2010
AB genome genetic maps			
A. hypogaea × *A. cardenasii*	46 $F_{10}C_9$	RAPDs	Garcia et al., 1995
A. hypogaea × (*A. batizocoi* × (*A. cardenasii* × *A. diogoi*))	78 BC_1F_1	370 RFLPs	Burow et al., 2001
ICG 12991 × ICGV-SM 93541	200 F_2	12 AFLPs	Herselman et al., 2004
TAG 24 × ICGV 86031	318 RILs	191 SSRs	Varshney et al., 2009c; Ravi et al., 2011
A. duranensis × (*A. ipäensis* × *A. duranensis*)	88 BC_1F_1	298 SSRs	Foncéka et al., 2009
Yueyou 13 × Zhen Zhuhei	142 RILs	132 SSRs	Hong et al., 2010a
Yueyou 13 × FU 95-5	84 RILs	109 SSRs	Hong et al., 2010a
Yueyou 13 × J 11	136 RILs	46 SSRs	Hong et al., 2010a
TAG 24 × GPBD 4	266 RILs	188 SSRs	Khedikar et al., 2010; Sujay et al., 2011
ICGS 44 × ICGS 76	188 RILs	82 SSRs	Gautami et al., in press
ICGS 76 × CSMG 84-1	176 RIL	119 SSRs	Gautami et al., in press
TG 26 × GPBD 4	146 RILs	181 SSRs	Sarvamangala et al., 2011, Sujay et al., 2011
SunOleic 97R × NC94022	190 RILs	170 SSR, 2 CAPS	Chen et al. 2010
			Chen et al. 2010
Tifrunner × GT-C20	158 RILs	238 SSR, 1 CAPS	

large number SNP markers will facilitate development of SNP-based genetic maps in coming years.

3.3. BAC libraries and physical map

Large insert genomic DNA libraries, such as bacterial artificial chromosome (BAC) libraries, provide a platform for physical mapping, map-based cloning of the genes for traits of interest, analysis of gene structure, and function and genome sequencing. In the case of peanut, the first BAC library with 182,784 clones was developed for the tetraploid peanut cultivar Florunner (Yüksel and Paterson, 2005). Recently, two BAC libraries have been developed for the probable ancestors of AA- (*A. duranensis*) and BB- (*A. ipäensis*) genomes of tetraploid peanut (Guimarães et al., 2008). These libraries possess ca. 7.4 and ca. 5.3 genome equivalents respectively for AA and BB genomes with low organelle contamination.

Table 5
List of QTLs identified for some economically important traits in peanut.

Traits studied	QTLs identified	Phenotypic variance explained (%)	References
Biotic stress			
Late leaf spot (LLS)	39	1.70–67.98	Khedikar et al. 2010, Sujay et al., in press
Leaf rust	27	1.70–82.96	Khedikar et al. 2010, Sujay et al., in press
Resistance to *Aspergillus flavus* invasion	6	6.2–22.7	Liang et al. 2009a
Aphid vector of groundnut rosette disease	8	1.18–76.16	Herselman et al. 2004
Abiotic stress			
Transpiration (T)	15	4.36–18.17	Varshney et al. 2009c, Ravi et al. 2011, Gautami et al., in press
Transpiration efficiency (TE)	14	4.47–18.12	Varshney et al. 2009c, Ravi et al. 2011, Gautami et al., in press
Specific leaf area (SLA)	13	3.48–13.29	Varshney et al. 2009c, Ravi et al. 2011
Leaf area (LA)	4	7.24–11.51	Varshney et al. 2009c, Ravi et al. 2011
SPAD chlorophyll meter reading (SCMR)	29	5.72–19.53	Varshney et al. 2009c, Ravi et al. 2011
Biomass	7	4.25–20.32	Varshney et al. 2009c, Ravi et al. 2011
Canopy conductance (ISC)	7	3.28–22.24	Varshney et al. 2009c, Ravi et al. 2011
Total dry matter (TDM)	7	4.34–22.39	Varshney et al. 2009c, Ravi et al. 2011, Gautami et al., in press
Agronomic traits			
Shoot dry weight (ShDW)	11	5.03–22.09	Varshney et al. 2009c, Ravi et al. 2011, Gautami et al., in press
Pod weight (PW)	7	4.17–8.73	Varshney et al. 2009c, Ravi et al. 2011
Seed weight (SW)	5	4.18–8.22	Varshney et al. 2009c, Ravi et al. 2011
Haulm weight (HW)	6	3.78–33.66	Varshney et al. 2009c, Ravi et al. 2011
Harvest index (HI)	3	6.39–40.10	Gautami et al., in press
Pod mass/plant	3	13.1–18.3	Liang et al. 2009a
Mature pods/plant	3	11.9–12.3	Liang et al. 2009a
Number of branches	7	8.1–17.3	Liang et al. 2009a
Number of fruit branches	1	17.5	Liang et al. 2009a
Height of main axis	7	8.2–12.8	Liang et al. 2009a
Stem diameter	4	7.8–24.1	Liang et al. 2009a
Leaf length, width and length/width ratio	7	12.4–18.9	Liang et al. 2009a
Yield parameters	5	9.19–17.69	Selvaraj et al. 2009
Oil content	7	1.5–9.5	Liang et al. 2009a, Selvaraj et al. 2009, Sarvamangala et al. 2011
Oil quality	10	1.4–9.7	Sarvamangala et al. 2011
Protein content	10	1.5–13.4	Liang et al. 2009a, Sarvamangala et al. 2011

The above mentioned BAC libraries have been used for a variety of applications. For instance, BACs derived from the AA genomic library have been used for the isolation of full length Resistance Gene Analogs (RGAs) and the first complete retroelement in *Arachis*-FIDEL (Nielen et al., 2010). In addition, an extended version of the AA library is currently being used to construct a physical map for the AA genome at the University of Georgia using the "overgo hybridization"

Table 6
List of expressed sequence data (ESTs) generated at different developmental stages and under various stresses in peanut.

Traits/stages of EST generation	Number of ESTs	GenBank ID	EST–SSRs developed	References
Under water stress conditions	1235	–	–	Jain et al. 2001
Wild and cultivated genotypes	1350	–	44	Luo et al., 2005a, 2005b
Leaves and roots of wild genotype (*A. stenosperma*)	8785	EH041934-048197	188	Proite et al. 2007
Aspergillus infection and aflatoxin contamination	21,777	ES702769-724546	–	Guo et al. 2008
Tomato spotted wilt virus and leaf spots	16,931	ES751523-768453	290	Guo et al. 2009
Different developmental stages	12,000	–	94	Song et al. 2010
Seed development	17,000	–	–	Bi et al. 2010
Roots, leaves and seedlings	10,102	–	3187	Koilkonda et al. 2011
Root, leaf and developing pod	84,229	–	2138	Nagy et al. 2010a
Total ESTs	173,405		5941	

technique (A. Paterson, pers. comm.). Very recently, 36,435 (28.6 Mbp) BAC-end sequences generated from the AA diploid library (*A. duranensis*) as well as tetraploid (Tifrunner) genotypes by the University of California, USA were further used for marker development (Wang et al. 2011 unpublished). The availability of BAC libraries from the allopolyploid cultivated peanut and the two wild ancestors have allowed the comparison of these genomes regarding microsynteny and repetitive DNA contents and the isolation of desirable genes and alleles from cultivated and wild species. This will facilitate undertaking genome sequencing in peanuts.

3.4. Transcriptome resources

Transcriptome resources are an alternative source to genome sequence in species where genome sequence is not available. Trancriptomic resources e.g. expressed sequence tags (ESTs), can be used to understand genome dynamics as well as for applied aspects such as development of gene-based markers and maps, transcript profiling for identification of candidate genes involved in expression of traits of interest, and the identification of transcriptional changes during plant immunity responses. The use of microarray and next generation sequencing (NGS) technologies (Varshney et al., 2009d) offers greater potential in such transcriptional profiling, with deep coverage sequence data, unbiased transcript representation, appropriate for gene discovery, including *de novo* detection of rare transcripts and their expression, as well as alternative splicing and gene sequence polymorphism detection. Such potential seen in NGS technologies makes them well placed for applications in functional genomic analysis in *Arachis*. One important application is focusing on host–pathogen interactions, as an approach for increasing our understanding of the molecular mechanisms controlling pathogen-associated molecular pattern (PAMP), triggered immunity (PTI), and effector triggered immunity (ETI) in plants.

Currently, a total of 173,405 peanut ESTs are available in the public domain in NCBI (www.ncbi.nlm.nih.gov, 25th April, 2011) (Fig. 1). Many of these have been generated from: (i) developing seeds for identification of resistance-related genes involved in defense mechanisms against *Aspergillus* infection and subsequent aflatoxin contamination (21,777 ESTs, Guo et al., 2008), (ii) leaf tissues of peanut cultivars resistant and susceptible to tomato spotted wilt virus and leaf spots (16,931 ESTs, Guo et al., 2009), (iii) diverse seed development stages (17,000 ESTs, Bi et al., 2010), (iv) different plant developmental stages (12,000 ESTs, Song et al., 2010), (v) root, seeds and leaves (10,102 ESTs, Koilkonda et al., 2011) and (vi) root tissues in *A. stenosperma* resistant to *Meloidogyne arenaria* (8000 ESTs, Proite et al., 2007) (Table 6). In addition to the above, some small-scale efforts have contributed to the EST repertoire in peanut (Jain et al., 2001; Luo et al., 2005a). Analysis of all publicly available ESTs using an in-house developed script showed a huge variation in length of submitted ESTs ranging from 37 bp to 2038 bp (average 534.4 bp) (Sarwar Azam, pers. Comm.).

The majority of the ESTs (83.01%) were of medium length (300–800 bp) while 11.48% were of small length (<300 bp) and 5.51% were of larger length (>800 bp) (Fig. 2). As expected from above-mentioned discussions, the majority of the public domain ESTs have come from seeds (57.26%), followed by roots (23.22%) and leaves (19.02%). Most recently, the University of Georgia (UGA), USA has developed about 1 million reads representing >350 Mb of transcript sequences from 17 tetraploid genotypes using 454-titanium sequencing technology. After combining these data with publicly-available sequences submitted under NCBI BioProject Accession PRJNA49471, a consensus transcriptome assembly has been developed for tetraploid peanut which is comprised 211,244 contigs. The largest contig was of 3907 bp with an average length of 563 bp and N50 (the length N for which 50% of all bases in the sequences are in a sequence of length L<N) of 609 bp. This assembly should be useful for aligning the transcript reads using next generation sequencing technologies (Varshney et al., 2009d) from different genotypes, and alignments can be used for SNP discovery. Another study utilizing Roche-454 GS FLX Titanium technology was conducted on two wild *Arachis* species under biotic (*A. stenosperma*/leaf spot fungi [resistant response]) and abiotic stress (*A. duranensis*/hydric stress [tolerant response]), producing a total of 743,232 ESTs. The average length of sequence reads was 280 bp, producing genome coverage of 85 and 78 Mbp for each respective genome. A total of 39,626 unigenes were functionally annotated for the two species, with those displaying *in silico* differential gene expression between stressed and non-stressed plants identified (Guimarães et al., 2011). Large scale SSR motif and SNP collections generated in this study are well suited for high-throughput mapping purposes, given that the above

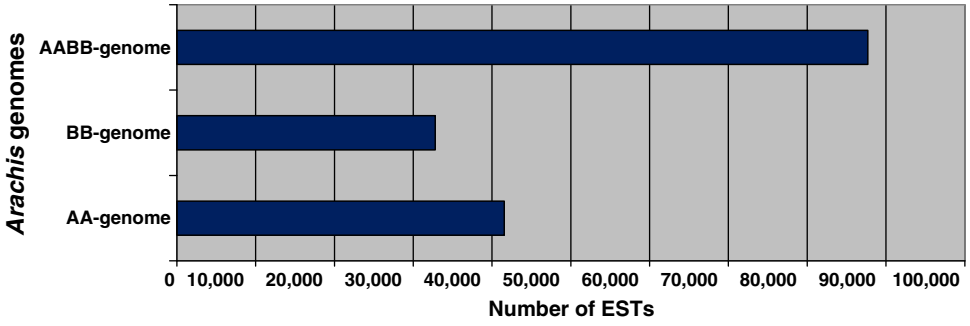

Fig. 1. Publicly available ESTs in different *Arachis* species.

species are also parentals of a diploid reference mapping population (Moretzsohn et al., 2005). The genomic resource developed in this study will, in association with other genetic and genomic tools already developed for wild *Arachis* species, contribute to accelerate genetics and breeding of peanut and the understanding of stress tolerance/resistance mechanisms in peanut and other legumes.

Developed EST or transcript sequences have also been used for identification of SSRs and SNPs and their conversion into genetic markers. For instance, a set of 44 EST–SSR markers was developed by Luo et al. (2005a). Similar efforts were made by several other groups (Table 2). In addition to SSRs, large scale SNP mining has been conducted based on 454/FLX transcript sequences. For instance, UGA has identified over 2000 SNPs after comparing the 454 sequence reads generated from 17 genotypes. Based on these SNPs, a GoldenGate assay for 1536 SNPs has also been developed. Molecular markers derived from gene/transcript sequences, generally referred as genic molecular markers (GMMs), on one hand, provide estimates on functional diversity in the germplasm collection and can be used to develop the 'functional' or 'perfect' markers (Varshney, 2010). On the other hand, transcript-based markers generally have low polymorphism compared to the markers derived from genomic DNAs (Varshney et al., 2005b, 2010b, 2010c).

3.5. Whole genome sequences

In recent years, genome sequences have become available for several crop species including rice (2005), maize (2009), sorghum (2009), and soybean (2010). The peanut (4X) genome is ~20 times larger than *Arabidopsis thaliana,* and 2–6 times larger than rice, sorghum or soybean. Because of its large genome size and amphidiploid nature of its genome, sequencing for peanut genome is only just been initiated. A Peanut Genome Consortium (PGC) (http://www.peanutbioscience.com/peanutgenomeproject.html) has been formed having the following specific goals: 1) a high quality chromosome scale draft of a tetraploid (cultivated species) as the reference genome sequence, plus high density maps of both progenitor and synthetic amphidiploids genomes; 2) high-throughput transcriptome characterization of the reference tetraploid cultivar; 3) characterization of gene space in amphidiploid and diploid (progenitor species) germplasm, 4) phenotypic association with mapped genetic markers, and 5) interactive bioinformatic resources for data curation and application in a breeders's toolbox to enable molecular breeding for enhancing peanut yielding ability, optimizing resistance to diseases and insects, tolerance to environmental stresses, and improved quality traits. In this context, in collaboration with BGI-Shenzhen (China), the Peanut Genome Project (PGP) is initiating sequencing of the peanut genome. The strategy is an integrated one of whole-genome shotgun sequencing of tetraploid and diploid genotypes, highly multiplexed BAC-by-BAC sequencing of one tetraploid cultivar (Tifrunner) with HiSeq2000 technology, and genetic ordering based on low-coverage sequencing of recombinant inbred lines. In parallel, efforts have been initiated at the University of California-Davis, USA (Richard Michelmore, pers. comm.) for generating ultra-high density genetic maps through low coverage, shotgun sequencing of diploid and tetraploid mapping populations and of the reference germplasm set (Froenicke et al., 2011). These populations include *A. duranensis* × *A. stenosperma* (AA genome); *A. ipäensis* × *A. magna* (BB genome); *A. hypogaea* cv. IAC Runner × synthetic amphidiploid of the two progenitor species and the reference set consisting of 300 accessions selected based on SSR genotyping data to represent the diversity of the international peanut germplasm collection (Upadhyaya et al., 2003), and the US mini-core collection (Holbrook and Dong, 2005) along with historical and modern US cultivars. The SNPs identified in the diversity panel will be analyzed for linkage disequilibrium (LD) and used to refine the genetic bins generated from the RIL segregation data of different mapping populations. These efforts are expected to both assist and complement the assembly of the reference genome sequence for peanut by validating genome assemblies and providing chromosomal genetic coordinates for contigs and scaffolds. It is, therefore, anticipated that a draft genome sequence along with extensive genome and trancriptome information will be available for peanut community within the next few years.

4. Application of genomic resources

Genomic resources developed as above are useful for accelerating both basic and applied research for peanut improvement (Fig. 3). Some selected areas demonstrating application of genomic resources are given below:

4.1. Identification of candidate genes

Apart from marker development and preparation of gene-based genetic maps, ESTs can be used for transcript profiling to identify the candidate genes for trait of interest as well as development of microarray to study differential expression of different genes at varied growth stages.

Identification of peanut genes were reported for encoding arachins and conglutins after sequencing and analyzing randomly selected clones from a cDNA library (Yan et al., 2005); cloning of a full length cDNA encoding aldehyde oxidase from leaves of peanut revealed 326 bp 5′ untranslated region and a 128 bp 3′ untranslated region including a poly (A) tail of 21 nucleotides (Yang et al., 2009); and identification and cloning of resistance gene to TSWV in cultivated peanut along with characterization of two peanut *oxalate oxidase genes* (Chen et al., 2008a, 2008b, 2011). Through a similar approach, Guo et al. (2008) identified nine and eight resistance related genes

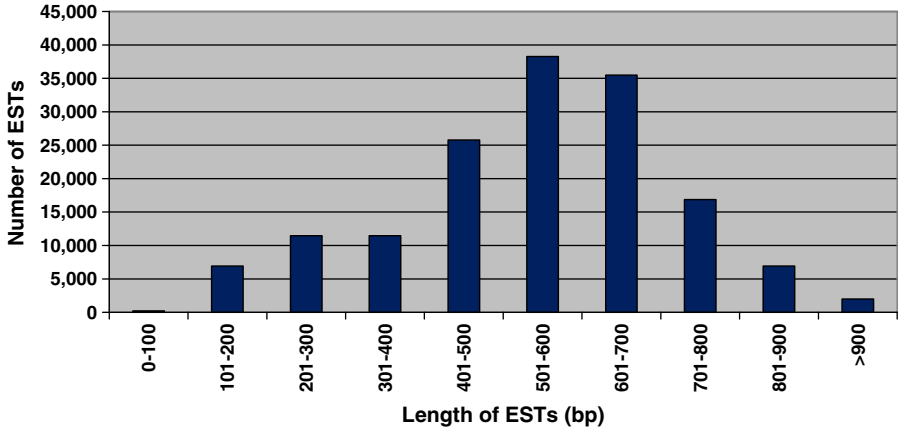

Fig. 2. Length distribution of publicly available ESTs in peanut.

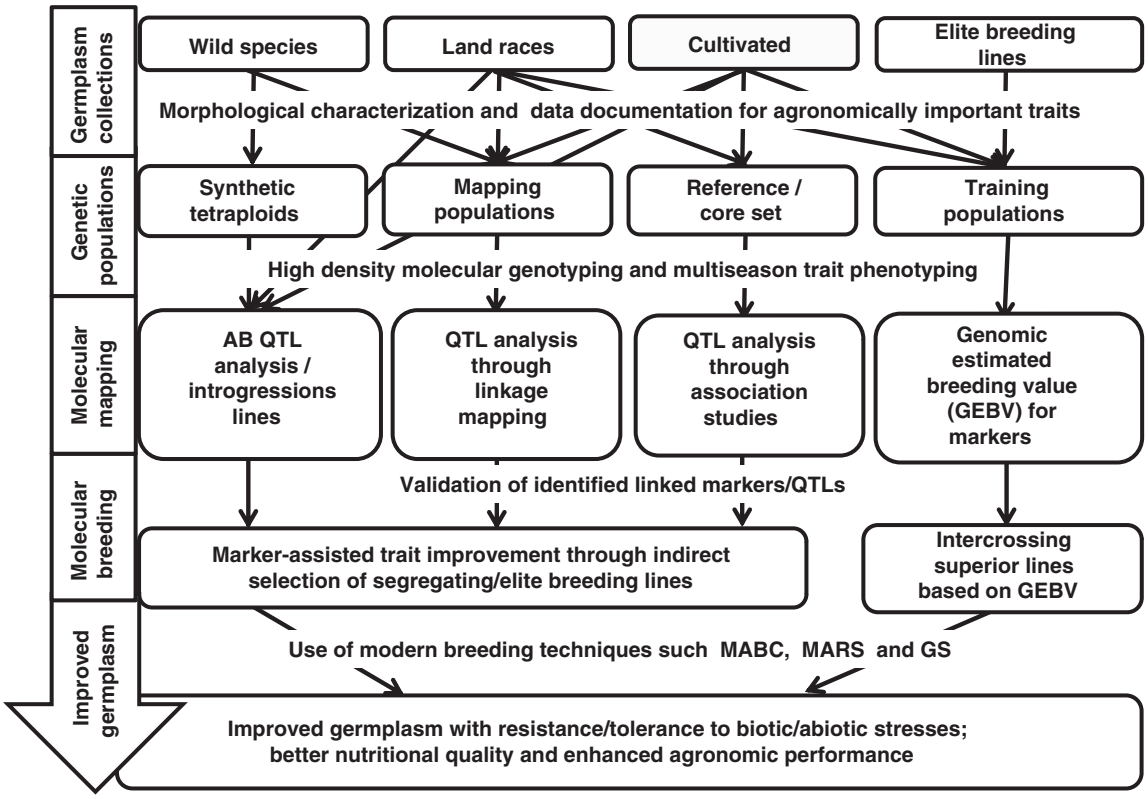

Fig. 3. A diagrammatic representation of different activities involved in improving the cultivated peanut.

with significant upregulation for *Aspergillus* infection and subsequent aflatoxin contamination, respectively, from the cDNA libraries obtained from 'GT-C20' and 'Tifrunner'.

Wild relatives of peanut have also been targeted as sources of disease resistant genes because they have high genetic diversity and have been selected during evolution in a range of environments and biotic stresses (Stalker and Simpson, 1995). Using in silico subtraction of ESTs and microarray analysis, eight genes were identified as differentially expressed in *A. stenosperma* roots during its resistance response to *M. arenaria.* The expression profile of the three most differentially expressed genes Auxin Repressed Protein (AsARP), Cytokinin Oxidase (AsCKX), Metallothionein Type 2 (AsMET2) was distinct in the resistant and susceptible species both after, and sometimes even before, challenge with nematodes, as shown by qRT-PCR and northern-blot analysis (Guimarães et al., 2010, 2011). In a separate study, Wang et al. (2010) adopted an integrated proteomic approach for identification of differentially expressed proteins for *A. flavus* infection under normal and drought conditions followed by further validation by real time RT-PCR analysis. The study identified a total of 29 protein spots showing differential expression between resistant and susceptible cultivars in response to *A. flavus* attack under drought stress. A significant decrease or down regulation of trypsin inhibitor caused by *A. flavus* in the resistant cultivar was also observed. In addition, proteomics approaches also helped in identification of drought tolerant cultivars (Basha, 1979; Katam et al., 2007), identification of cultivars with allergen proteins in seed (Kottapalli et al., 2008) and unique proteins related to metabolism in leaf (Katam et al., 2010).

Development of a large number of ESTs paved the way for the origin of microarrays to study differential expression of genes at various growth stages in a large set of genotypes. For example, the use of such arrays for expression profiling in a variety of peanut tissues to understand the regulatory role of microRNAs (miRNAs), which are known to control gene expression at the post-transcriptional level in peanut, resulted in identification of 14 novel miRNA families as well as 75 conserved miRNAs (Payton et al., 2009). Similarly, in an initial study, 25 potential unigenes associated with stress to either drought or to *A. parasiticus* were identified by monitoring up/down regulation using a microarray of 400 unigenes (Luo et al., 2005b). Another gene expression profiling study dealing with resistant and susceptible peanut cultivars infected with a mixture of *Aspergillus flavus* and *parasiticus* spores, 62 genes in resistant cultivars were identified that were up-expressed in response to *Aspergillus* infection (Guo et al., 2011). In addition, 22 putative *Aspergillus*-resistance genes were also identified, which were constitutively up-expressed in the resistant cultivar in comparison to the susceptible cultivar.

4.2. Trait mapping

Trait mapping is essential to identify tightly linked markers to economically agronomical traits through linkage/association mapping. This is the first step towards molecular breeding because these mapped gene/QTLs controlling desired traits either can be introgressed individually or as several genes/QTLs pyramided into an elite cultivar using tightly linked/perfect/functional markers (Fig. 3).

Several mapping populations were developed using diverse parents for a combination of traits in peanut by different research groups (Table 3). Initial mapping populations were developed in order to map the maximum number of loci in a single map by selecting parents with diverse origins. Later, mapping populations were developed targeting economically important traits such as biotic stresses (Tomato spotted wilt virus, early leaf spot, late leaf spot, leaf rust, aphid vector of groundnut rosette disease, cylindrocladium black rot disease, sclerotinia and nematode resistance, tomato spotted wilt virus), abiotic stress (drought tolerance), nutritional quality (aflatoxin contamination, oil content, oleic acid) and several agronomic traits. Attempts to map economically important traits prior to the availability of SSR markers in peanut were through bulk segregant analysis (BSA). BSA was used for identifying linked marker for nematode resistance (Burow et al., 1996; Garcia et al., 1996) and aphid vector of groundnut

rosette disease (Herselman et al., 2004) using RAPD and AFLP markers, respectively (Table 5). The above strategy was also applied for mapping the yield and yield parameters with SSR markers (Liang et al., 2009a; Selvaraj et al., 2009). However in recent years due to availability of advanced mapping populations such as RILs and relatively large number of molecular markers, linkage-mapping based marker analysis has been undertaken to identify the QTLs for drought tolerance related traits (Gautami et al., in press; Ravi et al., 2011; Varshney et al., 2009c), resistance to foliar disease (Khedikar et al., 2010; Sujay et al., in press) and nutritional quality traits (Sarvamangala et al., 2011). By using multi-environment phenotyping data for drought tolerance traits, 153 main-effect and 25 epistatic QTLs were identified (Gautami et al., in press; Ravi et al., 2011; Varshney et al., 2009c). On the other hand, one major QTL each for leaf rust (55.2% PVE, Khedikar et al., 2010; 82.96% PVE, Sujay et al., in press) and LLS (67.98% PVE, Sujay et al., in press) was detected. Some selected examples on trait mapping are summarized in Table 5. Although in some cases like resistance to nematode (Nagy et al., 2010b), leaf rust (Khedikar et al., 2010), LLS (Sujay et al., in press) and high-oleate trait (Chen et al., 2010; Chu et al., 2009) diagnostic molecular markers are available for deployment in molecular breeding, tightly linked molecular markers for are yet to be identified for several other important traits like ELS, GRD, etc. Availability of more genomic resources like SNPs, genome sequence, however, will accelerate trait mapping efforts in the near future.

4.3. Marker-assisted breeding

Introgression of recessive genes and pyramiding of multiple genes is very difficult using conventional breeding methods. However, marker-assisted selection (MAS) has proved its utility in several crops to overcome such problems and many genes can be pyramided either for the same trait or for different traits along with faster recurrent parent genome recovery through intense background selection. Furthermore, MAS can be used to introgress many recessive genes in less time than is possible through conventional breeding.

In the case of peanut, some efforts have already been initiated to use the molecular markers in breeding programs. Molecular markers linked with root-knot nematode (*M. arenaria*) resistance, introgressed through the amphidiploid pathway into cultivated peanut (Simpson, 2001), have been relatively easy to identify due to sequence divergence between diploid and tetraploid genomes (Chu et al., 2007a; Nagy et al., 2010b). These markers have been used during marker-assisted backcrossing (MABC) to screen for the introgressed DNA fragment carrying nematode resistance while simultaneously selecting for a recessive *AhFAD2B* allele necessary to recover lines with a high ratio of oleic:linoleic acid (O/L) (Chu et al., 2011). While homozygous recessive mutations in both *AhFAD2* homeologs are necessary to achieve high O/L, the frequency of a spontaneous loss-of-function allele of *AhFAD2A* already is high in ssp. *hypogaea* germplasm (Chu et al., 2007b) and fixed in most elite lines of US runner and Virginia market-type peanuts (Chu et al., 2009); therefore, breeding for high O/L may only require selection for the mutant allele of *AhFAD2B*. For disease resistance traits, the molecular markers associated with the QTL for leaf rust are being used to introgress leaf rust resistance in elite cultivars, namely ICGV 91114, JL 24 and TAG 24 through MABC at ICRISAT, India. At present, a total of 158 BC_2F_2 and 76 BC_3F_2 homozygous lines are available in the genetic background of all the above three parents. In the case of drought tolerance, the MABC approach is not the appropriate molecular breeding approach because there are many QTLs identified for drought tolerance and they contribute to small phenotypic variation (Gautami et al., in press; Ravi et al., 2011; Varshney et al., 2009c). In such cases, modern breeding approaches like marker-assisted recurrent selection (MARS) or genomic selection (GS) seems to be better approach (Bernardo, 2009; Bernardo and Charcosset, 2006; Bernardo and Yu, 2007; Charmet et al., 1999; Heffner et al., 2009; Hospital et al., 2000; Jannink et al., 2010; Ribaut and Ragot, 2007).

Apart from introgression of genes/QTLs linked to traits from the elite cultivars in the variety of interest, molecular markers are helpful for introgression the genes from wild species, which are generally inferior in agronomic performance, into elite cultivars (Fig. 3). In this context, molecular markers evenly distributed in the genome have been utilized for tracking genome recovery during backcrossing in several crops. While introgressing genes from wild relatives, stringent background selection is necessary to limit linkage drag by tracking the presence of unwanted genomic segments of the wild relatives. Using this feature along with limited available markers, genome-wide segment introgressions from a synthetic amphidiploid (*A. duranensis* × *A. ipaënsis*) were transferred into genetic background of a cultivated variety (Fleur 11) (Foncéka et al., 2009). Marker assisted selection of BC_1F_1 and then BC_2F_1 lines carrying the donor segments with the best possible return to the background of the cultivated variety provided a set of lines offering an optimal distribution of the synthetic genome introgressions. In summary, now it is possible to broaden the genetic base of the cultivated peanut by introgressing genomic segments from the wild species or synthetic amphidiploid genotypes with the help of molecular markers.

5. Summary and future prospects

One of the main challenges of the low level of genetic diversity has been partially tackled by developing several thousand molecular markers for peanut. However, it will still take some time before cost-effective SNP genotyping platforms are available for genotyping the tetraploid peanut germplasm collections or peanut mapping populations. Therefore, SSR markers are still considered the markers of choice for genetics and breeding applications in cultivated peanuts. Although some genetic maps have been developed for both diploid and tetraploid genome species, genetic mapping position is available only for few hundred SSR markers. One of the challenges to the peanut community is integrating as many markers as possible on genetic maps. Although a physical map has been developed for AA-genome species (http:/www.plantgenome.uga.edu), it is highly desirable to have the physical map for BB-genome also. As a result, it would be possible to integrate the genetic and physical maps that will facilitate gene cloning as well as molecular breeding in efficient manner. In terms of application of genomics to peanut breeding, though some efforts have been made for trait mapping and molecular breeding for resistance to some biotic stresses (e.g., leaf rust and root nematode). The other challenging areas which should be targeted in the coming years include production and quality constraints such as drought stress along with aflatoxin/mycotoxin contamination which has teratogenic and carcinogenic effect on humans and animals. Further, in the current scenario of yield stagnation, utilizing molecular breeding to increase oil content is another possible way to increase profitability of farmers. Improving nutritional quality needs attention of the peanut community through increase of oleic:linoleic acid (O/L) ratio which will increase shelf life of peanut oil and other peanut products along with providing health benefits by decreasing blood low density lipoprotein (LDL) levels, suppressing tumerogenesis and ameliorating inflammatory diseases. Due to ongoing advances in sequencing technologies and bioinformatics strategies, it is also anticipated that genome sequence or at the least gene space for peanut genome will soon be available. The genome/gene space sequence would provide the opportunities to link the phenotype with genes. To reach this goal, efforts should be devoted to high-throughput and precise phenotyping for different traits important to the peanut breeding community. In summary, the future of peanut genomics and use of molecular tools in breeding seems to be bright that will ensure the peanut improvement for different production as well as quality constraints.

Acknowledgments

Financial support received from National Fund of the Basic and Strategic Research in Agriculture (NFBSRA) of the Indian Council of Agriculture Research (ICAR), New Delhi, India and Generation Challenge Programme (http://www.generationcp.org) of CGIAR is gratefully acknowledged.

References

Alves DMT, Pereira RW, Leal-Bertioli SC, Moretzsohn MC, Guimarães PM, Bertioli DJ. Development and use of single nucleotide polymorphism markers for candidate resistance genes in wild peanuts (*Arachis* spp). Genet Mol Res 2008;7:631–42.

Basha SM. Identification of cultivar differences in seed polypeptide composition of peanuts (*Arachis hypogaea* L.) by two-dimensional gel electrophoresis. Plant Physiol 1979;63:301–6.

Bernardo R. Genome wide selection for rapid introgression of exotic germplasm in maize. Crop Sci 2009;49:419–25.

Bernardo R, Charcosset A. Usefulness of gene information in marker-assisted recurrent selection: a simulation appraisal. Crop Sci 2006;46:614–21.

Bernardo R, Yu J. Prospects for genome wide selection for quantitative traits in maize. Crop Sci 2007;47:1082–90.

Bi YP, Liu W, Xia H, Su L, Zhao CZ, Wan SB, et al. EST sequencing and gene expression profiling of cultivated peanut (*Arachis hypogaea* L.). Genome 2010;53:832–9.

Bravo JP, Hoshino AA, Angelici CMLCD, Lopes CR, Gimenes MA. Transferability and use of microsatellite markers for the genetic analysis of the germplasm of some *Arachis* section species of the genus *Arachis*. Genet Mol Biol 2006;29:516–24.

Brown AHD. Core collections: a practical approach to genetic resource management. Genome 1989;31:818–24.

Burow MD, Simpson CE, Paterson AH, Starr JL. Identification of peanut (*Arachis hypogaea* L.) RAPD markers diagnostic for root-knot nematode (*Meloidogyne arenaria* (Neal) Chitwood) resistance. Mol Breed 1996;2:369–79.

Burow MD, Simpson CE, Starr JL, Paterson AH. Transmission genetics of chromatin from a synthetic amphiploid in cultivated peanut (*A. hypogaea* L.): broadening the gene pool of a monophyletic polyploid species. Genetics 2001;159:823–37.

Charmet G, Robert N, Perretant MR, Gay G, Sourdille P, Groos C, et al. Marker-assisted recurrent selection for cumulating additive and interactive QTLs in recombinant inbred lines. Theor Appl Genet 1999;99:1143–8.

Chen X, Culbreath A, Brenneman T, Holbrook C, Guo B. Identification and cloning of TSWV resistance gene(s) in cultivated peanuts and development of markers for breeding selection. Phytopathology 2008a;98:S36.

Chen X, Brenneman T, Culbreath A, Holbrook C, Guo B. Characterization of two peanut *oxalate oxidase* genes and development of peanut cultivars resistant to stem rot (*Sclerotium rolfsii*). Phytopathology 2008b;98:S36.

Chen CY, Guo B, Qin H, Feng S, Dang P, Culbreadth A, et al. Morphological and molecular marker characterization of two RIL mapping populations in cultivated peanuts. proc: Plant and Animal Genome XIX Conference, 9–13th Jan; 2010.. San Diego, USA.

Chen X, Wang ML, Holbrook C, Culbreath A, Liang X, Brenneman T, et al. Identification and characterization of a multigene family encoding germin-like proteins in cultivated peanut (*Arachis hypogaea* L.). Plant Mol Biol Rep 2011;29:389–403.

Chu Y, Holbrook CC, Timper P, Ozias-Akins P. Development of a PCR-based molecular marker to select for nematode resistance in peanut. Crop Sci 2007a;47:841–5.

Chu Y, Ramos L, Holbrook CC, Ozias-Akins P. Frequency of a loss-of-function mutation in oleoyl-PC desaturase (*ahFAD2A*) in the minicore of the U.S. peanut germplasm collection. Crop Sci 2007b;47:2372–8.

Chu Y, Holbrook CC, Ozias-Akins P. Two alleles of *ahFAD2B* control the high oleic acid trait in cultivated peanut. Crop Sci 2009;49:2029–36.

Chu Y, Wu CL, Holbrook CC, Tillman BL, Person G, Ozias-Akins P. Marker-assisted selection to pyramid nematode resistance and the high oleic trait in peanut. The Plant Genome 2011;4:110–7.

Cuc LM, Mace ES, Crouch JH, Quang VD, Long TD, Varshney RK. Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea*). BMC Plant Biol 2008;8:55.

Dwivedi SL, Gurtu S, Chandra S, Weng YJ, Nigam SN. Assessment of genetic diversity among selected groundnut germplasm-I:RAPD analysis. Plant Breed 2001;120:345–9.

FAO. Food and Agricultural Organization of the United Nation, FAO Statistical Database. http://faostat.fao.org/faostat/collections?subset=agriculture 2009.

Ferguson ME, Burow MD, Schultze SR, Bramel PJ, Paterson AH, Kresovich S, et al. Microsatellite identification and characterization in peanut (*A. hypogaea* L.). Theor Appl Genet 2004;108:1064–70.

Foncéka D, Hodo-Abalo T, Rivallan R, Faye I, Sall MN, Ndoye O, et al. Genetic mapping of wild introgressions into cultivated peanut: a way toward enlarging the genetic basis of a recent allotetraploid. BMC Plant Biol 2009;9:103.

Froenicke L, Pandey M, Upadhyaya H, Moretzsohn MC, Guimarães P, Leal-Bertioli S, et al. Towards ultra-dense genetic maps of peanut generated by sequencing diploid and tetraploid RIL populations and a peanut diversity panel. proc: Advances in *Arachis* through Genomics and Biotechnology (AAGB), 13–15th June; 2011 Brasília, Brazil.

Garcia GM, Stalker HT, Kochert G. Introgression analysis of an interspecific hybrid population in peanuts (*Arachis hypogaea* L.) using RFLP and RAPD markers. Genome 1995;38:166–76.

Garcia GM, Stalker HT, Shroeder E, Kochert G. Identification of RAPD, SCAR and RFLP markers tightly linked to nematode resistance genes introgressed from *Arachis cardenasii* to *A. hypogaea*. Genome 1996;39:836–45.

Garcia GM, Stalker HT, Schroeder E, Lyerly JH, Kochert G. A RAPD-based linkage map of peanut based on a backcross population between the two diploid species *Arachis stenosperma* and *A. cardenasii*. Peanut Sci 2005;32:1–8.

Gautami B, Ravi K, Narasu ML, Hoisington DA, Varshney RK. Novel set of groundnut SSR markers for germplasm analysis and inter-specific transferability. Int J Integr Biol 2009;7:100–6.

Gautami B, Pandey MK, Vadez V, Nigam SN, Ratnakumar P, Krishnamurthy L, et al. Quantitative trait locus analysis and construction of consensus genetic map for drought tolerance traits based on three recombinant inbred line populations in cultivated groundnut (*Arachis hypogaea* L.). Mol Breed in press. doi:10.1007/s11032-011-9660-0.

Gimenes MA, Hosino AA, Barbosa AVG, Palmieri DA, Lopes CR. Characterization and transferability of microsatellite markers of cultivated peanut (*Arachis hypogaea*). BMC Plant Biol 2007;7:9.

Gobbi A, Teixeira C, Moretzsohn M, Guimarães P, Leal-Bertioli SC, Bertioli D, et al. Development of a linkage map to species of B genome related to the peanut (*Arachis hypogaea*-AABB). proc: Plant and Animal Genomes XIV Conference, 14–18th January; 2006.. San Diego, USA.

Guimarães PM, Garsmeur O, Proite K, Leal-Bertioli SC, Seijo G, Chaine C, et al. BAC libraries construction from the ancestral diploid genomes of the allotetraploid cultivated peanut. BMC Plant Biol 2008;8:14.

Guimarães PM, Brasileiro ACM, Araújo ACG, Leal-Bertioli SCM, da Silva FR, Morgante CV, et al. A study of gene expression in the nematode resistant wild peanut relative, *Arachis stenosperma*, in response to challenge with *Meloidogyne arenaria*. Trop Plant Biol 2010. doi:10.1007/s12042-010-9056-z.

Guimarães PM, Brasileiro ACM, Leal-Bertioli SCM, Pappas G, Togawa R, Bonfim O, et al. Comparative 454 pyrosequencing of transcripts from two wild *Arachis* genotypes under biotic and abiotic stress. proc: Plant and Animal Genome XIX Conference, 15–19th Jan; 2011.. San Diego, USA.

Guo BZ, Chen X, Dang P, Scully BT, Liang X, Holbrook CC, et al. Peanut gene expression profiling in developing seeds at different reproduction stages during *Aspergillus parasiticus* infection. BMC Dev Biol 2008;8:12.

Guo B, Chen X, Hong Y, Liang X, Dang P, Brenneman T, et al. Analysis of gene expression profiles in leaf tissues of cultivated peanuts and development of EST–SSR markers and gene discovery. Int J Plant Genomics 2009. doi:10.1155/2009/715605.

Guo Y, Nagy ED, Khanal S, Knapp SJ. Comparative mapping in intraspesific populations uncovers high degree of macrosynteny between A- and B-genome diploid species of peanut. proc: Plant and Animal Genome XIX Conference, 9–13th Jan; 2010.. San Diego, USA.

Guo BZ, Fedorova ND, Chen X, Wan CH, Wang W, Nierman WC, et al. Gene expression profiling and identification of resistance genes to *Aspergillus flavus* infection in peanut through EST and microarray startegies. Toxins 2011;3:737–53.

Gupta PK, Kumar J, Mir RR, Kumar A. Marker-assisted selection as a component of conventional plant breeding. Plant Breed Rev 2010;33:145–217.

Halward TM, Stalker HT, Kochert G. Development of an RFLP linkage map in diploid peanut species. Theor Appl Genet 1993;87:379–84.

He GH, Prakash CS. Evaluation of genetic relationships among botanical varieties of cultivated peanut (*Arachis hypogaea* L.) using AFLP markers. Genet Resour Crop Evol 2001;48:347–52.

He GH, Meng R, Newman M, Gao G, Pittman RN, Prakash CS. Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.). BMC Plant Biol 2003;3:3.

Heffner EL, Sorrells ME, Jannink JL. Genomic selection for crop improvement. Crop Sci 2009;49:1-12.

Herselman L. Genetic variation among Southern African cultivated peanut (*Arachis hypogaea* L.) genotypes as revealed by AFLP analysis. Euphytica 2003;133:319–27.

Herselman LR, Thwaites FM, Kimmins B, Courtois PJA, Merwe VD, Seal SE. Identification and mapping of AFLP markers linked to peanut (*Arachis hypogaea* L.) resistance to the aphid vector of groundnut rosette disease. Theor Appl Genet 2004;109:1426–33.

Hilu KW, Stalker HT. Genetic relationships between peanut and wild species of *Arachis* sect. *Arachis* (*Fabaceae*): evidence from RAPDs. Syst Evol 1995;198:167–78.

Holbrook CC, Dong W. Development and evaluation of a mini core collection for the U.S. peanut germplasm collection. Crop Sci 2005;45:1540–4.

Holbrook CC, Anderson WF, Pittman RN. Selection of a core collection from the United States germplasm collection of peanut. Crop Sci 1993;33:859–61.

Hong Y, Chen X, Liang X, Liu H, Zhou G, Li S, et al. A SSR-based composite genetic linkage map for the cultivated peanut (*Arachis hypogaea* L.) genome. BMC Plant Biol 2010a;10:17.

Hong YB, Chen XP, Liu HY, Zhou GY, Li SX, Wen SJ, et al. Development and utilization of orthologous SSR markers in *Arachis* through soybean (*Glycine max*) EST. Acta Agron Sin 2010b;36:410–21. in Chinese with English abstract.

Hopkins MS, Casa AM, Wang T, Mitchell SE, Dean RE, Kochert GD, et al. Discovery and characterization of polymorphic simple sequence repeats (SSRs) in peanut. Crop Sci 1999;39:1243–7.

Hospital F, Goldringer I, Openshaw S. Efficient marker-based recurrent selection for multiple quantitative trait loci. Genet Res 2000;75:1181–9.

Jain AK, Basha SM, Holbrook CC. Identification of drought responsive transcripts in peanut (*Arachis hypogaea* L.). Mol Biotechnol Genet 2001;4:2.

Jannink JL, Lorenz AJ, Iwata H. Genomic selection in plant breeding: from theory to practice. Brief Funct Genomics 2010;9:166–77.

Jesubatham AM, Burow MD. PeanutMap: an online genome database for comparative molecular maps of peanut. BMC Bioinformatics 2006;7:375.

Jiang HF, Ren XP, Liao BS, Huang JQ, Lei Y, Chen BY, et al. Peanut core collection established in china and compared with ICRISAT mini core collection. Acta Agron Sin 2008;34:25–30.

Jiang HF, Ren XP, Zhang XJ, Huang JQ, Lei Y, Yan LY, et al. Comparison of genetic diversity based on SSR markers between peanut mini core collections from China and ICRISAT. Acta Agron Sin 2010;36:1084–91.

Katam R, Basha SM, Vasanthaiah Naik KSS. Identification of drought tolerant genotypes employing proteomics approach. J SAT Agric Res 2007;5:1–4.
Katam R, Basha SM, Suravajhala P, Pechan T. Analysis of peanut leaf proteome. J Proteome Res 2010;9:2236–54.
Khedikar YP, Gowda MVC, Sarvamangala C, Patgar KV, Upadhyaya HD, Varshney RK. A QTL study on late leaf spot and rust revealed one major QTL for molecular breeding for rust resistance in groundnut (*Arachis hypogaea* L.). Theor Appl Genet 2010;121: 971–84.
Kilian A. DArT-based whole genome profiling and novel information technologies in support system of modern breeding of groundnut. proc: 3rd International Conference for Peanut Genomics and Biotechnology on "Advances in *Arachis* through Genomics and Biotechnology (AAGB)," 4–8th November; 2008.. Hyderabad, India.
Knoll JE, Ramos ML, Zeng Y, Holbrook CC, Chow M, Chen S, et al. TILLING for allergen reduction and improvement of quality traits in peanut (*Arachis hypogaea* L.). BMC Plant Biol 2011;11:81.
Kochert G, Stalker HT, Gimenes M, Galgaro L, Lopes CR, Moore K. RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). Am J Bot 1996;83:1282–91.
Koilkonda P, Sato S, Tabata S, Shirasawa K, Watanabe A, Wada T, et al. Large-scale development of expressed sequence tag-derived simple sequence repeat markers and diversity analysis in *Arachis* spp. Mol Breed 2011. doi:10.1007/s11032-011-9604-8.
Kottapalli KR, Payton P, Rakwal R, Agrawal GA, Shibato J, Burow M, et al. Proteomics analysis of mature seeds of four peanut cultivars using two-dimensional gel electrophoresis reveals distinct differential expression of storage, anti-nutritional and allergenic proteins. Plant Sci 2008;175:321–9.
Krapovickas A, Gregory WC. Taxonomia del género *Arachis* (Leguminosae). Bonplandia 1994;8:1-186.
Leal-Bertioli SCM, José AC, Alves-Freitas DM, Moretzsohn MC, Guimarães PM, Nielen S, et al. Identification of candidate genome regions controlling disease resistance in *Arachis*. BMC Plant Biol 2009;9:112.
Liang X, Zhou G, Hong Y, Chen X, Liu H, Li S. Overview of research progress on peanut (*Arachis hypogaea* L.) host resistance to aflatoxin contamination and genomics at the Guangdong Academy of Agricultural Sciences. Peanut Sci 2009a;36: 29–34.
Liang X, Chen X, Hong Y, Liu H, Zhou G, Li S, et al. Utility of EST-derived SSR in cultivated peanut (*Arachis hypogaea* L.) and *Arachis* wild species. BMC Plant Biol 2009b;9: 35.
Luo M, Dang P, Guo B, He G, Holbrook CC, Bausher MG, et al. Generation of expressed sequence tags (ESTs) for gene discovery and marker development in cultivated peanut. Crop Sci 2005a;45:346–53.
Luo M, Liang XQ, Dang P. Microarray-based screening of differentially expressed genes in peanut in response to *Aspergillus parasiticus* infection and drought stress. Plant Sci 2005b;169:695–703.
Mace ES, Varshney RK, Mahalakshmi V, Seetha K, Gafoor A, Leeladevi Y, et al. *In silico* development of simple sequence repeat markers within the aeschynomenoid/dalbergoid and genistoid clades of the leguminosae family and their transferability to *Arachis hypogaea*, groundnut. Plant Sci 2007;174:51–60.
McCallum C, Comai L, Greene E, Henikoff S. Targeting induced local lesions in genomes (tilling) for plant functional genomics. Plant Physiol 2000;123:439–42.
Milla SR. Relationships and utilization of *Arachis* germplasm in peanut improvement. Ph. D. Thesis, North Carolina State University, USA. 2003;1–150.
Moretzsohn MC, Hopkins MS, Mitchell SE, Kresovich S, Valls JFM, Ferreira ME. Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis of hyper variable regions of the genome. BMC Plant Biol 2004;4:11.
Moretzsohn MC, Leoi L, Proite K, Guimarães PM, Leal-Bertioli SCM, Gimenes MA, et al. A microsatellite-based, gene-rich linkage map for the AA genome of *Arachis* (*Fabaceae*). Theor Appl Genet 2005;111:1060–71.
Moretzsohn MC, Barbosa AVG, Alves DMT, Teixeira C, Leal-Bertioli SCM, Guimarães PM, et al. A linkage map for the B-genome of *Arachis* (Fabaceae) and its synteny to the A-genome. MBC Plant Biol 2009;9:40.
Nagy E, Guo S, Khanal S, Taylor C, Ozias-Akins P, Stalker HT, et al. Developing a high density molecular map of the A-genome species *A. duranensis*. proc: American Peanut Research and Education Society (APRES), 12–15th July; 2010a.. Florida, USA.
Nagy E, Chu Y, Guo Y, Khanal S, Tang S, Li Y, et al. Recombination is suppressed in an alien introgression in peanut harboring *Rma*, a dominant root-knot nematode resistance gene. Mol Breed 2010b;26:357–70.
Nelson MN, Phan HTT, Ellwood SR, Paula MM, Hane J, Williams A, et al. The first gene-based map of *Lupinus angustifolius* L. — location of domestication genes and conserved synteny with *Medicago truncatula*. Theor Appl Genet 2006;113: 225–38.
Nielen S, Campos-Fonseca F, Leal-Bertioli S, Guimarães PM, Seijo G, Town C, et al. FIDEL —a retrovirus-like retrotransposon and its distinct evolutionary histories in the A- and B-genome components of cultivated peanut. Chrom Res 2010;18:227–46.
Palmieri DA, Hoshino AA, Bravo JP, Lopes CR, Gimenes MA. Isolation and characterization of microsatellite loci from the forage species *Arachis pintoi* (Genus *Arachis*). Mol Ecol Notes 2002;2:551–3.
Palmieri DA, Bechara MD, Curi RA, Gimenes MA, Lopes CR. Novel polymorphic microsatellite markers in section Caulorrhizae (*Arachis,* Fabaceae). Mol Ecol Notes 2005;5: 77–9.
Pandey MK, Gautami B, Jayakumar T, Sriswathi M, Upadhyaya HD, Gowda MVC, et al. Highly informative genic and genomic SSR markers to facilitate molecular breeding in cultivated groundnut (*Arachis hypogaea*). Plant Breed 2011. doi:10.1111/j.1439-0523.2011.01911.x.
Payton P, Kottapalli K, Rowland D, Faircloth W, Guo B, Burow M, et al. Gene expression profiling in peanut using high density oligonucleotide microarrays. BMC Genomics 2009;10:265.
Proite K, Leal-Bertioli SC, Bertioli DJ, Moretzsohn MC, da Silva FR, Martins NF, et al. ESTs from a wild *Arachis* species for gene discovery and marker development. BMC Plant Biol 2007;7:7.
Ravi K, Vadez V, Isobe S, Mir RR, Guo Y, Nigam SN, et al. Identification of several small-effect main QTLs and larg number of epistatic QTLs for drought tolerance in groundnut (*Arachis hypogaea* L.). Theor Appl Genet 2011;122:1119–32.
Ribaut JM, Hoisington D. Marker-assisted selection: new tools and strategies. Trends Plant Sci 1998;3:236–9.
Ribaut1 JM, Ragot M. Marker-assisted selection to improve drought adaptation in maize: the backcross approach, perspectives, limitations, and alternatives. J Exp Bot 2007;58: 351–60.
Sarvamangala C, Gowda MVC, Varshney RK. Identification of quantitative trait loci for protein content, oil content and oil quality for groundnut (*Arachis hypogaea* L.). Field Crops Res 2011;122:49–59.
Seijo JG, Lavia GI, Fernández A, Krapovickas A, Ducasse D, Moscone EA. Physical mapping of 5S and 18S–25S rRNA genes evidences that *Arachis duranensis* and *A. ipäensis* are the wild diploid species involved in the origin of *A. hypogaea* (Leguminosae). Am J Bot 2004;91:1294–303.
Selvaraj MG, Narayana M, Schubert AM, Ayers JL, Baring MR, Burow MD. Identification of QTLs for pod and kernel traits in cultivated peanut by bulked segregant analysis. Electron J Biotech 2009;12. doi:10.2225/vol12.
Simpson CE. Use of wild *Arachis* species/introgression of genes into *A. hypogaea* L. Peanut Sci 2001;28:114–6.
Singh KP, Raina SN, Singh AK. Variation in chromosomal DNA associated with the evolution of *Arachis* species. Genome 1996;39:890–7.
Song GQ, Li MJ, Xiao H, Wang XJ, Tang RH, Xia H, et al. EST sequencing and SSR marker development from cultivated peanut (*Arachis hypogaea* L.). Electron J Biotechnol 2010;13. doi:10.2225/vol13.
Stalker HT. A new species in section *Arachis* of peanuts with a D genome. Am J Bot 1991;78:630–7.
Stalker HT, Simpson CE. Genetic resources in *Arachis*. In: Pattee HE, Stalker HT, editors. Advances in peanut scienceStillwater, OK: Am Peanut Res Educ Soc; 1995. p. 14–53.
Stalker HT, Dhesi JS, Parry D, Hahn JH. Cytological and interfertility relationships of *Arachis* section *Arachis*. Am J Bot 1991;78:238–46.
Subramanian V, Gurtu S, Nageswara Rao RC, Nigam SN. Identification of DNA polymorphism in cultivated groundnut using random amplified polymorphic DNA (RAPD) assay. Genome 2000;43:656–60.
Sujay V, Gowda MVC, Pandey MK, Bhat RS, Khedikar YP, Nadaf HL, et al. Quantitative trait locus analysis and construction of consensus genetic map for foliar disease resistance based on two recombinant inbred line populations in cultivated groundnut (*Arachis hypogaea* L.). Mol Breed in press. doi:10.1007/s11032-011-9661-z.
Upadhyaya HD, Bramel PJ, Ortiz R, Singh S. Developing a mini core of peanut for utilization of genetic resources. Crop Sci 2002;42:2150–6.
Upadhyaya HD, Ortiz R, Bramel PJ, Singh S. Development of a groundnut core collection using taxonomical, geographical and morphological descriptors. Genet Resour Crop Evol 2003;50:139–48.
Valls JFM, Simpson CE. Taxonomy, natural distribution, and attributes of *Arachis*. In: Kerridge PC, Hardy B, editors. Biology and agronomy of forage *Arachis*. Cali: CIAT; 1994. p. 1-18.
Varshney RK. Gene-based marker systems in plants: high throughput approaches for discovery and genotyping. In: Jain SM, Brar DS, editors. Molecular techniques in crop improvement, vol 2. The Netherlands: Springer; 2010. p. 119–42.
Varshney RK, Graner A, Sorrells ME. Genomics-assisted breeding for crop improvement. Trends Plant Sci 2005a;10:621–30.
Varshney RK, Graner A, Sorrells ME. Genic microsatellite markers in plants: features and applications. Trends Biotechnol 2005b;23:48–55.
Varshney RK, Hoisington DA, Tyagi AK. Advances in cereal genomics and applications in crop breeding. Trends Biotechnol 2006;24:490–9.
Varshney RK, Close TJ, Singh NK, Hoisington DA, Cook DR. Orphan legume crops enter the genomics era. Curr Opin Plant Biol 2009a;12:202–10.
Varshney RK, Hoisington DA, Nayak SN, Graner A. Molecular plant breeding: methodology and achievements. In: Somers D, Langridge P, Gustafson PJ, editors. Methods in molecular biology: plant genomics. Totowa, NJ, USA: The Humana Press; 2009b. p. 283–304.
Varshney RK, Bertioli DJ, Moretzsohn MC, Vadez V, Krishnamurthy L, Aruna R, et al. The first SSR-based genetic linkage map for cultivated groundnut (*Arachis hypogaea* L.). Theor Appl Genet 2009c;118:729–39.
Varshney RK, Nayak SN, May GD, Jackson SA. Next-generation sequencing technologies and their implications for crop genetics and breeding. Trends Biotechnol 2009d;27: 522–30.
Varshney RK, Gowda MVC, Radhakrishnan T, Pandey MK, Gautami B, Sujay V, et al. Development and application of genomic resources for molecular breeding in groundnut (*Arachis hypogaea L*). Proc: The 3rd International Conference on Plant Molecular Breeding (ICPMB), 5–9th September; 2010a. Beijing, China.
Varshney RK, Thudi M, May GD, Jackson SA. Legume genomics and breeding. Plant Breed Rev 2010b;33:257–304.
Varshney RK, Glaszmann JC, Leung H, Ribaut JM. More genomic resources for less-studied crops. Trends Biotechnol 2010c;28:452–60.
Wang CT, Yang XD, Chen DX, Yu SL, Liu GZ, Tang YY, et al. Isolation of simple sequence repeats from groundnut. Electron J Biotechnol 2007;10(3).
Wang CT, Yu SL, Zhang SW, Wang XZ, Tang YY, Zhang JC, et al. Novel protocol to identify true hybrids in normal oleate × high oleate crosses in peanut. Electron J Biotechnol 2010;13. doi:10.2225/vol13-issue5-fulltext-18.
Xu Y, Crouch JH. Marker-assisted selection in plant breeding: from publications to practice. Crop Sci 2008;48:391–407.
Yan YS, Lin XD, Zhang YS, Wang L, Wu K, Huang SZ. Isolation of peanut genes encoding *arachins* and *conglutins* by expressed sequence tags. Plant Sci 2005;169:439–45.

Yang X, Guo Y, Yan J, Zhang J, Song T, Rocheford T, et al. Major and minor QTL and epistasis contribute to fatty acid compositions and oil concentration in high-oil maize. Theor Appl Genet 2009;120:665–78.

Young ND, Weeden NF, Kochert G. Genome mapping in legumes (Family *Fabaceae*). In: Paterson AH, editor. Genome mapping in plants. Texas: Landes Biomedical Press; 1996. p. 211–77.

Yuan M, Gong L, Meng R, Li S, Dang P, Guo B, et al. Development of trinucleotide $(GGC)_n$ SSR markers in peanut (*Arachis hypogaea* L.). Electron J Biotechnol 2010. doi: 10.2225/vol13-issue6-fulltext-6.

Yüksel B, Paterson HA. Construction and characterization of peanut *HindIII* BAC library. Theor Appl Genet 2005;111:630–9.