Validation of SSR markers linked to oil content in groundnut (*Arachis hypogaea* L.)

N DEVASENA, B K ANITHA, N MANIVANNAN*, G NALLATHAMBI[1], P JANILA[2], M K PANDEY[2] AND R K VARSHNEY[2]

Department of Oilseeds, Tamil Nadu Agricultural University, Coimbatore-641 003, Tamil Nadu

(Received: January 25, 2017; Revised: May 26, 2017; Accepted: June 12, 2017)

ABSTRACT

A set of 14 SSR markers that are specific for six QTLs for oil content reported by various authors was selected for validation in the present study. The $F_{4:5}$ mapping population of the cross ICGV 00440 x ICGV 03128 was used to validate the markers as well as QTLs specific for oil content. Among the selected markers, six markers were polymorphic for the parents ICGV 00440 and ICGV 03128. Among the six polymorphic SSR markers, two markers IPAHM103 and PGS16F10 revealed a strong association for oil content with a PVE of 15.3 and 19.6 respectively in single marker analysis. Validation of QTLs was also performed through composite interval mapping analysis. A QTL with flanking markers IPAHM103 and PM36 could be considered as a potential tool for marker assisted selection of the trait oil content in groundnut.

Keywords: Groundnut, Oil content, QTL, SSR markers, Validation

Groundnut (*Arachis hypogaea* L.) is a major oilseed crop worldwide and one of the most widely produced legume (Weiss, 2000; CGIAR, 2010). It is used as food (raw, roasted or boiled), animal feed (pods, seeds and plant material) and for industrial raw material. The products derived from groundnut include flour, oil, peanut butter, confectionary and paste. Chemical and epidemiological studies conducted by Dean *et al.* (2009) suggest that the nutritional properties of peanuts are favorable due to the fatty acid profiles, a high quality protein and a source of naturally occurring folate. Regular consumption of peanuts in diet has been confirmed to have positive impacts on human health by providing significant source of protein (20 to 36%), edible oil (45 to 50%) containing essential fatty acids, carbohydrates, fiber, folacin, phosphorus, magnesium, zinc, iron, potassium, calcium, vitamins (riboflavin, thiamine, niacin) and three forms of fat-soluble tocopherols (α, γ and δ), which are the most important lipid soluble antioxidants discovered by Herbert Evans in 1922 (Packer *et al*., 2001; Munné-Bosch, 2005; Mondal and Badigannavar, 2016; Ajay *et al*., 2016).

Groundnut occupies a unique position among oilseeds as it contains approximately 50 per cent oil which is comparably higher than other vegetable oil. About two-thirds of world groundnut production is crushed for cooking oil and the remaining one-third is used in the form of edible products. The production of groundnut is largely confined to Asian and African countries. Asia accounts for about 50 per cent of area and 60 per cent of world production of groundnut with largest share of India (>20%) in the groundnut coverage, followed by China (>18%). However, China accounts for highest share (37%) in the total production of groundnut in the world. As per 2013-14 (GOI, 2014), India being the second largest groundnut producing nation occupies an area of 52.6 lakh ha with a production of 96.73 lakh tonnes which accounts for a productivity of 1750 kg/ha. Generally, the groundnut yields are low in the developing countries where the focus is more on to get rid of the obstacles for improving yield. The present production and productivity has to increase at a much higher rate to meet the growing needs of the oil market. It is estimated that each one per cent increase in oil content would raise the processor's benefit by seven per cent (Liao and Holbrook, 2005).

Presently, groundnut breeding programme is faced with the challenge of improving oil content and enhancing yield. Yield and oil content is a complex trait, polygenic in nature with significant environmental influences. Though selection for yield is practiced in early generations, selection for oil content is practiced only in advanced breeding lines as biochemical estimation or through NIR/NMR in segregating populations is cumbersome and demands high resources, as well as time. It is possible to specifically target such traits for improvement, and also enhance the efficiency of overall breeding program through use of molecular breeding strategies.

Molecular markers have been used to identify genomic regions (QTL) involved in expressing seed oil content, through available molecular information/approaches and then attain the genetic gain by selection of QTL loci. SSR marker

[1]Department of Millets, Tamil Nadu Agricultural University, Coimbatore-641 003, Tamil Nadu;
[2]International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, Telangana, India;
*E-mail: nmvannan@gmail.com

has been the most widely used at identifying molecular variation within the cultivated species of groundnut and considerable progress has been made in tagging economically important traits (Selvaraj *et al*., 2009; Ferguson *et al*., 2004). The markers associated with QTLs cannot be directly used in MAS as it requires validation in other population. An attempt was made to validate the markers and reported QTLs associated with oil content in the $F_{4:5}$ population of the cross ICGV 00440 x ICGV 03128.

MATERIALS AND METHODS

Plant material: The present study was done in the $F_{4:5}$ generation developed from cross ICGV 00440 x ICGV 03128. It consists of 103 RILs developed by single seed descent method at Department of Oilseeds, Tamil Nadu Agricultural University, Coimbatore. The parent ICGV 00440 is Spanish bunch type with an oil content of 42-45 per cent. Parent ICGV 03128 is Virginia bunch type with oil content of 52-55 per cent and drought tolerant (Table 1).

Experimental design and phenotyping: A total of 103 recombinant inbred lines (RILs) in their F_4 generation were sown along with their parents in 3m rows with spacing of 30 x 10 cm in progeny rows during rainy season, 2013. Single plants were tagged in each RIL. The DNA samples were collected from the tagged plants and used for genotyping. The tagged F_4 progenies along with three checks were raised in 3 m row following Augmented Block Design I to study the $F_{4:5}$ performance during rainy season 2014. Spacing adopted was 30 x 10 cm. Oil content in progeny row bulk seed samples were measured on a BRUKER Matrix-I NIR spectrometer in absorbance mode. The validation was carried out using 55 samples with known oil content as estimated using Sox-plus for oil content estimation. The samples were scanned in a wave number range of 6000 - 4000 cm^{-1} with a resolution of $2cm^{-1}$ (Sundaram *et al*., 2011). The sample was scanned 32 times. After validation with an R^2 value of 98.69 per cent, the analysis of oil content in 103 RILs, parents and checks was estimated.

Molecular marker analysis: For DNA extraction, leaves were collected from 103 RILs as well as from the two parents in two leaf stages. DNA extraction was performed according to the cetyl trimethyl ammonium bromide (CTAB) method (Doyle and Doyle, 1990). The extracted DNA content was measured using DNA standards in agarose gel (0.8 % w/v). The 10 µl PCR cocktail contained 20 ng of 2 µl DNA, 1 µl of 10XTaq buffer, 0.2 µl of 25 mM $Mgcl_2$, 1.0 µl of 0.2 mM of dNTP, 0.5 µl of 0.5 uM of each forward and reverse primer, 4.5 µl of sterile water and 0.3 µl of 0.03 IU Taq DNA polymerase. DNA amplification was performed in a 96-Well Fast Thermal Cycler (Applied Biosystems Inc., Foster city, CA). DNA samples were denatured initially at 94°C for 3 min, then subjected to the following 20 cycles: 94°C for 30 s, 63°C for 30 s with a decrement of 0.5°C per cycle, and 70°C for 1 min. This was followed by another 20 cycles of 94°C for 15 s, 55°C for 30 s, and 70°C for 1 min. A 10 min extension was performed at 72°C as the last step. Amplified products were analyzed using 6% polyacrylamide gel electrophoresis at 150 volts DC for 4 hrs and silver stained in accordance with the protocol described by (Benbouza *et al*., 2006). Fourteen SSR markers that account for six already reported QTLs (Table 2 and 3) were used to study the polymorphism between the two parents. Among these, six markers *viz*., IPAHM103, PGS16F10, PM 36, GM1890, GM1878 and GM1961 were found to be polymorphic. The F_4 generation was genotyped with the six polymorphic primers.

Marker trait association: Clear and unambiguous bands were scored for their presence or absence with the score 1 indicating their presence and 0 indicating their absence. The data matrix of binary codes thus obtained was subjected to further analysis *viz.,* single marker analysis and composite interval mapping analysis.

Single marker analysis: Phenotypic values of all the 103RILs were subjected to associate with corresponding marker score for its significance by using simple regression in SPSS software (version.16). Simple linear regression method (Haley and Knott, 1992) was used to identify significant marker trait association. The linear equation formed was as follows:

$$Y = \mu + f(\text{marker}) + \text{error}$$

where, Y = phenotypic trait value; μ = population mean and f (marker) = function of the molecular marker.

The potential relationship between the marker and trait was established considering the significance of the regression coefficient at 5 and 1 per cent probability. The phenotypic variance explained (PVE) was expressed in terms of adjusted R^2 values.

Composite interval mapping: QTL ICI mapping version 3.2 (Wang *et al*., 2012) was used to analyze the data for composite interval mapping. The reported map position of three QTLs by various authors (Pavithradevi, 2013; Sarvamangala *et al*., 2011; Pandey *et al*., 2014) was used to validate through inclusive composite interval mapping of additive and dominant (ICIM-ADD) method. The ICIM-ADD was performed with mapping parameters as scan interval of 1cM, 0.001 probability in step-wise regression and LOD threshold of 3.0.

RESULTS AND DISCUSSION

One of the most important use of QTL mapping is to apply them in marker-assisted selection (MAS) for genetic improvement of quantitative traits. The already identified QTLs has to be validated in a different population to use it further in marker assisted selection. Once the tightly linked markers have been identified, the traits can be selected indirectly using MAS. In the present study, phenotype data were obtained from F_5 population of ICGV 00440 x ICGV 03128 and genotypic data from F_4 population. The analysis of variance for the F_5 generation of the cross ICGV 00440 x ICGV 03128 revealed significant differences for oil content. Frequency distribution for the oil content revealed a typical normal distribution indicating their quantitative nature of inheritance (Fig. 1).Validation of reported SSR markers and known QTLs was performed in the $F_{4:5}$ generations. Microsatellites (SSRs) are markers of choice in groundnut as they are ubiquitous, co-dominant and multi-allelic in nature. A total of 14 markers that account for reported QTLs linked to oil content were selected, of which only six markers were polymorphic. The reason for low level polymorphism is due to the narrow genetic base between the two parents selected for mapping population. Young *et al.* (1996) reported that low level of genetic polymorphism in cultivated groundnut was attributed to its origin from a single polyploidization event that occurred relatively on an evolutionary time scale.

For single marker analysis, the markers were subjected to single factor regression analysis using the marker as independent and the respective phenotype as dependent. The result of the single marker analysis for the trait oil content is presented in Table 2. Among the six markers, only three markers IPAHM103, PGS16F10 and PM 36 revealed a strong association for oil content with a PVE of 15.3,19.6 and 3.8, respectively (Table 4). Persuasive results were reported by Anitha *et al*. (2014) for the marker IPAHM103. Sarvamangala *et al*. (2011) also reported similar results for the same marker for two different seasons. The marker PGS16F10 was reported by Pavithradevi (2013) in the QTL region PM36_PGS16F10 for the trait oil content. The marker PM36 recorded a comparatively low PVE of 3.8. This was also reported in the QTL region IPAHM103_ PM36 by Sarvamangala *et al*. (2011).

Three QTL region reported by previous workers (Pavithradevi, 2013; Sarvamangala *et al*., 2011; Pandey *et al*., 2014) for oil content were used for QTL validation. Among the three linkage groups, IPAHM103_PM36 recorded significant QTL with 3.07 LOD and 12.82% PVE (Table 5). This reiterates the findings of Sarvamangala *et al*. (2011). The distance between the flanking markers is 9.5cM.Thus these two markers can be of utmost importance in MAS for oil content. Inclusion of some more markers in this region will further improve the success of MAS.

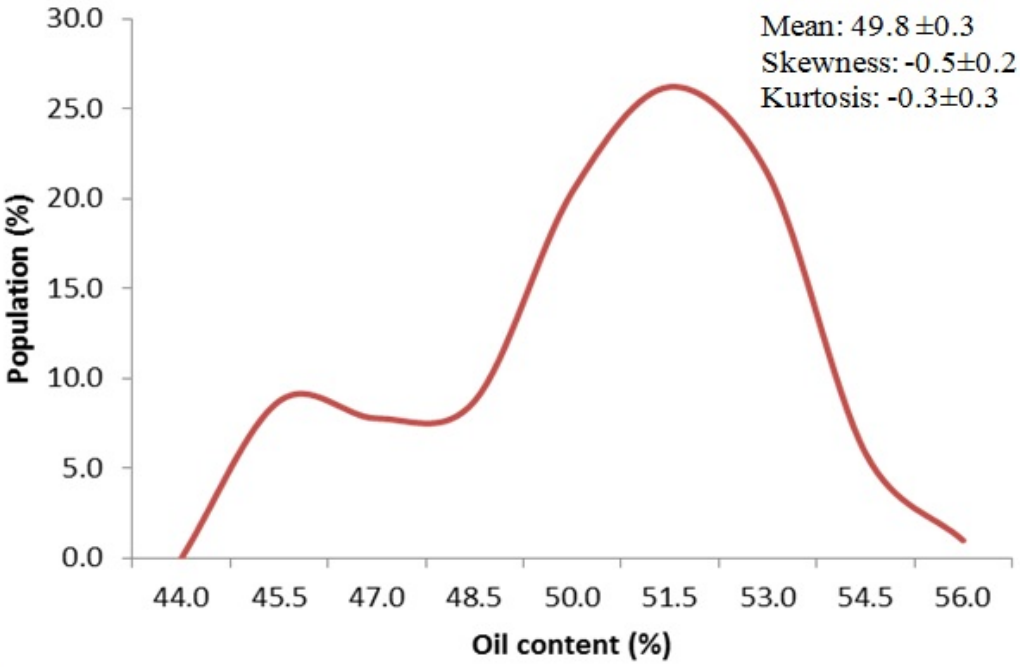

Fig. 1. Frequency distribution for oil content in F_5 progenies of the cross ICGV 00440 x ICGV 03128

Table 1 Details of parents used in the mapping population

Particulars	ICGV 00440	ICGV 03128
Pedigree	(ICGV 88386 x ASHFORD) x ICGV 95172	(ICGV 99160 x ICGV 99240)
Habit	Spanish bunch	Virginia bunch
Oil content (%)	42-45	52-55
Oleic (%)	45-50	35 -43
Linoleic (%)	28-30	40-57
Special features	Low oil and Confectionery line	High oil and Drought tolerant

Table 2 List of reported QTLs and markers linked to oil content and used for validation

Linkage group	Identified QTLs		Position (cM)	Interval (cM)	PVE (%)	Markers	Reference
	Left marker	Right marker					
2	AC2AO6	GM1907	120	126.1-165.6	23.7	AC2AO6,GM1907	Pavithradevi (2013)
2	PM36	PGS16F10	220	209.9-243.3	24.5	PM36, PGS16F10	
3	IPAHM103	PM36	28	25.5-34.0	10.2	IPAHM103,PM36	Sarvamangala *et al*. (2011)
5	GM1702	GM1878		0-23.8	10.23	GM1702,GM1878	Pandey *et al*. (2014)
5	GM1878	GM1890	28	25.8-31.8	25.52	GM1878-GM1890	
8	GM2690	IPAHM606	-	37.4-50.5	14.07	GM2690, IPAHM123 GM2689,GM1961, PM505,IPAHM606	

Table 3 List of fourteen SSR primers and sequences used for F_5 genotyping

SSR primer	Forward (5'-3')	Reverse (5'-3')
GM1702	GATTGGGAAGCAGCAAGAAG	CAACCAGCTCCTTCTCTACCC
GM1878	TCAGTGGTTCAGTGCATCAAG	GTCCCTTGGTCATCTTCGATT
GM2690-1	GACGCCGTGGTTTATGACTT	CAACCAGCTCCTTCTCTACCC
IPAHM123	CGGAGACAGAACACAAACCA	TACCCTGAGCCTCTCTCTCG
GM2689	GACGCCGTGGTTTATGACTT	CAACCAGCTCCTTCTCTACCC
GM1961	TGTATTCTCCCTGAAATGACGA	CTTCTTCCTCCATCCTCCCTA
PM505	TCCTCACATTGACGATGACC	CGGAGAACGAGAGGTTGAAG
IPAHM606	CCTAACTCAGCCTGCGAAAC	CAGAGGTGTTTGGAGAACTAGGA
GM1890	CTCTCCGATTATAGGCCAACC	TGGCTTCTCCGTGAAAATAAC
AC2A06	ATCATCTCGATCCATCCTTCTG	CTCCTTCTTCTCGCGTATTTGT
GM1907	CACTGTCCTCTTCCCTCACTCT	GGTGGACGAAGAAGAAGAAGAA
PM36	ACTCGCCATAGCCAACAAAC	CATTCCCACAACTCCCACAT
IPAHM103	GCATTCACCACCATAGTCCA	TCCTCTGACTTTCCTCCATCA
PGS16F10	TGGAGGGAAAAACATTTTGG	CCTGGAGGGGTGAGAGGT

Table 4 Single marker analysis for oil content in $F_{4:5}$ population of ICGV 00440 x ICGV 03128

Marker name	PVE	Significance
IPAHM103	15.3	**
PGS16F10	19.3	**
PM36	3.8	**
GM1890	2.4	not significant
GM1878	1.2	not significant
GM1961	0.9	not significant

Table 5 Details of QTL validated for oil content

Trait Name	Position cM	Left Marker	Right Marker	LOD	PVE (%)	Additive Effect
Oil content (%)	25.5	IPAHM103	PM36	3.07	12.82	ICGV 03128

ACKNOWLEDGEMENTS

This research is a part of the first author's Ph. D. research work which was supported by ISOPOM-ICRISAT project funded by Integrated Scheme of Oilseeds, Pulses, Oilpalm and Maize (ISOPOM), Department of Agriculture and Co-operation (DAC), Ministry of Agriculture, Government of India.

REFERENCES

Ajay B C, Meena H N, Singh A L, Dagla M C, Kumar N, Bera S K, Makwana A D and Kalariya K A 2016. Genotypic variability and inheritance of iron, phosphorus, potassium and zinc contents in segregating generations of peanut (*Arachis hypogaea* L.). *Journal of Oilseeds Research*, **33**: 102-107.

Anitha B K 2013. Identification of quantitative trait loci for oil yield and marker assisted backcross for high oleic acid in groundnut (*Arachis hypogaea* L.). Ph.D. Thesis, Tamil Nadu Agricultural University, Coimbatore.

Benbouza H, Jacquemin J M, Baudoin, J P and Mergeai G 2006. Optimization of a reliable, fast, cheap and sensitive silver staining method to detect SSR markers in polyacrylamide gels. *Biotechnology, Agronomy, Society and Environment*, **10**: 77-81.

CGIAR 2010. World Geography of the Peanut. Consultative Group of International Agricultural Research, http://lanra.anthro.uga. edu/peanut/database.

Dean L L, Henrix K W, Holbrook C C and Sanders T H 2009. Content of some nutrients in the core of the peanut germplasm collection. *Peanut Science*, **36**: 104-120.

Doyle J J and Doyle J L 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, **19**: 11-15.

Ferguson M E, Burow M D, Schulze S R, Bramel P J, Paterson, A H, Kresovich S and Mitchell S 2004. Microsatellite identification and characterization in peanut (*A. hypogaea* L.). *Theoretical and Applied Genetics*, **108**: 1064-1070.

Haley C S and Knott S A 1992. A simple regression method for mapping quantitative trait *loci* in line crosses using flanking markers. *Heredity*, **69**: 315-324.

Liao B S and Holbrook C C 2005. Groundnut. *In: Genetic resources, chromosome engineering, and crop improvement. Oilseed Crops,* (Singh R J ed.), Boca Raton, FL, USA: CRC Press, pp. 51-57.

Mondal S and Badigannavar A M 2016. Development of high protein groundnut (*Arachis hypogaea* L.) mutant through induced mutagenesis. *Journal of Oilseeds Research*, **33**: 98-101.

Munné-Bosch S 2005. The role of alpha-tocopherol in plant stress tolerance. *Journal of Plant Physiology*, **162**: 743-748.

Packer L, Weber S U and Rimbach G 2001. Molecular aspects of α-tocotrienol antioxidant action and cell signalling. *Journal of Nutrition*, **131**: 369-373.

Pandey M K, Wang M L, Qiao L, Feng S, Khera P, Wang H, Tonnis B, Barkley N A, Wang J, Holbrook C C, Culbreath A K, Varshney R K and Guo B 2014. Identification of QTLs associated with oil content and mapping FAD2 genes and their relative contribution to oil quality in peanut (*Arachis hypogaea* L.). *BMC Genetics*, **15**: 133.

Pavithradevi S 2013. Identification of quantitative trait loci (QTLs) for yield and yield component traits under drought stress in Spanish bunch groundnut (*Arachis hypogaea* L.). Ph.D. Thesis, Tamil Nadu Agricultural University, Coimbatore.

Sarvamangala C, Gowda, M V C and Varshney R K 2011. Identification of quantitative trait loci for protein content, oil content and oil quality for groundnut (*Arachis hypogaea* L.). *Field Crop Research*, **1**: 49-59.

Selvaraj M G, Manivannan N, Schubert A M, Ayers J L, Baring M R and Burow M D 2009. Identification of QTLs for pod and kernel traits in cultivated peanut by bulked segregant analysis. *Electronic Journal of Biotechnology*, **12**: 1-13.

Sundaram J, Kandala C V, Butt C L, Chen C Y and Sobolev V 2011. Non-destructive NIR reflectance spectroscopic method for rapid fatty acid analysis of peanut seeds. *Peanut Science*, **38**: 85-92.

Voorrips R E 2006. MapChart 2.2: software for the graphical presentation of linkage maps and QTLs. Wageningen: Plant Research International, Wageningen, The Netherlands.

Wang J, Li H, Zhang L and Meng L 2012. User's Manual of QTL ICIMapping version 3.2. The Quantitative Genetics group, Institute of crop science, Chinese Academy of Agricultural Science (CAAS), Beijing, China and Genetic Resources Programme, CIMMYT, Mexico, 208 pp.

Weiss E A 2000. *Oil seed Crops.* London: Blackwell Science, 364 pp.

Young N D, Weeden N F and Kochert G 1996. Genome mapping in legumes (Family Fabaceae). *In: Genome Mapping in Plants,* Austin T X and Landes R G (eds), pp. 211-227.