Field Crops Research 122 (2011) 49–59

Contents lists available at ScienceDirect

Field Crops Research

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

# Identification of quantitative trait loci for protein content, oil content and oil quality for groundnut (*Arachis hypogaea* L.)

Cholin Sarvamangala [a,b,c], M.V.C. Gowda [a], R.K. Varshney [b,d,*]

[a] University of Agricultural Sciences (UAS), Krishinagar, Dharwad 580 005, Karnataka, India
[b] International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, Andhra Pradesh, India
[c] Genetics Division, Agharkar Research Institute (ARI), Pune 411004, Maharashtra, India
[d] Generation Challenge Programme (GCP), c/o CIMMYT, 06600 Mexico DF, Mexico

ARTICLE INFO

*Article history:*
Received 24 December 2010
Received in revised form 14 February 2011
Accepted 14 February 2011

*Keywords:*
Groundnut
Nutritional traits
Molecular markers
SSRs
QTLs
Molecular mapping

ABSTRACT

Very few efforts have been made to improve the nutritional quality of groundnut, as biochemical estimation of quality traits is laborious and uneconomic; hence, it is difficult to improve them through traditional breeding alone. Identification of molecular markers for quality traits will have a great impact in molecular breeding. An attempt was made to identify microsatellite or simple sequence repeat (SSR) markers for important nutritional traits (protein content, oil content and oil quality in terms of oleic acid, linoleic acid and oleic/linoleic acid ratio) in a mapping population consisting of 146 recombinant inbreed lines (RILs) of a cross TG26 × GPBD4. Phenotyping data analysis for quality traits showed significant variation in the population and environment, genotype × environment interaction and high heritability was observed for all the traits. Negative correlation between protein content and oil content, oleic acid and linoleic acid indicated their antagonistic nature. After screening >1000 SSR markers, a partial genetic linkage map comprising of 45 SSR loci on 8 linkage groups with an average inter-marker distance of 14.62 cM was developed. QTL analysis based on single marker analysis (SMA) and composite interval mapping identified some candidate SSR markers associated with major QTLs as well as several minor QTLs for the nutritional traits. Validation of these major QTLs using a wider genetic background may provide the markers for molecular breeding for improving groundnut for nutritional traits.



## 1. Introduction

Groundnut also called peanut is one of the principal oilseed as well as economic crops of the world. It is utilized for human consumption as a vegetable oil and protein, as fodder for livestock and as green manure. With about 26% protein, 48% oil and 3% fiber and high content of calcium, thiamine, and niacin, it has all the potential to be used as an economic food supplement to fight malnutrition. Thus, groundnut is nature's gift to man in general and to children, pregnant or nursing women and the poor in particular (Misra, 2006).

About 80% of total groundnut production in India is crushed for oil extraction, thus improvement in oil content and quality is of interest to plant breeders and millers. Development of cultivars in groundnut varies with the purpose for which it is put to use (Bandyopadhyay and Desai, 2000). For example, the most important quality requirements of groundnut as a source of oil are high protein and oil content in seed and high oleic acid resulting in high oleic/linoleic acid (O/L) ratio for longer oil stability. Cultivars with high O/L ratio, low oil/fat and high protein are suitable for confectionary purpose. Nutritional quality of oil is determined by its fatty acid composition. In groundnut, there are mainly eight fatty acids viz. palmitic (16:0), stearic (18:0), oleic (18:1), linoleic (18:2), arachidic (20:0), eicosenoic (20:1), behenic (22:0) and lignoseric (24:0). Among them, oleic acid, a monounsaturated fatty acid and linoleic acid, a polyunsaturated fatty acid account for 75–80% of the total fatty acids in groundnut oil. The remaining 20% is contributed by other fatty acids, among them; palmitic acid (10%) has the largest proportion (Kavera, 2008). From the nutritional point of view, oleic acid lowers bad cholesterol (LDL) as effectively as linoleic acid, but does not affect good cholesterol (HDL) levels (Kris-Eterton et al., 2001) hence it balances cholesterol, which is desirable for healthy heart. Saturated fatty acids are hyper-cholesterolemic, polyunsaturated fatty acids are hypo-cholesteromic, but monounsaturated fatty acids are neutral in this regard (Groff et al., 1996). Oils with higher proportion of unsaturated fatty acids can be heated to high temperatures without smoking, leading to faster cooking time and absorption of less oil (Miller et al., 1987).

* Corresponding author at: Centre of Excellence in Genomics (CEG), Building # 300, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, Greater Hyderabad, India. Tel.: +91 40 3071 3305; fax: +91 40 3071 3074/3075.
*E-mail address:* r.k.varshney@cgiar.org (R.K. Varshney).

0378-4290/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.fcr.2011.02.010

Larger genetic variation is available for these quality traits in the groundnut germplasm (Norden et al., 1987; Branch et al., 1990; Upadhyaya et al., 2005). However, selection for seed quality is practiced only in advanced breeding lines, as biochemical estimation for these traits in segregating populations is high resource requiring, cumbersome and time consuming. Biochemical analysis of most of these traits is postmortem and is also substantially influenced by genotype (G) × environment (E) interaction. Thus, it seems very complex and challenging to the breeders to undertake quality improvement in large-scale breeding programmes through conventional breeding approaches.

Molecular markers offer great scope for improving the efficiency of conventional plant breeding. With the advent of molecular markers, by using segregating populations for trait of interest for breeders, it has now become routine to map gene or quantitative traits loci (QTLs) and identify valuable alleles for the corresponding traits. Once the trait is mapped, the markers associated with them are efficiently employed in breeding programmes through marker-assisted selection (MAS). Markers not only eliminate the need of bio-chemical analysis and phenotypic evaluation in the early generation breeding programme, but also minimize the time required to develop new genotypes with desirable traits in the seedling stage itself, instead of waiting until harvest.

Molecular mapping studies have been conducted in past in groundnut for several traits, e.g. rust resistance (Mace et al., 2006; Varma et al., 2005; Mondal et al., 2007; Khedikar et al., 2010), nematode resistance (Burrow et al., 1996; Garcia et al., 1996), resistance to aphid vector causing groundnut rosette disease (Herselman et al., 2004), resistance to seed infection by *Aspergillus flavus* (Yong et al., 2005), drought tolerance traits (Varshney et al., 2009; Ravi et al., 2011). For the oil quality traits, some studies have been undertaken. For instance, based on conventional genetics and breeding studies, two recessive alleles $ol_1$ and $ol_2$ were identified for high and low oleic acid genotypes (Lopez et al., 2000). Loss of function of oleoyl-PC desaturase activity has been reported being solely responsible for the high oleic/linoleic acid (O/L) trait (Ray et al., 1993). Two homeologous genes, *ahFAD2A* and *ahFAD2B* have been found to control the oleolyl-PC desaturase activity (Jung et al., 2000) and cleaved amplified polymorphic sequences (CAPS) markers were developed to differentiate mutant and wild-type *ahFAD2A* alleles (Chu et al., 2007). Recently some efforts have been made to tag oil content, 100-seed weight and other yield contributing traits based on bulk segregant analysis (BSA) by using SSR markers (Gomez et al., 2009). However, to the best of our knowledge, not much effort has been made to locate genes/QTLs responsible for protein content, oil content, oleic acid, linoleic acid and O/L ratio in groundnut. Therefore the present study has been undertaken to develop a genetic map and identify the QTLs for the above traits by using SSR markers and TG26 × GPBD4 mapping population of groundnut.

## 2. Materials and methods

### 2.1. Plant material

The $F_9$:$F_{10}$ generations of 146 recombinant inbred lines (RILs) obtained from a cross TG26 × GPBD4 by single seed descent method from $F_2$ onwards developed at University of Agricultural Sciences (UAS), Dharwad, Karnataka (India) was used for the study. TG26 is an improved Spanish bunch variety, it is a semi dwarf, erect with high pod growth rate, high harvest index, greater partitioning efficiency, tolerance to bud necrosis and has high linoleic acid content but it is susceptible to rust and late leaf spot (Kale et al., 1997; Badigannavar et al., 2002). GPBD4 is an improved Spanish bunch groundnut variety developed at UAS, Dharwad, it is popular in Karnataka and Southern states of India (Gowda et al., 2002). It has a desirable combination of early maturity, high yield, high pod growth rate, desirable pod and kernel features, high oil and protein content, optimum oleic/linoleic acid (O/L) ratio and resistant to late leaf spot and rust.

### 2.2. Experimental design and phenotyping

A total of 146 recombinant inbred lines (RILs) were sown in a randomized block design (RBD) in two replications at UAS, Dharwad. Ten seeds of each RIL were planted in 1 m row with 30 cm and 10 cm inter and intra-row spacing, respectively. Two parental genotypes (TG26 and GPBD4) were sown as controls after every 50 rows. Phenotyping was done for protein content, oil content, oleic acid, linoleic acid and O/L ratio in two experiments viz., *Kharif* 2007 (1st environment – E1) and summer 2007 (2nd environment – EII). Observations were recorded in two replications for each line in both the experiments.

Phenotyping for protein content (%), oil content (%), oil quality with respect to oleic acid, linoleic acid were estimated using Near Infrared (NIR) spectroscopy model 6500 (Foss NIR systems, France) and O/L ratio was calculated as the ratio of oleic acid and linoleic acid. 15–20 g sample seed from each RIL and parents in two replications was used for analysis. The calibration equations were developed using principle component regression (PCR), partial least square and modified partial least square (mPLS) regression models. Wavelengths at interval of 8 nm across the entire visible-plus-near-infrared spectrum (visible: 408–1092 nm; Near Infrared: 1108–2492 nm) were used for calibration. The standard error of calibration (SEC), standard error of cross-validation (SECV), correlation coefficient ($r$), and 1-VR statistics were used to select the best calibration equations. The performance of the calibration equations were monitored using the cross validation and external validation of set of samples ($n$ = 100). SECV, standard error of prediction (SEP) and $r$ were used to determine the accuracy of prediction (Kavera, 2008). The best equation for determining the protein content, oil content and fatty acid composition were developed and used for the subsequent analysis of fatty acid profile for parents and 146 RILs of the mapping population.

### 2.3. Molecular marker analysis

For DNA extraction, young leaf tissues were collected from parents and RILs at $F_{10}$ generation from two weeks old plants. SIGMA Genelute plant genomic DNA extraction kit was used to isolate DNA as per the manufacturer's recommendations. DNA quality and quantification was checked on 0.8% agarose gel with known concentrations of uncut lambda DNA standard.

Polymerase chain reactions (PCRs) with SSR markers on DNA of parental genotypes or RILs were performed in 5 μl reaction mixture using GeneAmp® PCR system ABI 9700 (Applied Biosystems, USA) as mentioned in Khedikar et al. (2010). Amplified products were tested on 1.2% agarose gel to check for amplification before the size separation. TPCR annealing temperature varied between 60 °C and 65 °C depending on the primers.

Separation of amplified DNA fragments were performed on 6% polyacrylamide gel electrophoresis (PAGE) and capillary electrophoresis i.e. ABI 3700 Genetic Analyzer (Applied Biosystems, USA) depending on the use of normal and florescent dyes labeled primers respectively as mentioned in Khedikar et al. (2010). Allele sizing and scoring based on capillary electrophoresis data was carried out using Genescan 3.1 and Genotyper 3.1 softwares (Applied Biosystems, USA) while manual scoring was done on PAGE data. In summary, alleles obtained were scored as A, B, H and O, where, A represents homozygosity for the allele from female parent (TG26), B indicates the homozygosity for the allele from male parent (GPBD4), H represents the heterozygotes i.e. the presence of both A