

The first SSR-based genetic linkage map for cultivated groundnut (*Arachis hypogaea* L.)

R. K. Varshney · D. J. Bertioli · M. C. Moretzsohn · V. Vadez ·
L. Krishnamurthy · R. Aruna · S. N. Nigam · B. J. Moss · K. Seetha ·
K. Ravi · G. He · S. J. Knapp · D. A. Hoisington

Received: 25 July 2008 / Accepted: 4 November 2008 / Published online: 2 December 2008
© The Author(s) 2008. This article is published with open access at Springerlink.com

Abstract Molecular markers and genetic linkage maps are pre-requisites for molecular breeding in any crop species. In case of peanut or groundnut (*Arachis hypogaea* L.), an amphidiploid (4X) species, not a single genetic map is, however, available based on a mapping population derived from cultivated genotypes. In order to develop a genetic linkage map for tetraploid cultivated groundnut, a total of 1,145 microsatellite or simple sequence repeat (SSR) markers available in public domain as well as unpublished markers from several sources were screened on two genotypes, TAG 24 and ICGV 86031 that are parents of a recombinant inbred line mapping population. As a result,

144 (12.6%) polymorphic markers were identified and these amplified a total of 150 loci. A total of 135 SSR loci could be mapped into 22 linkage groups (LGs). While six LGs had only two SSR loci, the other LGs contained 3 (LG_AhXV) to 15 (LG_AhVIII) loci. As the mapping population used for developing the genetic map segregates for drought tolerance traits, phenotyping data obtained for transpiration, transpiration efficiency, specific leaf area and SPAD chlorophyll meter reading (SCMR) for 2 years were analyzed together with genotyping data. Although, 2–5 QTLs for each trait mentioned above were identified, the phenotypic variation explained by these QTLs was in the range of 3.5–14.1%. In addition, alignment of two linkage groups (LGs) (LG_AhIII and LG_AhVI) of the developed genetic map was shown with available genetic maps of AA diploid genome of groundnut and *Lotus* and *Medicago*. The present study reports the construction of the first genetic map for cultivated groundnut and demonstrates its utility for molecular mapping of QTLs controlling drought tolerance related traits as well as establishing relationships with diploid AA genome of groundnut and model legume genome species. Therefore, the map should be useful for the community for a variety of applications.

Communicated by T. Lübberstedt.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-008-0933-x) contains supplementary material, which is available to authorized users.

R. K. Varshney (✉) · V. Vadez · L. Krishnamurthy ·
R. Aruna · S. N. Nigam · B. J. Moss · K. Seetha · K. Ravi ·
D. A. Hoisington
International Crops Research Institute for the Semi-Arid Tropics
(ICRISAT), Patancheru, Greater Hyderabad 502 324, India
e-mail: r.k.varshney@cgiar.org

D. J. Bertioli
Universidade Católica de Brasília (UCB), Campus II,
SGAN 916, Brasília, DF CEP 70.790-160, Brazil

M. C. Moretzsohn
Embrapa Recursos Genéticos e Biotecnologia, C.P. 02372,
Brasília, DF CEP 70.770-900, Brazil

G. He
Tuskegee University, Tuskegee, AL 36088, USA

S. J. Knapp
The University of Georgia, Athens, GA 30602, USA

Introduction

Groundnut or peanut (*Arachis hypogaea* L.) is an important food and cash crop for resource-poor farmers in Asia and Africa. It is primarily grown for edible oil (48–50%) as well as for direct consumption by people. In addition, groundnut haulms and groundnut cake (after oil extraction) are excellent animal feed. For the subsistence farmers, groundnut contributes significantly to household food

security and cash income through the sale of groundnut products. Groundnut productivity in Western and Central Africa (WCA) and Eastern and Southern Africa (ESA) is below the world average yield of 1.55 tons/ha. Although groundnut productivity in Asia (1.8 tons/ha) exceeds the world average, it is still lower than the yields in developed countries (3 tons/ha). One of the main reasons for low productivity of this crop in these regions is the exposure of the crop to severe abiotic and biotic stresses. For instance, groundnut producing regions in WCA, ESA and Asia represent typically the semi-arid tropics (SAT) environment which is characterized by short and erratic rainfall and then long periods with virtually no rain. Water deficit is one of the most severe stresses that threaten sustainable crop production in SAT regions as the yield losses each year due to drought alone are estimated to be around US\$520 million (Johansen and Nigam 1994).

Water capture by roots and water-use efficiency are two major components of the yield architecture, as defined by Passioura (1977), that are important for crops growing under water-limited environments. Water use efficiency can be considered as a drought avoidance trait, which deals with using soil water more efficiently for biomass production, therefore to “avoid” drought. Drought avoidance is considered to be the major trait of interest for expanding production to presently uncropped areas and the post-rainy fallows in SAT regions. Crop productivity per unit of water has become an important consideration in breeding programs dealing with drought. Higher water use efficiency or transpiration efficiency (TE) is therefore a major component for improving yield under water deficit. Several groundnut genotypes with higher transpiration efficiency (TE, in g of biomass per kg of water transpired) have been identified at ICRISAT. A recombinant inbred line (RIL) mapping population has been developed by crossing ICGV 86031 and TAG 24 (respectively high and low TE under the conditions in which they were tested) that segregates for TE as well as several of its surrogate traits such as specific leaf area (SLA) and SPAD chlorophyll meter reading (SCMR).

Groundnut breeders and physiologists have been working across the world to improve the yield of the crop under water deficit conditions but the complexity of the drought issue and the difficulty to accurately measure plant response to drought requires some modern methods to unmistakably identify the genotypes having superior performance under stress conditions. Recent advances in the area of crop genomics have offered tools to assist breeding (Varshney et al. 2005, 2006). Molecular markers and molecular genetic linkage maps are the pre-requisites for undertaking molecular breeding activities in any crop. Such tools would then simply speed up the process of introgression of beneficial traits into preferred varieties, especially for complex traits such as drought. However, for

groundnut, although several hundreds of microsatellite markers have been developed (see Varshney et al. 2007), no molecular genetic map based on a cultivated \times cultivated cross has been published to date. The main reasons for this is the low level of genetic diversity present in cultivated germplasm, at least the level which can not be detected with the detection tools that are currently available. A genetic map based on a cross of a synthetic amphidiploid (TxAG-6) and a US variety (Florunner) was developed earlier using restriction fragment length polymorphism (RFLP) loci (Burow et al. 2001). However, RFLP is labor intensive and not very suitable for use in breeding programs. Therefore, several research groups have developed microsatellite or simple sequence repeat (SSR) markers (see Varshney et al. 2007) but to date, these SSR markers have only been integrated into a diploid *Arachis* AA genome map (Moretzsohn et al. 2005). This map is based on a cross of the most probable AA genome donor to cultivated groundnut *A. duranensis* (Kochert et al. 1996; Seijo et al. 2007) with a closely related species. The aim of developing this map was to provide a reference map based on a highly polymorphic population. This high polymorphism means that a very high percentage of candidate DNA markers are informative, thus permitting their inter leveraging on an integrated map with cultivated \times cultivated, cultivated \times synthetic amphidiploid and even other legumes (see below). With this in mind, since its publication, the AA genome map has been enriched, at UCB/Embrapa (Brazil), with other markers, including candidate genes and Universal Legume Anchor Markers (Leg markers). Leg markers are based on PCR primers that bind conserved sequences flanking introns in legume homologues of genes present in only a single copy in the *Arabidopsis* genome. As such they work in a wide range of legume species and allow the alignment and integration of different genetic maps (Hougaard et al. 2008; unpublished data).

The present study was initiated to develop a molecular genetic map of groundnut based on the cultivated \times cultivated mapping population and SSR markers. Furthermore the application of this genetic map was demonstrated for mapping WUE and related surrogate traits in groundnut. In addition, it was possible to align some of the LGs of this map with the reference AA genome map and consequently with the genome sequences of the model legumes.

Material and methods

Plant material and DNA isolation

A RIL mapping population comprising of 318 F_8/F_9 lines, developed from a cross between ICGV 86031 \times TAG 24

was used. DNA was extracted from the parental genotypes and the RILs according to a modified CTAB-based procedure, as described in Cuc et al. (2008).

Phenotyping

The complete set of 318 F₈/F₉ lines was used for phenotyping for the following drought related traits for two consecutive years, 2004 and 2005: (i) transpiration (T), (ii) TE, (iii) SLA, and (iv) SCMR. Methodology for measuring the above mentioned traits are given in a separate study (Krishnamurthy et al. 2007).

Marker polymorphism and analysis

SSR markers from different sources, given in Table 1, were used to screen the polymorphism between ICGV 86031 and TAG 24 and subsequently identified polymorphic markers were used to genotype all 318 RILs.

PCR reactions for all SSR markers were performed in 10 µl reaction volume in an ABI 9700 thermal cycler (Applied Biosystems, USA), in 384-well PCR plates (Applied Biosystems, USA), consisting of 2 pmoles of primer, 1.5 mM MgCl₂, 2 mM dNTPs, 0.1U of *Taq* DNA polymerase (Qiagen, Germany) and 1X PCR buffer (Qiagen, Germany). A touch down PCR amplification profile with 3 min of initial denaturation cycle, followed by first five cycles of 94°C for 20 s, 60°C for 20 s and 72°C

for 30 s, with 1°C decrease in annealing temperature per cycle, then 30 cycles of 94°C for 20 s with constant annealing temperature (56°C) and 72°C for 30 s followed by a final extension for 20 min at 72°C. The amplified products were tested on 1.2% agarose gel to check for the amplification of the PCR products.

Amplified products for majority of SSR markers were separated by electrophoresis on 6% polyacrylamide gels and visualized through silver staining (Tegelstrom 1992). In some cases, where resolving polymorphism was difficult, the PCR was done using the forward primer labeled with one of four fluorescence dyes, 6-FAM, VIC, NED, or PET (Applied Biosystems, USA). Such PCR amplicons were size fractionated using capillary electrophoresis on an ABI 3700 automatic DNA sequencer (Applied Biosystems, USA). Allele sizing of the electrophoretic data thus obtained was done using Genescan 3.1 software (Applied Biosystems, USA) and Genotyper 3.1 (Applied Biosystems, USA).

Map construction

Genotyping for identified polymorphic markers was done on 318 F₈ RILs. Marker segregation was subjected to the χ^2 test to examine distortion from the expected 1:1 segregation. Linkage analysis was performed using Mapmaker Macintosh version 2.0 (Lander et al. 1987). LGs were established using a minimum LOD score of 6.0 and a

Table 1 Summary on marker polymorphism between ICGV 86031 and TAG 24

Source	Marker name	Number of markers screened	Number of polymorphic markers	Number of amplified loci	Percent polymorphism
ICRISAT					
Ferguson et al. (2004)	pPGPseq, pPGSseq	226	39	40	17.3
Mace et al. (2007)	Chaet, Dal, Lup, Stylo, Ades, Amor	51	0	0	0
Cuc et al. (2008)	IPAHM	104	17	18	16.3
Embrapa/Catholic Univ					
Moretzsohn et al. (2004, 2005)	Ah, gi, RN, ML, RI, TC, AC	338	37	39	10.9
Proite et al. (2007)	RM, RN	53	4	4	7.5
Tuskegee Uni					
He et al. (2003), unpublished	PM	59	17	18	28.8
Uni Georgia					
Unpublished (S J Knapp)	GM	153	24	25	15.7
Miscellaneous					
Hopkins et al. (1999)	Ah	26	1	1	3.8
Palmieri et al. (2002, 2005)	AP	18	1	1	5.6
Nelson et al. (2006) (COS markers)	LG, Lup	103	2	2	1.9
Gimenes et al. (2007)	Ah, Ag	14	2	2	14.3
Total		1,145	144	150	12.6

maximum recombination fraction (θ) of 0.35. The most likely marker order within each LG was estimated by comparing the log-likelihood of the possible orders using multipoint analysis (“compare” command) or by the matrix correlation method using the “first order” command, for groups containing more than six markers. The LOD score was then decreased to 3.0 in order to include new markers in the groups, by two-point analysis (“group” command). The exact position of the newly included markers within each group was determined by using the “try” command, which compares the maximum-likelihood of each marker order after placing the markers, one by one, into every interval of the established order. The new marker orders were confirmed by permuting all adjacent triple orders (“ripple” command). Recombination fractions were converted into map distances in centimorgans (cM) using the Kosambi’s mapping function.

Quantitative trait locus (QTL) analysis

Genotyping data and phenotyping data obtained for T, TE, SLA and SCMR were analyzed for mapping QTLs by using the method composite interval mapping (CIM), proposed by Zeng (1993, 1994) in the WinQTL Cartographer, version 2.5 (Wang et al. 2007). CIM analysis was performed using the Model 6, scanning the genetic map and estimating the likelihood of a QTL and its corresponding effects at every 1 cM, while using significant marker cofactors to adjust the phenotypic effects associated with other positions in the genetic map. The number of marker cofactors for the background control was set by forward–backward stepwise regression. A window size of 10 cM was used, and therefore cofactors within 10 cM on either side of the QTL test site were not included in the QTL model. Thresholds were determined by permutation tests (Churchill and Doerge 1994; Doerge and Churchill 1996), using 1,000 permutations and a significance level of 0.05. The significant QTLs were plotted in graphics. Graphic presentation of the LGs and the QTLs was obtained by using MapChart, version 2.1 (Voorrips 2002).

Comparative mapping

Some SSR markers that were mapped onto the genetic map in this study and in the diploid AA genome map (Moretzsohn et al. 2005) were the same. Selected markers mapped in this study that had not been screened earlier for polymorphisms in the AA genome parentals were screened and, wherever possible, genotyped and mapped onto AA genetic map using the same methodology as described above.

Legume anchor markers were developed essentially as described by Fredslund et al. (2006a, b). Key features of the marker development were:

1. Identification of ESTs from multiple legume species, usually *Lotus*, soya and *Medicago*, with single strong BLAST hits against all predicted *Arabidopsis* proteins and the alignment of these ESTs.
2. Alignment of ESTs to a corresponding genomic region from *Lotus* or *Medicago* and inference of intron positions.
3. Identification of conserved intron-flanking sequences, and design of primers to bind these conserved sequences.

The rationale behind this procedure is that:

- (1) Markers to unique sequences within a genome facilitate the comparison of genetic maps, and genes that are single copy in *Arabidopsis* have a high probability of being single copy in legume genomes.
- (2) Introns are more variable than coding regions, and therefore they are better for marker development.
- (3) Primers that bind to sequences that are conserved are more likely to be transferable to other species.

The primers were used in PCR with the progenitors of the *Arachis* mapping population. Polymorphisms were identified by size- or sequence variation. In the latter case, most markers developed were cleaved amplified polymorphic sequences (CAPS) or dCAPS (Neff et al. 2002; Hougaard et al. 2008; unpublished data).

The methodology for determining synteny of the AA genome map with *Lotus* and *Medicago* will be described in detail elsewhere. Briefly, all legume anchor markers (Leg markers; Fredslund et al. 2005, 2006a, b) and most other markers mapped in the AA genome were sequence characterized. These sequences were used in BLAST as queries against the *Lotus* database from Kazusa DNA Research Institute (Japan), and against the pseudomolecules of *Medicago* using CViT blast (Chromosome Visualization Tool, http://www.medicago.org/genome/cvit_blast.php). For *Lotus*, genetic positions were available for most transformation-competent artificial chromosome (TAC)/bacterial artificial chromosome (BAC) clones (Sato et al. 2008), but where necessary, TAC/BACs were sequenced and microsatellite markers were developed for genotyping and mapping in Gifu × MG-20 and/or in *L. filicaulis* × *L. japonicus* Gifu (Sandal et al. 2006). All map positions are given with respect to the former map.

Results and discussion

Marker polymorphism

A total of 1,145 SSR markers, available in public domain as well as unpublished markers were screened on ICGV

86031 and TAG 24 (Table 1), and 144 markers showed polymorphism between these genotypes. The very low level of polymorphism (12.6%) observed in the present study is not unexpected, as similar levels of polymorphism has been observed in several other studies (see Varshney et al. 2007). Low level of genetic polymorphism in cultivated groundnut has been attributed to its origin from a single polyploidization event that occurred relatively recently on an evolutionary time scale (Young et al. 1996). However, additional contributing factors to the low levels of molecular polymorphism observed to date could be due to the marker techniques used. This emphasizes the urgent need to develop a critical mass of highly polymorphic molecular markers in groundnut. Indeed, development of SSR markers, from longer SSR-enriched libraries, BAC-end sequences, and SNP (single nucleotide polymorphism) markers using next generation sequencing technologies is underway in several laboratories including the Embrapa Recursos Genéticos e Biotecnologia/Universidade Católica de Brasília (UCB) (Brazil), University of Georgia (USA) and University of California-Davis (USA).

All identified 144 polymorphic markers were used for genotyping 318 RILs of the mapping population. While genotyping the mapping population, segregation data were scored for one locus for 139 markers, for two loci for four markers (pPGSseq9H8, IPAHM108, PM733, GM635) and for three loci for one marker (TC3G01). As a result, segregation data were obtained for a total of 150 SSR loci. Amplification of more than one fragment by primer pair/marker in groundnut has been reported in earlier studies (Hopkins et al. 1999; Krishna et al. 2004; Kottapalli et al. 2007). In these studies observation of more than one fragment per marker has been attributed to either amplification of duplicated loci or different loci, because of the tetraploid genome.

Genetic map for cultivated groundnut

Genotyping data obtained for all 150 loci were checked for segregation ratio using χ^2 test. A total 93 loci showed the expected 1:1 segregation ratio ($P < 0.05$) and were initially used to establish the LGs. Using a minimum LOD score of 6.0 and a maximum recombination fraction (θ) of 0.35, 84 marker loci were mapped into 20 LGs. The LOD score was then decreased to 3.0 in order to include other SSR loci, (basically markers that showed segregation distortion), by two-point analysis. As a result, additional 51 SSR loci could be integrated and two new LGs were formed. The LG_AhXI, composed of five distorted loci out of six mapped, and the LG_AhXVII composed by two SSR loci, being one distorted. Thus, in total, 135 loci were integrated into a total of 22 LGs, covering 1,270.5 cM of total map distance.

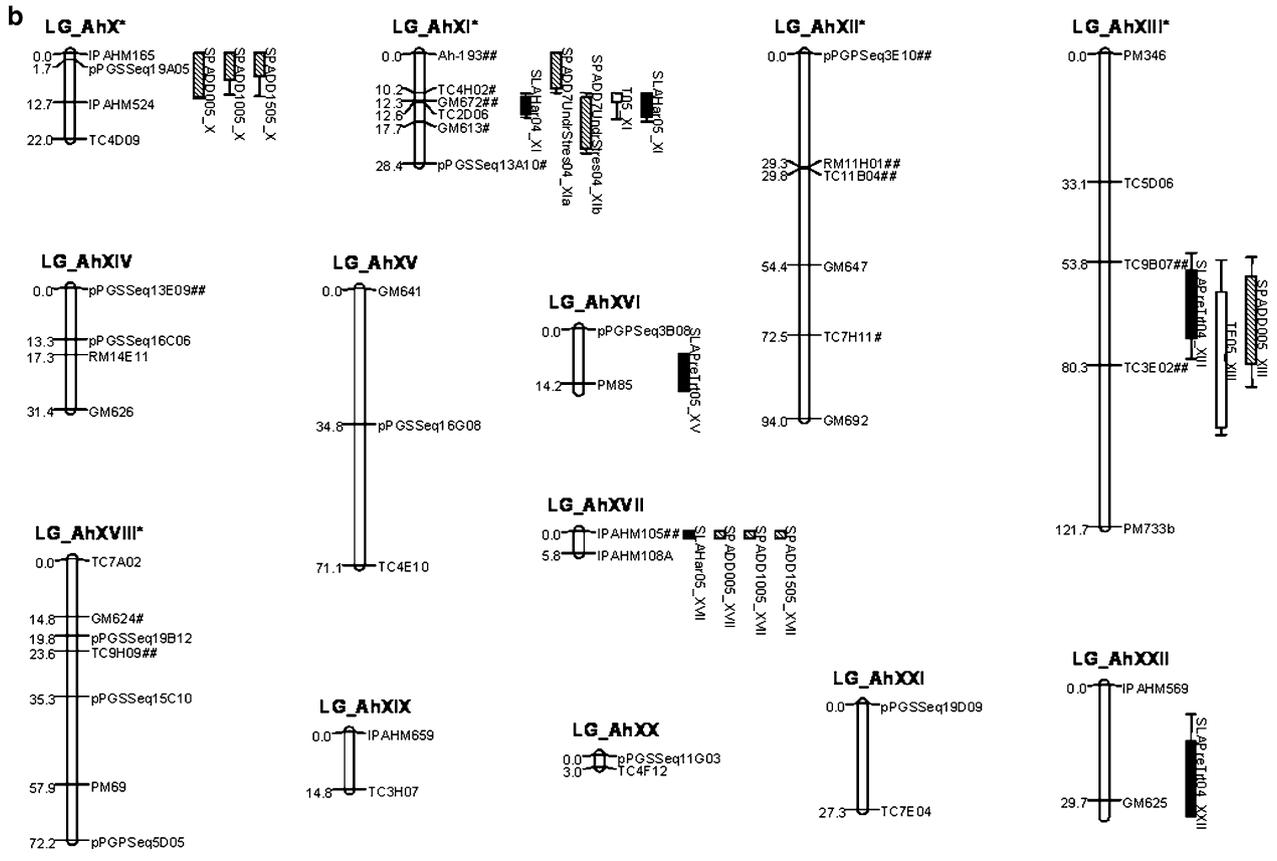
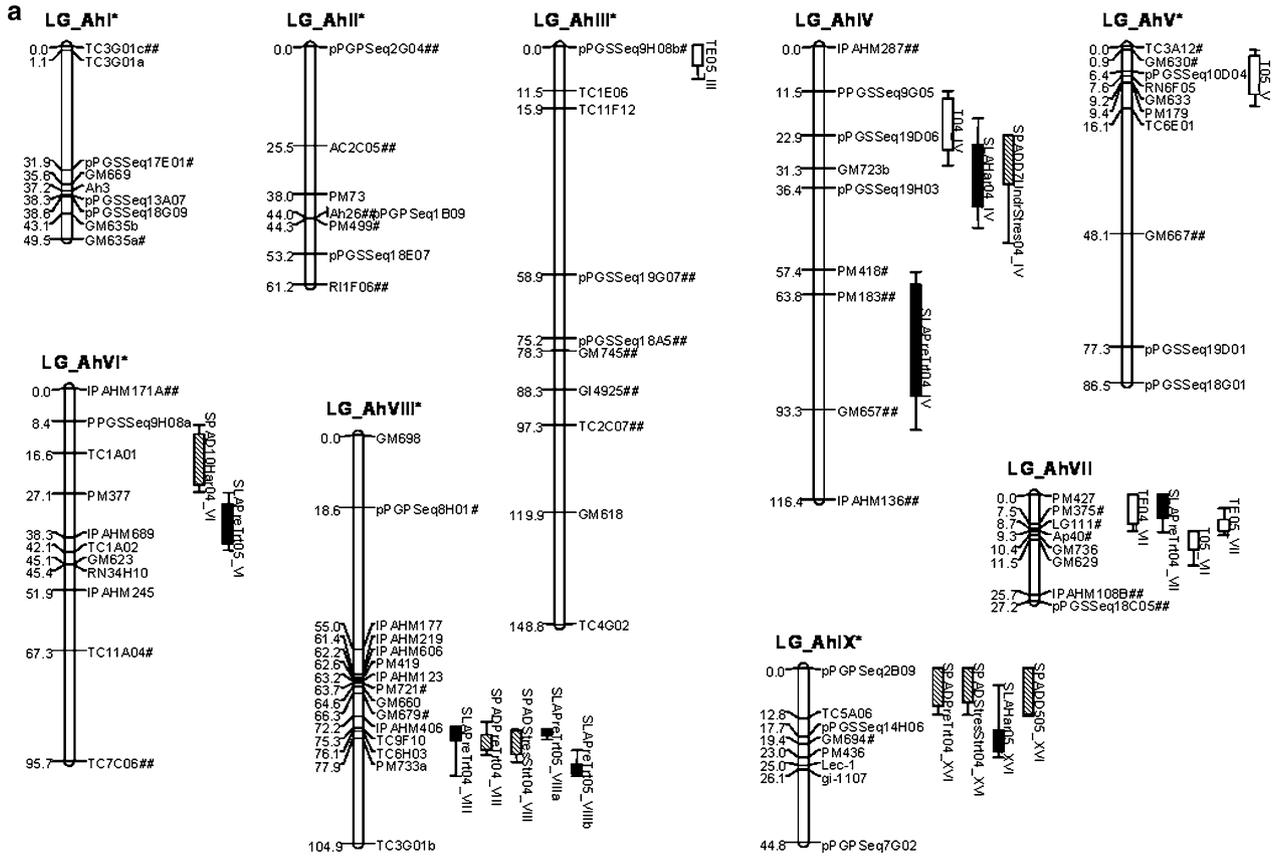
LGs were numbered according to the LG numbers of the AA genome map (Moretzsohn et al. 2005) by the identification of syntenic markers. These groups are identified with an asterisk in Fig. 1. The LGs with no common markers were numbered sequentially according to the number of mapped markers from higher to lower. Syntenic markers allowed the identification of four putative homeologous LGs (AhI/AhXI, AhII/AhXII, AhIII/AhXIII and AhVIII/AhXVIII). The map has a total coverage of 1,270.5 cM with an average distance of 9.4 cM between markers. While six LGs (AhXVI, AhXVII, AhXIX, AhXX, AhXXI, AhXXII) had only two marker loci, the other LGs contained 3 (AhXVI) to 15 (AhVIII). Marker loci were given the same name as primer name used by source laboratory, given in Table 1. Distorted markers were indicated with # ($P < 0.05$) or ## ($P < 0.01$). The markers detecting more than one locus were identified by letters-a, -b and -c after the marker names (Fig. 1).

To the best of our knowledge this is the first genetic map of groundnut based only on cultivated genotypes. Although a genetic map for tetraploid groundnut genome was developed earlier (Burow et al. 2001), this was based on RFLP markers on 78 BC₁F₁ lines derived from a cross of TxAG-6, a synthetic amphidiploid (Simpson et al. 1993) with Florunner. Because different marker types were used by Burow et al. (2001) and in this study, a direct comparison can not be made between these two maps.

In contrast, the map of the AA genome of *Arachis* (Moretzsohn et al. 2005 and unpublished data) was developed using SSR markers. Mapping of common markers thus allowed the alignment of these two maps in some regions (see later). As SSR markers are the markers of choice for plant geneticists and breeders (Gupta and Varshney 2000) and a larger number of SSR markers are available for groundnut (see Varshney et al. 2007; Table 1), it is anticipated that future groundnut genetic maps will involve SSR markers. Therefore, the developed SSR-genetic map of cultivated groundnut should be very useful to the community to compare the future genetic maps of groundnut with the map developed in the present study.

Trait phenotyping and QTL analysis

The parental genotypes of the mapping population, ICGV 86031 and TAG 24, were found to show variation in transpiration (T), TE and also for its surrogate traits-SLA, and SCMR (Serraj et al. 2004; Nautiyal et al. 2002). Therefore all 318 RILs were phenotyped for the above mentioned traits for two consecutive years at ICRISAT, Patancheru. Phenotyping of RILs for T, TE and other surrogate traits for 2 years, overall, showed a fairly good consistency across seasons/years/watering regimes, in spite



◀ **Fig. 1** A genetic linkage map for cultivated groundnut. Linkage groups containing syntenic markers between this map and the AA genome map (Moretzsohn et al. 2005) have been identified with *asterisk sign*. Distorted markers are indicated with *hash* ($P < 0.05$) or *double hash sign* ($P < 0.01$). QTLs identified for drought related traits, as mentioned in Table 3, have been shown by bars on right hand side of the linkage groups. Numbers on the left of each group are Kosambi map distances

of the range of variations among RILs being much lower in one season, and of the differences in the evaporative demands in the two seasons. Details about the phenotyping data and the reasons for range of variations in different seasons are given elsewhere (Krishnamurthy et al. 2007).

For both T and TE, moderate and consistent variations with moderate heritability were observed among RILs across the years (Table 2; ESM 1, 2). Distribution of RILs for TE in both experiments showed a quantitative inheritance of the trait, with TAG 24 showing the lowest TE value and ICGV 86031 the above average value (ESM 2; Krishnamurthy et al. 2007). QTL analysis of T and TE showed one QTL each for T and TE in 2004 while three QTLs each for T and TE in 2005 (Table 3; Fig. 1). Although these QTLs were identified with high LOD values, their phenotypic variation ranged between 4.7 and 10.3%.

SLA was measured at the start of drought stress imposition as well as at the time of harvest and showed variation with moderate levels of heritability in both years (Table 2). QTL analysis of SLA at the start of drought stress imposition showed five QTLs in 2004 and four QTLs in 2005. However for SLA measured at the time of harvest, two QTLs were identified in 2004 and three QTLs were identified in 2005. The phenotypic variation contributed by

these QTLs ranged from 3.5 to 17.6%. As Krishnamurthy et al. (2007) did not find any relation of SLA with TE, the QTLs for SLA are not of much importance.

SCMR at the start of stress imposition in both seasons, at 7 and 10 days after imposing the stress in 2004, and at 5, 10 and 15 days after imposing the stress in 2005 showed large and significant variation among RILs (Table 2). Indeed, the heritability values observed for the SCMR were the highest among all the traits studied, particularly during 2005. For each season data, eight QTLs were identified for SCMR measured at different time points. However like the other traits mentioned above, the phenotypic variation explained by these QTLs was in the range of 2.9–11.0% (Table 3).

Alleles with moderate additive effects were identified for most of the evaluated traits. These alleles, which should confer more tolerance to drought, were derived from both the tolerant (positive additive effect) and the susceptible (negative effect) parents (Table 3). Alleles that improve the trait being derived from parents agronomically inferior have been identified for several plant species (Xiao et al. 1998; Frary et al. 2004; Wang et al. 2004; Yoon et al. 2006).

To the best of our knowledge, this is the first report on identification of QTLs for drought related traits in groundnut. As a result, no comparison can be made on QTLs identified in this study with other studies in groundnut. It is, however, important to mention here that though several QTLs were identified for each trait in both seasons, none of the identified QTLs explained a high phenotypic variation that could be used for marker assisted breeding. However, given the highly polygenic nature of the traits analyzed (Krishnamurthy et al. 2007) and the

Table 2 Trait phenotyping data on ICGV 86031 and TAG 24 and its mapping population

Trait	Year	ICGV 86031	TAG 24	Variation in RILs	S.Ed (\pm)	σ_g^2 (SE)	Heritability
Transpiration (T, kg)	2004	1.42	1.47	1.27–1.51	0.051	0.002 (0.0004)	0.17
	2005	4.39	3.78	3.45–4.80	0.212	0.080 (0.0089)	0.35
Transpiration efficiency (TE, g kg ⁻¹)	2004	3.29	2.65	2.60–3.55	0.282	0.066 (0.011)	0.18
	2005	2.19	1.99	1.92–2.36	0.116	0.013 (0.002)	0.21
Specific leaf area (SLA, cm ² g ⁻¹) at start of stress	2004	147	153	137–169	7.6	51.4 (7.9)	0.30
	2005	165	192	161–201	6.9	84.1 (9.3)	0.36
Specific leaf area (SLA, cm ² g ⁻¹) at harvest	2004	130	148	117–171	9.3	101.4 (12.3)	0.29
	2005	143	156	131–158	8.4	71.6 (10.2)	0.23
SCMR at start of treatment	2004	49.2	43.8	40.7–50.1	2.14	5.0 (0.64)	0.26
	2005	46.9	42.3	39.1–50.4	1.91	5.3 (0.63)	0.31
SCMR after 5 days of treatment	2005	43	36.9	34.9–46.0	1.81	4.7 (0.54)	0.34
SCMR after 1 week of treatment	2004	52.8	50.3	43.1–55.2	2.59	6.7 (0.92)	0.23
SCMR after 10 days of treatment	2005	45.1	41.9	39.3–48.0	1.53	3.8 (0.41)	0.38
SCMR after 15 days of treatment	2005	47.6	42.2	39.9–50.9	1.50	3.9 (0.41)	0.39
SCMR at harvest	2004	51.5	49	42.9–55.8	2.25	5.0 (0.69)	0.23

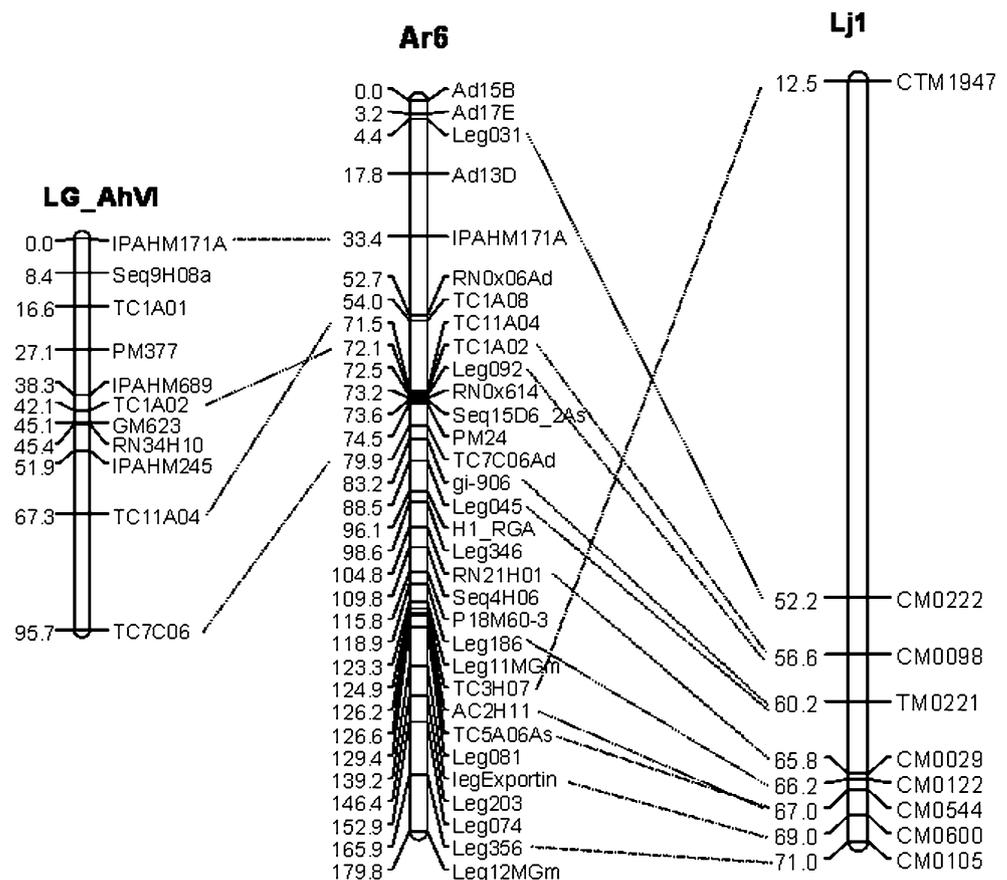
Table 3 Quantitative trait loci for drought tolerance-related traits identified by composite interval mapping (CIM) method

Trait	QTL name	Nearest marker	Position (cM)	Highest LOD score (threshold)	Phenotypic variation (r^2 %)	Additive effect
Transpiration (T)	T04_IV	pPGSSeq9G05	17.6	4.26 (2.8)	10.3	0.0123
	T05_VII	GM736	10.4	4.49 (2.6)	5.7	-0.0593
	T05_V	RN6F05	7.6	5.29 (2.6)	7.9	0.0651
	T05_XI	GM672	12.3	3.46 (2.6)	4.7	0.0544
Transpiration efficiency (TE)	TE04_VII	PM427	2.1	5.96 (2.8)	8.9	-0.0552
	TE05_III	pPGSSeq9H08b	0.1	4.00 (2.8)	5.8	-0.0183
	TE05_VII	LG111	8.7	6.04 (2.8)	8.1	-0.0226
	TE05_XIII	TC9B07	78.5	2.88 (2.8)	8.0	0.0223
Specific leaf area (SLA)	SLAPreTrt04_VII	PM427	0.1	2.97 (3.1)	3.5	1.0267
	SLAPreTrt04_VIII	PM721	76.2	7.89 (3.1)	17.6	1.7221
	SLAPreTrt04_XIII	TC9B07	64.9	3.96 (3.1)	7.6	-1.3542
	SLAPreTrt04_IV	PM183	73.8	3.01 (3.1)	8.4	1.6203
	SLAPreTrt04_XXII	GM625	30.0	5.57 (3.1)	13.0	-1.8143
	SLAHar04_XI	GM672	12.3	4.33 (3.0)	6.1	-2.0947
	SLAHar04_IV	pPGSSeq19H03	33.3	3.57 (3.0)	5.1	-1.8638
	SLAPreTrt05_VI	IPAHM689	37.1	4.53 (2.8)	6.5	-1.8793
	SLAPreTrt05_VIIIa	PM721	76.1	6.30 (2.8)	14.1	2.2488
	SLAPreTrt05_VIIIb	TC9F10	86.5	3.96 (2.8)	5.8	1.8663
	SLAPreTrt05_XV	PM85	15.8	3.43 (2.8)	5.3	-1.8296
	SLAHar05_XI	TC2D06	12.6	3.72 (3.3)	5.6	-1.2918
	SLAHar05_XVI	GM694	19.4	3.30 (3.3)	4.2	1.1874
	SLAHar05_XVII	IPAHM105	0.1	3.45 (3.3)	4.2	-1.2709
SPAD chlorophyll meter reading (SCMR)	SPADPreTrt04_VIII	GM679	78.7	5.28 (2.9)	7.1	-0.5219
	SPADPreTrt04_XVI	pPGPSeq2B09	2.3	6.02 (2.9)	10.6	0.6423
	SPADD005_XIII	TC9B07	66.5	5.41 (2.8)	11.0	0.5238
	SPADD005_X	pPGSSeq19A05	5.7	4.11 (2.8)	6.5	-0.4969
	SPADD005_XVII	IPAHM105	0.1	4.47 (2.8)	6.0	0.4783
	SPADD505_XVI	pPGPSeq2B09	4.0	4.18 (2.6)	7.6	0.5076
SPAD at stage of harvest	SPADStresStrt04_VIII	GM660	77.1	5.16 (3.0)	8.7	-0.4585
	SPADStresStrt04_XVI	pPGPSeq2B09	2.1	4.58 (3.0)	8.2	0.4982
	SPADD7UndrStres04_Xia	Ah-193	2.1	3.21 (2.9)	5.1	0.4528
	SPADD7UndrStres04_Xib	TC2D06	14.6	3.20 (2.9)	4.9	0.4475
	SPADD7UndrStres04_IV	GM723b	28.9	3.13 (2.9)	4.7	0.4359
	SPAD10Har04_VI	TC1A01	18.6	4.58 (2.8)	7.4	0.4595
	SPADD1005_X	IPAHM165	0.1	4.23 (2.8)	5.7	-0.3918
	SPADD1005_XVII	IPAHM105	0.1	3.03 (2.8)	4.1	0.3543
	SPADD1505_X	IPAHM165	0.1	3.64 (1.5)	5.1	-0.7242
	SPADD1505_XVII	IPAHM105	0.1	2.08 (1.5)	2.9	0.5748

relatively high number of progenies, it is not surprising to get QTLs with lower phenotypic variation (R^2 values). Based on QTL mapping studies in other species, it can be generalized that higher phenotypic variation for the given trait in the mapping population and high/reasonable marker density genotyping data are the pre-requisites to identify the major QTLs explaining higher phenotypic variation. However, in the present study, on one hand, range of

variations for the targeted traits was not very high in RILs, the marker density on the developed genetic map is also not very satisfactory. For instance, here the range of TE value was only between 2.60 and 3.55 g kg^{-1} water transpired in 2004 and between 1.92 and 2.36 g kg^{-1} water transpired in 2005. The marker density on this genetic map will be further improved after integrating more number of polymorphic markers. Further, it may be possible to identify

Fig. 2 Alignment of linkage group 6 (LG_AhVI) of developed map with diploid AA genome (linkage group Ar 6) and Lotus (linkage group Lj1)



more (and major) QTLs, for different traits, explaining higher phenotypic variation. Work is also in progress on methods to capture larger variations for some of the phenotypic traits.

Comparative maps with *Arachis* and model legume genomes

Two linkage groups LG 3 (AhIII) and LG 6 (AhVI) were chosen for aligning to the AA genome map (Moretzsohn et al. 2005; unpublished data). These groups consist of ten and eleven markers, respectively. Four markers from each group could be mapped in AA genome map giving consistent points of correspondence to Ar3 and Ar6, respectively. The alignment of cultivated LGs with the diploid allows the accumulation of information from both maps, and new possibilities to be explored. Through the use of Leg markers, supplemented with other sequence characterized markers, Ar3 and Ar6 were aligned with the genetic maps of *Lotus*, *Medicago*, and common bean (unpublished results; Hougaard et al. 2008). Therefore, in spite of the fact that a direct alignment of the groundnut map with these other legumes is impossible, an indirect alignment of cultivated genetic map through the reference AA genome map is feasible. Ar3 is syntenous with *Lotus* chromosomes 3 and 4, *Medicago* chromosomes 4 and 7 and

with bean LGs 2 and 11, therefore it can be inferred that AhIII is also syntenous with these same chromosomes/LGs. In the same way, AhVI corresponds to Ar6 which is syntenous with *Lotus* chromosome 1, *Medicago* chromosome 7 and, in part, with bean LG 6. The alignment of AhVI with Ar6 and Lj1 is illustrated in Fig. 2. From this alignment, it can be seen that AhVI is equivalent to the c.52.2–60.2 cM region of *Lotus* chromosome 1, the sequence of which has recently been published (Sato et al. 2008). Integration of more and common markers between AA genome map (Moretzsohn et al. 2005) and AABB genome map developed in the present study would establish more anchoring points among AA, AABB genomes maps of groundnut and other legume genome maps. The alignment of some parts of the present map with maps of AA genome of groundnut and *Lotus* and *Medicago*, demonstrates the possibility of using sequenced model legume genomes for improving our understanding of the groundnut genome, the generation of candidate genes, etc.

Conclusions

The present study reports the development of the first genetic map for cultivated groundnut after screening a

large number of SSR markers, available in public domain as well as unpublished ones. Low level of polymorphism, observed in the present study like in earlier studies, emphasises the need to develop a critical mass of polymorphic (SSR and SNP) markers, so that cultivated groundnut genetic maps with reasonable marker density can be developed in future. The present study also demonstrates the application of developed genetic map for identification of QTLs for drought tolerance related traits and comparative mapping. In summary the developed genetic map should be useful for the groundnut community to align the future genetic maps with it, and to transfer the sequence information from model legume species like *Lotus* and *Medicago* for enhancing the knowledge of comparative genome evolution of legumes as well as groundnut improvement.

Acknowledgments Authors are thankful to Mr A Gafoor and Mr G Somaraju for their help in conducting some experiments and collection of data. Financial support from National Fund of Indian Council of Agricultural Research (NBFSA), New Delhi, India and Generation Challenge Programme (<http://www.generationcp.org>) of CGIAR is gratefully acknowledged to sponsor this study.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Burow MD, Simpson CE, Starr JL, Paterson AH (2001) Transmission genetics of chromatin from a synthetic amphidiploid to cultivated peanut (*Arachis hypogaea* L.) broadening the gene pool of a monophyletic polyploid species. *Genetics* 159:823–837
- Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. *Genetics* 138:963–971
- Cuc LM, Mace ES, Crouch JH, Quang VD, Long TD, Varshney RK (2008) Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea*). *BMC Plant Biol* 8:55
- Doerge RW, Churchill GA (1996) Permutation tests for multiple loci affecting a quantitative character. *Genetics* 142:285–294
- Ferguson ME, Burow MD, Schulze SR, Bramel PJ, Paterson AH, Kresovich S, Mitchell S (2004) Microsatellite identification and characterization in peanut (*A. hypogaea* L.). *Theor Appl Genet* 108:1064–1070
- Frary A, Fulton TM, Zamir D, Tanksley SD (2004) Advanced backcross QTL analysis of a *Lycopersicon esculentum* × *L. pennellii* cross and identification of possible orthologs in the Solanaceae. *Theor Appl Genet* 108:485–496
- Fredslund J, Madsen LH, Hougaard BK, Nielsen AM, Bertioli D, Sandal N, Stougaard J, Schauser L (2006a) A general pipeline for the development of anchor markers for comparative genomics in plants. *BMC Genomics* 7:207
- Fredslund J, Madsen LH, Hougaard BK, Sandal N, Stougaard J, Bertioli D, Schauser L (2006b) GeMprospector—online design of cross-species genetic marker candidates in legumes and grasses. *Nucleic Acids Res* 34:W670–W675
- Fredslund J, Schauser L, Madsen LH, Sandal N, Stougaard J (2005) PriFi: using a multiple alignment of related sequences to find primers for amplification of homologs. *Nucleic Acids Res* 33:W516–W520
- Gimenes MA, Hoshino AA, Barbosa AVG, Palmieri DA, Lopes CR (2007) Characterization and transferability of microsatellite markers of the cultivated peanut (*Arachis hypogaea*). *BMC Plant Biol* 7:9
- Gupta PK, Varshney RK (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. *Euphytica* 113:163–185
- He G, Meng R, Newman M, Gao G, Pittman RN, Prakash CS (2003) Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.). *BMC Plant Biol* 3:3
- Hopkins MS, Casa AM, Wang T, Mitchell SE, Dean RE, Kochert GD, Kresovich S (1999) Discovery and characterization of polymorphic simple sequence repeats (SSRs). *Crop Sci* 39:1243–1247
- Hougaard BK, Madsen LH, Sandal N, Moretzsohn MC, Fredslund J, Schauser L, Nielsen AM, Rohde T, Sato S, Tabata S, Bertioli DJ, Stougaard J (2008) Legume anchor markers link syntenic regions between *Phaseolus vulgaris*, *Lotus japonicus*, *Medicago truncatula* and *Arachis*. *Genetics* 179:2299–2312
- Johansen C, Nigam SN (1994) Importance of drought stress and its alleviation in legumes. *Crop Sci* 24:17–19
- Kochert G, Stalker HT, Gimenes M, Galgano L, Lopes CR, Moore K (1996) RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). *Am J Bot* 83:1282–1291
- Kottapalli KR, Burow MD, Burow G, Burke J, Puppala N (2007) Molecular characterization of the U.S. peanut mini core collection using microsatellite markers. *Crop Sci* 47:1718–1727
- Krishna GK, Zhang J, Burow M, Pittman RN, Delikostadinov SG, Lu Y, Puppala N (2004) Genetic diversity analysis in Valencia peanut (*Arachis hypogaea* L.) using microsatellite markers. *Cell Mol Biol Lett* 9:685–697
- Krishnamurthy L, Vadez V, Jyotsna Devi M, Serraj R, Nigam SN, Sheshshayee MS, Chandra S, Aruna R (2007) Variation in transpiration efficiency and its related traits in a groundnut (*Arachis hypogaea* L.) mapping population. *Field Crops Res* 103:189–197
- Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174–181
- Mace ES, Varshney RK, Mahalakshmi V, Seetha K, Gafoor A, Leeladevi Y, Crouch JH (2007) In silico development of simple sequence repeat markers within the aescynomenoid/dalbergoid and genistoid clades of the Leguminosae family and their transferability to *Arachis hypogaea*, groundnut. *Plant Sci* 174:51–60
- Moretzsohn MC, Hopkins MS, Mitchell SE, Kresovich S, Valls JFM, Ferreira ME (2004) Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis of hypervariable regions of the genome. *BMC Plant Biol* 4:11
- Moretzsohn MC, Leoi L, Proite K, Guimaraes PM, Leal-Bertioli SCM, Gimenes MA, Martins WS, Valls JFM, Grattapaglia D, Bertioli D (2005) A microsatellite-based, gene-rich linkage map for the AA genome of *Arachis* (Fabaceae). *Theor Appl Genet* 111:1060–1071
- Nautiyal PC, Rachaputi NR, Joshi YC (2002) Moisture-deficit induced changes in leaf-water content, leaf carbon exchange rate and biomass production in groundnut cultivars differing in specific leaf area. *Field Crops Res* 74:67–79

- Neff MM, Turk E, Kalishman M (2002) Web-based primer design for single nucleotide polymorphism analysis. *Trends Genet* 18:613–615
- Nelson MN, Phan HTT, Ellwood SR, Moolhuijzen Paula M, Hane J, Williams A, O'Lone CE, Nyarko JF, Scobie M, Cakir M, Jones MGK, Bellgard M, Książkiewicz M, Wolko B, Barker SJ, Oliver RP, Cowling WA (2006) The first gene-based map of *Lupinus angustifolius* L.-location of domestication genes and conserved synteny with *Medicago truncatula*. *Theor Appl Genet* 113:225–238
- Palmieri DA, Bechara MD, Curi RA, Gimenes MA, Lopes CR (2005) Novel polymorphic microsatellite markers in section *Caulorrhizae* (*Arachis*, Fabaceae). *Mol Ecol Notes* 5:77–79
- Palmieri DA, Hoshino AA, Bravo JP, Lopes CR, Gimenes MA (2002) Isolation and characterization of microsatellite loci from the forage species *Arachis pintoii* (Genus *Arachis*). *Mol Ecol Notes* 2:551–553
- Passioura JB (1977) Grain yield, harvest index and water use of wheat. *J Aust Inst Agric Sci* 43:117–120
- Proite K, Leal-Bertioli SC, Bertioli DJ, Moretzsohn MC, da Silva FR, Martins NF, Guimaraes PM (2007) ESTs from a wild *Arachis* species for gene discovery and marker development. *BMC Plant Biol* 7:7
- Sandal N, Petersen TR, Murray J, Umehara Y, Karas B, Yano K, Kumagai H, Yoshikawa M, Saito K (2006) Genetics of symbiosis in *Lotus japonicus*: recombinant inbred lines, comparative genetic maps, and map position of 35 symbiotic loci. *Mol Plant Microbe Interact* 19:90–91
- Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, Sasamoto S, Watanabe A, Ono A, Kawashima K, Fujishiro T, Katoh M, Kohara M, Kishida Y, Minami C, Nakayama S, Nakazaki N, Shimizu Y, Shinpo S, Takahashi C, Wada T, Yamada M, Ohmido N, Hayashi M, Fukui K, Baba T, Nakamichi T, Mori H, Tabata S (2008) Genome structure of the legume, *Lotus japonicus*. *DNA Res* 1–13
- Seijo GJ, Lavia GI, Fernandez A, Krapovickas A, Ducasse D, Bertioli DJ, Moscone DEA (2007) Genomic relationships between the cultivated peanut (*Arachis hypogaea*—Leguminosae) and its close relatives revealed by double GISH. *Am J Bot* 94:1963–1971
- Serraj R, Krishnamurthy L, Jyotsna Devi M, Reddy MJV, Nigam SN (2004) Variation in transpiration efficiency and related traits in a groundnut mapping population. *Int Arachis Newslett* 24:42–45
- Simpson CE, Starr JL, Nelson SC, Woodard KE, Smith OD (1993) Registration of TxAG-6 and TxAG-7 peanut germplasm. *Crop Sci* 33:1418
- Tegelstrom H (1992) Detection of mitochondrial DNA fragments. In: Hoelzel AR (ed) *Molecular genetic analysis of populations: a practical approach*. IRL Press, Oxford, pp 89–114
- Varshney RK, Graner A, Sorrells ME (2005) Genomics-assisted breeding for crop improvement. *Trends in Plant Sci* 10:621–630
- Varshney RK, Hoisington DA, Tyagi AK (2006) Advances in cereal genomics and applications in crop breeding. *Trends in Biotechnol* 24:490–499
- Varshney RK, Hoisington DA, Upadhyaya HD, Gaur PM, Nigam SN, Saxena K, Vadez V, Sethy NK, Bhatia S, Aruna R, Gowda MVC, Singh NK (2007) Molecular genetics and breeding of grain legume crops for the semi-arid tropics. In: Varshney RK, Tuberosa R, Dordrecht (eds) *Genomic assisted crop improvement genomics applications in crops*. Springer, The Netherlands, pp 207–2420
- Voorrips RE (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. *J Hered* 93:77–78
- Wang S, Basten CJ, Zeng ZB (2007) Windows QTL cartographer 2.5. <http://statgen.ncsu.edu/qtlcart/WQTLCart.htm>. Cited 2007
- Wang D, Graef GL, Procopiuk AM, Diers BW (2004) Identification of putative QTL that underlie yield in interspecific soybean backcross populations. *Theor Appl Genet* 108:458–467
- Xiao J, Li J, Grandillo S, Ahn SN, Yuan L, Tanksley SD, McCouch SR (1998) Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. *Genetics* 150:899–909
- Yoon DB, Kang KH, Kim HJ, Ju HG, Kwon SJ, Suh JP, Jeong OY, Ahn SN (2006) Mapping quantitative trait loci for yield components and morphological traits in an advanced backcross population between *Oryza grandiglumis* and the *O. sativa japonica* cultivar Hwaseongbyeon. *Theor Appl Genet* 112:1052–1062
- Young ND, Weeden NF, Kochert G (1996) Genome mapping in legumes (Fam. *Fabaceae*). In: Paterson AH, Austin (eds) *Genome mapping in plants*. Landes Company, USA, pp 211–227
- Zeng ZB (1993) Theoretical basis of separation of multiple linked gene effects on mapping quantitative trait loci. *Proc Natl Acad Sci USA* 90:10972–10976
- Zeng ZB (1994) Precision mapping of quantitative trait loci. *Genetics* 136:1457–1468