Published in Field Crops Research

Volume 122, Issue 1, 28 April 2011, Pages 49-59

http://dx.doi.org/10.1016/j.fcr.2011.02.010

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

Cholin Sarvamangala[1,2,3], M.V.C. Gowda[1], and R.K.Varshney[2,4,*]

[1]University of Agricultural Sciences (UAS), Krishinagar, Dharwad-580 005, Karnataka, India

[2]International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, India

[3]Genetics Division, Agharkar Research Institute (ARI), Pune-411004, Maharashtra, India

[4]Generation Challenge Programme (GCP), c/o CIMMYT, 06600 Mexico DF, Mexico

*Author for correspondence

Address:

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: 0091 40 30713305, Fax: 0091 40 3071 3074/ 3075

E-mail: r.k.varshney@cgiar.org

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 x GPBD 4. Phenotyping data analysis for quality traits showed significant variation in the population and environment, genotype x 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 QTL 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.

Key words: Groundnut, nutritional traits, molecular markers, SSRs, QTLs, molecular mapping

1. Introduction

Groundnut also called peanut is one of the principal oil seed 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 per cent protein, 48 per cent oil and 3 per cent 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, 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 to 80 per cent of the total fatty acids in peanut oil. The remaining 20% is contributed by other fatty acids, among them; palmitic acid (10%) has the largest proportion (Kaveri 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).

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) x environment (E) interaction. Thus, it seems very complex and

challenging to the breeders to undertake quality improvement in large-scale breeding programs 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 chemical analysis and phenotypic evaluation in the early generation breeding program, 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, eg. 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 maded to tag oil content, 100-seed weight and other yield contributing traits based on bulk segregate analysis 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 x GPBD 4 mapping population of groundnut.

2. Material and methods

2.1 Plant material

The F_9:F_{10} generations of 146 recombinant inbred lines (RILs) obtained from a cross TG 26 x GPBD 4 by single seed descent method from F_2 onwards developed at U.A.S. Dharwad, Karnataka (India) was used for the study. TG 26 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). GPBD 4 is an improved Spanish bunch groundnut variety developed at University of Agricultural Sciences, 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 U.A.S. Dharwad. Ten seeds of each RIL were planted in one meter row with 30 cm and 10 cm inter and intra-row spacing, respectively. Two parental genotypes (TG 26 and GPBD 4) were sown as controls after every 50 rows. Phenotyping was done for protein content, oil content, oleic acid, linoleic acid and O/L ratioin two experiments *viz*., *Kharif* 2007 (1^{st} environment-E1) and summer 2007 (2^{nd} 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. 15-20grams 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 five 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 60^0C to 65^0C 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 (TG 26), B indicates the homozygosity for the allele from male parent (GPBD 4), H represents the heterozygotes *i.e* the presence of both A and B alleles and O represents off types (neither A nor B) or missing values.

2.4 Statistical analysis

2.4.1 Phenotypic data

The replicated phenotypic data obtained for protein content, oil content, oleic acid, linoleic acid and O/L ratio was used for analysis of variance (ANOVA) pooled over environments (PE). Mean and range among the RILs in comparison with parents and genetic variability components such as, phenotypic coefficient of variation (PCV), genotypic coefficient of variation (GCV) and heritability in broad sense (h^2b.s.) were estimated in individual (EI and EII) as well as pooled across environments (PE). Correlation coefficients (r) among these traits and the frequency

distributions were also estimated using pooled data (PE) only as there was no difference in the pattern of correlation and distribution in individual seasons All the necessary computation for the field trial was performed with the software packages, Genstat 10^{th} edition (Payne et al., 2007) and SPSS 16^{th} version.

2.4.2 Construction of Linkage map

The Chi-square test was used to assess goodness of fit to the expected 1:1 segregation ratio for each marker. All the markers including those with distorted segregation were employed for linkage analysis using MAPMAKER Macintosh version 3.0 (Lander *et al*., 1987) as mentioned in Khedikar *et al.* (2010). Recombination fraction was converted into map distances in centiMorgans (cM) using Kosambi mapping function (Kosambi 1944). The inter-marker distances calculated from MAPMAKER were used to construct linkage map using MAPCHART version 2.2 (Voorrips 2006).

2.4.3 Marker-trait association

Marker trait association was conducted by using single marker analysis (SMA) and composite interval mapping (CIM). For SMA, the mean phenotypic data of all the traits in individual environments and the genotypic data of 53 markers pertaining to 146 RILs were analysed using simple linear regression method with help of Genstat (10^{th} edition) programme (Haley and Knott 1992). The phenotypic variance explained was expressed in adjusted R^2 values. For CIM analysis (Zeng 1994; Jansen and Stam 1994), the software package PLABQTL version 1.1w (Utz and Melchinger, 1996) was used to identify QTLs and to estimate the additive effects with their phenotypic variance.The phenotypic data for the traits in two individual environments (EI and EII) and pooled data over two environments (PE) were combined with marker genotyping data to identify QTLs in individual environment and QTL x Environment respectively.

3. Results

3.1 Phenotypic data analysis

Analysis of variance in individual environments revealed significant variation within the population (data not shown) and analysis of variance pooled over the environments revealed significant variation among the genotypes, environments and significant G x E interaction for protein content, oil content, oleic acid, linoleic acid and O/L ratio indicating the existence of environmental interaction (Table 2).

The mean protein content was high in EI for both parents and RILs compared to EII but broader range was observed in EII (21.12-37.51%) compared to EI (24.02-36.64 %) and hence, the magnitude of variation (PCV, GCV) was higher in EII as compared to EI. Very high heritability (>80%) was observed in both the environments but in the pooled data the heritability was lower (<50 %). The male parent GPBD 4, was a higher value parent for all the favorable traits *i.e* for protein content, oil content, oleic acid and O/L ratio and TG 26 was higher value parent for linoleic acid. Higher mean oil content was observed for EI compared to EII in both parents and RILs but unlike protein content, the range was high in EI (40.76-49.03) compared to EII (42.40-49.55). For oil content, magnitude of variation was very low but the heritability was high in both the environments and pooled across the environments. Among the oil quality traits, broader range was observed in EII for oleic acid (29.96-65.11), linoleic acid (16.57-47.63) and for O/L ratio (0.63-3.93) hence, the magnitude of variation was also high in that environment compared to EI. The heritability was very high in individual environments compared to pooled data. Frequency distribution of pooled data for all the traits revealed a typical normal distribution indicating their quantitative nature of inheritance. For all the traits, majority of the RILs were within the parental limit and few transgressive segregants were observed in both the directions (Table 3, Fig 1).

Association analysis between protein content and oil content (r=-0.294) revealed a significant negative correlation. Among the oil quality traits, strong negative correlation was observed between oleic and linoleic acids (r=-0.987) and linoleic acid with O/L ratio (r=-0.970). As far as protein content and oil quality are concerned there was a negative correlation between protein content with oleic acid (-0.302) and O/L ratio (-0.350) and positively correlated with linoleic acid (0.316) but the correlation pattern between oil content and oil quality was exactly inverse to that of protein content and oil quality (Table 4).

3.2 Marker polymorphism and linkage analysis

Initially a total of 1043 SSR markers, listed in Table 1, were screened on the parental cultivars (TG 26 and GPBD 4). Out of these, 894 primers produced scorable bands and 53 markers showed polymorphism between the parents. These polymorphic markers were used for generating the marker genotyping data on the population. Of these data, 15 markers, however, showed segregation distortion, which does not fit the 1:1 ratio based on the chi-square analysis. Nevertheless, due to availability of less marker genotyping data, both distorted and non-distorted markers were utilized for linkage analysis. As a result, a total of 45 markers were mapped on eight linkage groups with the total map distance of 657.90cM, with an average inter-marker distance of 14.62 cM. Eight markers, however, remained ungrouped. The length of the linkage group varied from 29.00 cM (LG5) to 145.30 cM (LG1) and the number of markers on each linkage group varied from 4 (LG2, LG6, LG7 and LG8) to 8 marker (LG1){Fig.2}.

3.3 Marker-Trait association (SMA and CIM analyses)

3.3.1 Protein content

A total of seven markers were identified for protein content based on SMA (Table 5) with the phenotypic variance ranging from 2.54 - 9.78%, the highest contribution was from TC6H03 (9.78 %). QTL analyses identified six QTLs (1.50 to 10.70%) for protein content, among them two QTLs *viz*., TC6H03-TC11A04 on LG1 and TC2E05-TC3E2 on LG 4 showed significant contribution (>10.0 R^2) to variance. One QTL near TC6H03 on LG1 had highest contribution as revealed by both single marker analysis (9.78 %) and QTL analysis (10.70 %). The favorable allele for this QTL was contributed by TG 26 and for other QTL, contribution was from GPBD 4 parent (Table 6, Fig 2).

For analyzing QTL x Environment interaction, individual environments data and mean data across the environments was pooled with genotypic and mapping data to identify QTLs across environments for all the above traits. Two QTLs were identified across the environments (also detected in EI) with the phenotypic variance ranged from 6.2 to 8.9 % in the marker interval of TC1D12-pPGSSeq19D6 and TC2E05-TC3E2 but the favorable allele has come from different parents i.e. TG26 and GPBD 4 respectively (Table6, Fig 3).

3.3.2 Oil content

For oil content, a total of four QTLs with the phenotypic variance ranging from 1.5 to 9.1 % were identified from two environments and among them one QTL in the marker interval IPAHM103-PM36 on LG3 was common between environments with the phenotypic variance of 7.1- 9.1 % and SMA also supported this with 5.72 and 6.98 % adjusted R^2 also supported this. One QTL near TC2E05 on LG 4 was common for oil and protein but the direction of favorable allele was different for oil (TG 26) and protein (GPBD 4) with the additive effect of 0.199 and 1.030, respectively. Another minor QTL near TC2B09 was also common between oil (6.80 %) and protein (1.50 %) with TG 26 contributing the favorable allele for both the traits (Table 5 and 6, Fig 2).

Across the environment, three QTLs were detected on LG1, LG3 and LG4 with the phenotypic variance ranging from 2.2 to 10.2 % and the LOD was 3.47 to 5.27. Here also, as in individual environments, the contribution of IPAHM103 was highest and the favorable allele was from GPBD 4 but the contribution of other two QTLs was from TG26 parent (Table 6, Fig 3).

3.3.3. Oil quality

In the present study, all the three QTLs (TC6H03-TC11A04, TC5A07-IPAHM395 and TC3A12-PM433) identified for oil quality traits were common for oleic acid and linoleic acid which is also supported by their strong negative correlation (Table6, Fig 2). Among these, a QTL flanked by TC6H03-TC11A04 had a significant contribution to variance for oleic acid (9.70%), linoleic acid (9.00 %) and O/L ratio (6.80%). GPBD 4 contributed the favorable allele for this QTL for all the three traits. The contribution of other two QTLs (TC5A07-IPAHM395 and TC3A12-PM433) was minor for both oleic (5.60 and 7.20%) and linoleic acids (5.10 and 7.20 %). Single marker analyses identified three common markers for oleic acid, linoleic acid and O/L ratio (PM137, TC6H03 and IPAHM395). In EI, SMA detected six markers each for oleic acid (2.2-3.6), linoleic acid (2.6-5.2) and five markers for O/L ratio (2.86-5.09) (Table 5). No QTLs were identified in EI for oil quality, but when the pooled data was considered, three common QTLs were identified for oleic acid and linoleic acid (PM137-TC6H03, TC5A07-IPAHM395 and TC3A12-PM433 on LG1, 7 and 8 respectively). The phenotypic variance was highest by PM137-TC6H03 for both oleic (6.1 %) and linoleic acid (6.8 %) and the favorable allele was from GPBD 4parent for both. For O/L ratio also the same QTL contributed maximum phenotypic variance (5.1 %) across the environments and the other QTLs were minor QTLs (Table6, Fig 3).

3. Discussion

Traits associated with seed quality are difficult and uneconomic to measure in large segregating generations and are substantially influenced by genotype x environment interaction. Thus, breeding progress in these traits by conventional techniques has had a limited success. Therefore, MAS is highly justified option for indirect selection of these traits in groundnut. Not much information pertaining to the studies on tagging of molecular markers for nutritional quality is available in groundnut. The present study, therefore, was undertaken to identify the SSR markers linked with most important quality traits of groundnut. Although large numbers of SSR markers were used, due to very low polymorphism, only a partial linkage map was constructed. On the other hand, a good variation was observed for the traits examined in the RIL population. Detailed analysis has identified some QTLs for several nutritional quality traits.

4.1 Phenotypic evaluation

The population consisting of 146 RILs showed significant variation among the lines and significant G x E interaction for all the traits indicating the sufficient variation within the population and existence of environmental interaction. The maximum protein content observed in the parent GPBD 4 and in few RILs of the population ($\geq$ 34.0%) can be considered best among the groundnut germplasm for protein content reported till now. The range of protein content in groundnut was earlier reported between 16.0 and 34.0 (Dwivedi *et al*., 1993; Singh *et al*., 1998) hence, these lines may be useful in improving the protein content of groundnut. Although there was not much improvement in the oil content in the population, the parent GPBD 4 was favorable for combination of traits as shown by its high mean value for various traits. The highest oleic acid observed in the population was in the RIL no. 95 (65.90) with highest O/L ratio (4.06) and lowest linoleic acid (16.11) (data not shown). Hence, this line may be useful for improvement of fatty acid/oil quality in the future breeding programmes. Although, many RILs have shown transgressive segregation in positive direction for various traits, none of them had an improved version for all the combination of traits like GPBD 4. Hence, it is difficult to improve all these favorable traits in a single line through conventional breeding approaches, thus justifying the identification of molecular markers for improvement of nutritional quality of groundnut.

A typical normal distribution for all the nutritional traits indicates their quantitative nature of inheritance. The magnitude of variation was lower for protein content and oil content, moderate for oleic acid and linoleic acid,

high for O/L ratio and the heritability was high for all the five traits. Hence, phenotypic selection alone may be effective for these traits but biochemical estimation of fatty acid composition is not economic and mostly quantitatively inherited, hence, identification of efficient markers helpsin marker-assisted introgression.

Antagonistic nature of protein content and oil content as revealed by their negative correlation indicates difficulty in simultaneous improvement of both traits and one can be increased at the expense of the other (Table 4). This relationship is more helpful for developing cultivars for confectionary purpose where high protein but low oil/fat is preferred. However, for developing cultivars for oil content this relationship poses a problem. Negative relationship between oil and protein has also been observed in earlier studies (Kale *et al*., 1998; Parmer *et al*., 2002; Yashoda 2005 and Kavera 2008). Among the oil quality traits, strong negative correlation existed between oleic acid and linoleic acids in the present study, which is in accordance with earlier reports (Sekhon *et al*., 1980; Bovi, 1983; Anderson *et al*., 1998 and Kavera, 2008). Linoleic acid is unstable at higher temperature and has an inverse relationship with oil stability (Braddock *et al*., 1995; O'Keefe *et al*., 1993 and Holley and Hammons 1968). Hence, increased oleic acid normally resulted in reduced linoleic acid, which is desirable from the point of health and oil stability. The correlation for all the traits between the environments was significantly positive (data not shown) indicating their consistent performance across the environments.

From the phenotypic data analysis, it can be concluded that, although heritability was high for the traits in the present study, their negative correlation, quantitative nature of inheritance, G x E interaction and the cost and difficulty involved in phenotyping makes them difficult to improve by conventional breeding techniques.

4.2. Marker polymorphism and a partial genetic map

Construction of genetic linkage map is necessary to apply marker assisted selection tool in crop improvement program. Very few reports on the construction of genetic linkage map based on SSR markers are available in groundnut (Moretzsohn *et al*., 2005; Gobbi *et al*., 2006; Varshney *et al*., 2009; Khedikar *et al.,* 2010, Hong *et al*., 2010). The per cent polymorphism obtained in the present study is very less compared to earlier reports e.g. 23% by Hopkins *et al*. (1999); 33.9% by He *et al*.(2003) (33.90%); 70.80- 81.00% by Ferguson *et al*.(2004); 29.23% by He *et al*.(2005); 52% by Mace *et al*. (2006); 47.10% by Moretzsohn *et al*.(2005); 52.08% by Nimmakayala *et al*. (2007),12.60% by Varshney *et al*. (2009) and 6.15% by Khedikar *et al*. (2010). In general, being a highly self-pollinated plant and its origin by single event hybridization followed by polyploidization, cultivated

peanut exhibits limited polymorphism (Halward *et al.,* 1991; Young *et al.,* 1996). The parents used in developing the mapping population in the present study are only two cultivars and limited polymorphism could be due to narrow genetic base of the parents compared to the reports based on the wider germplasm used in other studies. As a result, even after screening >1000 SSR markers, a partial genetic map with only 45 SSR loci could be developed. Hence, it becomes imperative to select the diverse parents for developing the mapping population so that good genetic maps can be developed.

Some markers (e.g. IPAHM103) identified with QTLs in the present study also showed association with rust resistance QTLs in another study (Khedikar *et al*., 2010). Hence, to know the marker orientation and possible markers used for further saturation of the rust QTL region on LG3 of this population, linkage maps of other reports (Varshney *et al*., 2009; Khedikar *et al*.,2010; Hong *et al*., 2010) were compared with the help of MAPCHART. As a result, four markers (PM183, IPAHM103, seq19D6 and IPAHM272) of LG3 of the present study were found syntenic to LG6 of TAG 24 x GPBD 4 population and LG7, LG5 and LG1 of present study were homologous to LG13, LG12 and LG2 of TAG 24 x GPBD 4 mapping population respectively (Khedikar *et al*., 2010). Similarly, two markers on LG 3 (PM183 and pPGSSeq19D06) of the present study were syntenic to the LG IV of the TAG 24 x ICGV 86031 population (Varshney *et al*., 2009) but the order was inversed. Three markers (PM36, pPGSseq19D6 and TC1D12) of LG19 from composite linkage map of three RIL populations (Yueyou 13 × Zhenzhuhei, Yueyou 13 × J11 population, Yueyou 13 × Fu 95-5) from the report of Hong *et al*. (2010) were also found syntenic to the corresponding markers on LG3 in the present population (Fig 4). In summary, the markers present in the same genomic regions in the other genetic maps, as mentioned above, can be used to saturate the QTL regions of the linkage groups of this study.

4.3 Marker-Trait association

As QTLs identified in one season/location may not express in the other as the expression of QTL is influenced by environments (Khedikar *et al*., 2010), hence, in the present study, QTL analysis was carried out both in individual environments (EI and EII) and pooled across the environments (PE). For protein content, two QTLs contributed substantial phenotypic variance of >10.0 % such as TC2E05-TC3E2 on LG4 in EI and TC6H03-TC11A04 on LG1 in EII. None of the QTLs was common between environments but the two QTLs identified across the environments (PE) for protein content were identified for EI (TC1D12-TC9B08 on LG3 and TC2E05-TC3E02

on LG4) and together both accounted for a total of 14.2% of phenotypic variance in EI and 15.1% in PE. Validating these markers across the locations will have a role in MAS for selection of high protein content varieties stable across the seasons/locations. Seasonal variation for protein content has been reported earlier (Dwivedi *et al*., 1996).

For oil content, a QTL in the marker interval IPAHM103-PM36 on LG3 was common in both individual (EI and EII) as well as PE with the phenotypic variance of 7.1- 10.2%. The position of the QTL was 2.5 cM away from the marker IPAHM103. The same marker has been associated with a major QTL for rust resistance in TAG24 X GPBD 4 mapping population (6.9-55.2 %) (Khedikar *et al*., 2010). For rust resistance, this marker is located at 0.5cM downstream of IPAHM 103 on LG3 in this population. It can be inferred that the QTL identified for oil content and the rust resistance is only 2.0cM away from each other indicating probable pleiotropic effect or tight linkage, which could be confirmed by fine mapping of this region. The incidence of rust was very high in both the seasons in which the oil content was estimated, indicating the impact of rust resistant QTL on oil accumulation, which is also supported by negative correlation between higher incidences of disease with oil content (Sarvamangala and Gowda 2010). Although, identified QTLs for oil content are minor but a QTL near IPAHM103 is consistent and it is also identified as major QTL for rust resistance, so, the use of this marker in MAS could simultaneously improve the oil content while developing the cultivars for rust resistance. It is interesting to note that, the SSR marker PM36 identified for oil content (16.60 %) by Gomez *et al.* (2009) based on bulk segregate analysis and in the present study based on single marker analysis (2.24 %) is linked with IPAHM103 with 8.5cM distance. PM36 is reported to be located on fifth linkage group in the AA genome (Moretzsohn *et al.,* 2005) which indicates the probable location of IPAHM103 on the same linkage group. It has also been suggested in the study of Gomez *et al.* (2009) that because of the availability of less polymorphic markers, BSA allows identification of markers up to 20cM from a gene in either direction.

One QTL in the marker interval of TC2E05-TC3E02 on LG 4 was common for oil content and protein but the direction of favorable allele was different for oil content (TG 26) and protein content (GPBD 4) with the additive effect of 0.199 and 1.030, respectively. Such QTLs can lead to antagonistic relations between the traits as revealed by their negative correlation and it could be important for developing confectionary groundnut where, low oil and high protein are preferred. Another minor QTL (TC2B09-RN16F05) was also common between oil content (6.80 %) and protein content (1.50 %) with TG 26 contributed the favorable allele for both the traits indicating the possible role of genomic region near TC2B09 in reducing the negative correlation between oil and protein. Similar findings

on common QTLs for oil and protein has given by Panthee *et al*. (2005) in soybean population. The favorable positive association of this QTL for oil content and protein content is useful for simultaneous improvement of both the traits. Although these two QTLs were common between protein and oil, the difference in the position of these QTLs on linkage groups once again indicates the necessity of saturation of the linkage map to confirm their significance.

As for the oil quality traits are concerned, all the three QTLs (TC6H03-TC11A04, TC5A07-IPAHM395 and TC3A12-PM433) were common for oleic acid and linoleic acid which is also supported by their strong negative correlation. Among them, a QTL flanked by TC6H03-TC11A04 had a significant contribution to variance for oleic acid (9.70%), linoleic acid (9.00 %) and O/L ratio (6.80%). GPBD 4 contributed the favorable allele for this QTL for all the three traits. The same QTL was also identified for protein with substantial contribution to variance (10.70%) thus revealing its major impact on nutritional quality. The genomic clone contributing the SSR marker TC6H03 was found associated with "Ras related GTP binding protein" in *Arabidopsis* and with GTP-binding protein in Fabaceae (David Bertioli, personal comm.), thus it could be an important candidate gene associated with nutritional traits. Further validation in different genetic backgrounds may prove the efficiency of this QTL. No QTLs were identified in rainy season for the oil quality, indicating their sensitiveness to the environment. But across the environments, one QTL on LG1 (PM137-TC6H03) had a substantial contribution to all the three oil quality traits with the significant phenotypic variance and again the location of this QTL is near to TC6H03 indicating the importance of this marker.

Conclusion

Although the present study is a preliminary study for identification of SSR markers for nutritional quality in groundnut, it was possible to identify few candidate markers for these traits. Assuming the criteria of major QTLs (>10% R^2), the identified QTLs for oil content (IPAHM103) and protein content (TC6H03) are considered as major QTLs. As shown by the frequency distribution, all the above traits are polygenic in nature; hence, the identified QTLs had phenotypic variance of <20.0 %. Validation of the identified QTLs, however, is required in the wider genetic background before they can be recommended to use it in MAS.

Acknowledgments

Authors are thankful to the National Fund of the Basic and Strategic Research in Agriculture of the Indian Council of Agriculture (RKV, MVCG), Department of Biotechnology (DBT) of the Government of India (MVCG) and SP2 Leader Discretionary Grant of Generation Challenge Programme (RKV) for financial support to undertake this study. Thanks are also due to Dr. David J Bertioli, for providing the information on BLAST search for the SSR markers and to Dr Yoginder P Khedikar for his help in experimental work. CS acknowledges Drs. H.L. Nadaf and Kaveri for providing the facility for phenotyping and calibration of NIR for quality traits, to Jawaharlal Nehru Memorial Fund for providing fellowship during the study and two anonymous reviewers for their valuable suggestions.

References

Anderson P.C., Hill, K., Gorbet, D.W., Brodbeck, B.V., 1998. Fatty acid and amino acid profiles of selected peanut cultivars and breeding lines. J. Food Compos. Anal. 11, 100-111.

Badigannavar, A.M., Kale, D.M., Murty, G.S.S., 2002. Assessment of yielding ability of Trombay groundnut varieties through growth analysis. J. oilseeds Res. 19, 38-43.

Bandyopadhyay, A., Desai, S., 2000. Groundnut as a source of oil and food-the facts and national perpespective. In: Nat Sem Oilseeds and oils-Res Develop Needs in the Milleniam Directorate of oilseed Research Hyderabad 2-4 Feb 72-87.

Bovi, M.L.A., 1983. Genotypic and environmental effects on fatty acid composition Iodine value and oil content of peanut (*Arachis hypogaea L.*). Diss Abst Int 44, 406.

Braddock, J.C., Sims, C.A., O' Keefe, S.K., 1995. Flavor and oxidative stability in roasted high oleic peanuts. J. Food Sci. 60, 489-493.

Branch, W.D., Nakayama, I., Chinnan, A., 1990. Fatty acid variation among U S runner-type peanut cultivars. J. Amer. Oilseed Chem. 67, 591-593.

Burrow, M.D., Starr, J.L., Simpson, C.E., Paterson, A.H., 1996. Identification of RAPD markers in peanut (*Arachis hypogaea* L) associated with root-knot nematode resistance derived from *A cardenasii.* Mol. Breed. 2, 307–319.

Chu, Y., Ramos, L., Holbrook, C.C., Ozias-Akins, P., 2007. Frequency of a loss-of-function mutation in Oleaolyl-PC desaturase (ahFAD2A) in the mini-core of the US peanut germplasm collection. Crop Sci. 47, 2372-2378.

Cuc, L.M., Mace, E.S., Crouch, J.H., Quang, V.D., Long, T.D., Varshney, R.K., 2008. Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea* L.). BMC Plant Biol. 8, 55–65

Dwivedi, S.L., Nigam, S.N., Jambunathan, R., Sahravat, K.L, Nagabheshanam, G.V.S., Raghunath, K., 1993. Effect of genotypes and environment on quality parameters and their correlation in peanut. Peanut Sci. 20, 84-89

Dwivedi, S.L., Nigam, S.N., Nageswara Rao,R.C., Singh, U., Rao, K.V.S.,1996. Effect of drought on oil fatty acids and protein contents if groundnut Seeds (*Arachis hypogaea* L). Field Crop Res. 48, 125-123

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

Garcia, G.M., Stalker, H.T., Shroeder, E., Kochert, G.A.,1996. Identification of RAPD SCAR and RFLP markers tightly linked to nematode resistance genes introgressed from *Arachis cardenasii hypogaea*. Genome 39,836–845

Gobbi, A., Teixeira, C., Moretzsohn, M., Guimaraes, P., Leal-Bertioli, S., Bertioli, D., Lopes, C.R., Gimenes, M, 2006. Development of a-linkage map to species of B genome related to the peanut (*Arachis hypogae*) In : Plant and animal genomes XIV conference, San Diego, CA, USA. P679.

Gomez, M.S., Manivannan, N., Schubert, M.A., Ayers, J.L., Baring, M.R., Burow, M.D., 2009. Identification of QTLs for pod and kernel traits in cultivated peanut by bulked segregant analysis. Electronic J. Biotech. 12(2).

Gowda, M.V.C., Motagi, B.N., Naidu, G.K., Diddimani, S.B., Sheshagiri, R., 2002. GPBD 4: A Spanish bunch groundnut genotype resistant to rust and late leaf spot. Int. Arachis Newsletter 22, 29-32

Groff, J.L., Gropper, S.S., Hunt, S.M., 1996. Lipids: In Advanced Nutrition and Human metabolism; West Publishing; Minneapolis/St Paul MN 113-146

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

Halward, T.M., Stalker, H.T., Larue, E.A. Kochert, G., 1991. Genetic variation detectable with molecular markers among unadapted germplasm resources of cultivated peanut and related wild species. Genome 34, 1013-1020.

He, G., Meng, R., Newman, M., Gao, G., Pittman, R.N., Prakash, C.S., 2003. Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.). BMC Plant Biol. 3, 3-9.

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

Herselman, L., Thwaites, R., Kimmins, F.M., Courtois, B.V., Ander Merwe, P.J.A., Seal, S.E., 2004. Identification and mapping of AFLP markers linked to peanut (*Arachis hypogaea*L) resistance to the aphid vector of groundnut rosette disease. Theor. Appl. Genet. 109. 1426–1433

Holley, K.T., Hammons, R.O., 1968. Strain and Environmental effects on peanut characteristics Research Bulletin 32 University Georgia College Agriculture Experiment Station pp 1-27

Hong ,Y., Xiaoping, C., Liang, X., Haiyan, L., Guiyuan, Z., Shaoxiong, L., Shijie, W., Holbrook, C., Guo, B., 2010. A SSR-based composite genetic linkage map for the cultivated peanut (*Arachis hypogaea* L.) genome. BMC Plant Biol. 10, 17

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

Jansen, R.C., Stam, P., 1994. High resolution of quantitative traits into multiple loci via interval mapping. Genetics 136, 1445-1447.

Jung, S., Powell, G., Moore, K., Abbot, A., 2000.The high oleate trait in the cultivated peanut (*Arachis hypogaea* L) II Molecular basis and genetics of the trait. *Mol. Genet.* 263, 806-811

Kale, D.M., Gadgil, J.D., Murthy, G.S.S., 1998. Relationship between seed size oil and protein contents in groundnut. J. Oilseeds Res. 5, 170-175

Kale, D.M., Mouli, C., Murthy, G.S.S., Rao, M.V.P., 1997. Development of a new groundnut variety TG 26 by using induced mutations in cross breeding. Mutat. Breed. Newslet. 43, 25-27

Kavera, B., 2008. Oil quality improvement in Groundnut (*Arachis hypogaea* L.) through induced mutagenesis. PhD Thesis Univ. Agric Sci Dharwad (India)

Khedikar, Y.P., Gowda, M.V.C., Sarvamangala, C., Patgar, K.V., Upadhyaya, H.D., Varshney, R.K., 2010. 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. DOI 101007/s00122-010-1366-x

Kosambi, D.D., 1944. The estimation of map distances from recombination values. Ann. Eugen. 12,172–175

Kris-Eterton, M.P., Pearson, A.T., Ying-Wan Hargrove, I.R., Moriarty, K., Fishell, V., Etherton, D.T., 2001. High monounsaturated fatty acid diets lower both plasma cholesterol and triacylglycerol concentrations. Amer. Soc. Clin. Nutr. 70, 1009-1015

Lander, E.S., Green, P., Abrahamson, J., Barlow, A., Daly, M.J., Lincoln, S.E., Newburn, L., 1987. MAPMAKER : An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1, 174–181

Lopez, Y., Nadaf, H.L., Smith, O.D., Connell, J.P., 2000. Isolations and characterization of the Δ^{12} fatty acid desaturase in peanut (*Arachis hypogaea* L.) and search for polymorphism for the high oleate trait in Spanish market-type lines. Theor. Appl. Genet. 101, 1131-1138

Mace, E.S., Phong, D.T., Upadhyaya, H.D., Chandra, S. Crouch, J.H., 2006. SSR analysis of cultivated groundnut (*Arachis hypogaea* L.) germplasm resistant to rust and late leaf spot diseases. Euphytica 152, 317–330

Miller, J.F., Zimmerman, D.C., Vick, B.A., 1987. Genetic control of high oleic acid content in sunflower oil. Crop Sci. 27, 923-926

Misra,, J.B., 2006. Nutritive value of groundnut and composition of Indian groundnut cultivars In: Groundnut Research in India Ed Basu M.S. and Sing N. B. National research centre for groundnut Junagarh India Pp273-291

Mondal, S., Badigannavar, A.M., Murthy, G.S.S., 2007. RAPD markers linked to a rust resistance gene in cultivated groundnut (*Arachis hypogaea* L). Euphytica 159, 233-239

Moretzsohn, M.C., Leoi, L., Proite, K., Guimara, P.M., Leal-Bertioli, S.C.M., Gimenes, M.A., Martins, W.S., Valls, J.F.M., Grattapaglia, D., Bertioli, D.A.J., 2005. Microsatellite-based gene-rich linkage map for the AA genome of *Arachis* (Fabaceae). Theor. Appl. Genet. 111, 1060-1071

Moretzsohn, M.C., Hopkins, M.S., Mitchell, S.E., Kresovich, S., Valls, J.F.M., Ferreira, M.E., 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.

Nimmakayala, P., Jeong, J., Asturi, S.R., Thomason, Y., Tallury, S., Reddy, U.K., 2007. Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on AFLP and microsatellite polymorphisms. In: Plant and animal genomes XV conference, San Diego, CA, USA. P 427.

Norden, A.J., Gorbet, D.W., Knauft, D.A., Young, C.T.,1987. Variability in oil quality among peanut genotypes in Florida breeding program. Peanut Sci. 14, 7-11

O' Keefe, S.F., Wiley, V.A., Knauft, D.A., 1993. Comparison of oxidation stability of high and normal oleic peanut oils. J. Amer. Oilseed Chem. 70, 489-492

Panthee, D.R., Pantlone, V.R., West, D.R., Saxton, A.M., Sams, C.E., 2005. Quantitative trait loci for seed protein and oil concentration and seed size in soybean. Crop Sci. 45, 2015-2022

Parmer, D.L., Ratna Kumar, Bharodia, P.S., 2002. Genetics and interrelationship of oil and protein content in crosses involving confectionary genotypes of groundnut (*Arachis hypogaea* L). J. Oilseeds Res. 19, 32-34

Payne, R.W., Murray, D.A., Harding, S.A.,Baird, D.B., Soutar, D.M., 2007. Gen stat for Windows (10^{th} edition) Introduction. VSN Interrnational, Hemel Hepstead. ISBN 1-904375-32-4.

Ravi, K., Vadez, V., Isobe S., Mir, R.R., Guo Y., Nigam, S.N., Gowda, M.V.C., Radhakrishnan, T., Bertioli, D.J., Knapp, S.J., Varshney, R.K., 2010. Identification of several small main-effect QTLs and a large number of epistatic QTLs for drought tolerance related traits in groundnut (*Arachis hypogaea* L.). Theor. Appl. Genet. DOI 10.1007/s00122-010-1517-0

Ray, T.K., Holly, S.P., Knauft, D.A., Abbott, A.G., Powell, G.L., 1993. The primary defect in developing seed from the high oleate variety of peanut (*Arachis hypogaea* L.) is the absence of Δ12-desaturase activity. Plant Sci. 91, 15-21

Sarvamangala, C., Gowda, M.V.C., 2010. Phenotypic evaluation of RILs for rust and late leaf spot and their association on productivity traits in groundnut. J. Oilseed Res. 26(2), 151-153

Sekhon, K.S., Gupta, S.K., Ahuja, K.L., Jaswal, S.V., 1980. Variability in fatty acid composition in semi-spreading peanut types. Oleagineux 35, 409-411

Singh, A.K., Smartt, J., Simpson, C.E., Raina, S.N., 1998., Genetic variation vis-à-vis molecular polymorphism in groundnut, *Arachis hypogaea* L. Genet. Resour. Crop Evol. 45, 119–126.

Upadhyaya, H.D., Mallikarjunaswamy, B.P., Kenchnagoudar, P.V., Kulliswamy, B.Y., 2005. Identification of diverse groundnut germaplasm through multi–environment evaluation of a core collection for Asia. Field Crop Res. 93, 293-299

Utz, H.F., Melchinger, A.E.,1996. PLABQTL: A program for composite interval mapping of QTL Quantitative Trait Loci On-line publication: vol 2 paper 1 Sponsored by the Crop Science Society of America

Varma, T.S.N., Dwivedi, S.L, Pande, S.Gowda, M.V.C., 2005. SSR markers associated with resistance to rust (*Puccinia arachidis* Speg) in groundnut (*Arachis hypogaea* L.) SABRAO J. Breed. Genet. 37, 107-119

Varshney, R.K., Bertioli, D.J., Moretzsohn, M.C., Vadez, V., Krishnamurthy, L., Aruna, R., Nigam, S.N., Moss, B.J., Seetha, K., Ravi, K., He, G., Knapp, S.J., Hoisington, D.A., 2009. The first SSR-based genetic linkage map for cultivated groundnut (*Arachis hypogaea*L). Theor. Appl. Genet. 118, 729–739

Voorrips, R.E., 2006. MapChart 22 : Software for the graphical presentation of linkage maps and QTLs Plant Research International Wageningen The Netherlands

Yashoda, M.H.,2005. Evaluation of bold seeded groundnuts for confectionary characteristics in groundnut. Dessertion: MSc (Agri) Thesis UAS Dharwad

Yong, L., Boshou, L., Wang, S.Y., Dong, L., Jiang, H., 2005. Identification of AFLP markers for resistance to seed infection by *Aspergillus flavus* in peanut (*Arachis hypogaea* L). Acta. Agron. Sinica. 31, 1349-1353

Young, N.D., Weeden, N.F., Kochert, G., 1996. Genome mapping in legumes. In: Genome mapping in plants (Eds.) Paterson, A. H., Landes, R. G., Austin, T. X.

Zeng, Z.B.,1994. Precision mapping of quantitative trait loci. Genetics 136, 1457-1468.

Table 1: Summary on marker polymorphism study on TG 26 x GPBD 4 population

Source	No. of markers screened	No of polymorphic markers identified
Hopkins *et al*.(1999)	26	0
He *et al*.(2003)	158	4
Ferguson *et al*. (2004)	226	10
Moretzshon *et al*.(2004, 2005)	338	20
Mace *et al*. (2006)	79	0
Proite et al. (2007)	46	3
Cuc *et al*. (2008)	170	16
Total	**1043**	**53**

Table 2: Pooled ANOVA for protein content, oil content, oleic acid, linoleic acid and O/L ratio in TG 26 x GPBD 4 mapping population

Traits/Source of variation	Mean sum of squares					
	Environment	**Replication**	**E x R**	**Genotypes**	**E X G**	**Error**
Df	1	1	1	145	145	290
Protein content (%)	65.31**	2.5	2.19	26.69**	13.66**	1.45
Oil content(%)	38.87**	0.63	0.06	9.23**	5.96**	1.18
Oleic acid	292.63**	0.38	0.63	82.84**	35.81**	2.13
Linoleic acid	1678.63**	0.06	0.06	59.68**	27.03**	1.72
O/L ratio	3.36**	0	0.01	0.01*	0.26**	0.01

** indicates significance at 1% level of probability, df- Degrees of Freedom, ExR-Environment x Replication, E x G-Environment x Genotype

Table 3: Mean, range and variability components in individual and pooled across two environments for protein content, oil content, oleic acid, linoleic acid and O/L ratio in TG26 X GPBD 4 mapping population

Traits	Environment	Parental means		Recombinant Inbred Lines				
		TG26	GPBD 4	Mean	Range	PCV	GCV	h^2
Protein content (%)	*Kharif* 2007 (EI)	26.03	34.79	30.12	24.02-36.64	8.76	8.23	88.3
	Summer 2007 (EII)	23.96	32.57	29.45	21.12-37.51	11.05	9.9	80.3
	Pooled (PE)	24.99	33.68	29.79	23.44-34.50	6.35	4.92	49.1
Oil content (%)	*Kharif* 2007 (EI)	44.17	49.68	45.31	40.76-49.03	3.14	2.67	72.6
	Summer 2007 (EII)	43.59	46.27	46.02	42.40-49.55	2.59	2.04	62.2
	Pooled (PE)	43.88	47.98	45.67	43.11-48.00	2.88	1.62	66.9
Oleic acid	*Kharif* 2007 (EI)	40.15	51.35	46.88	34.68-59.98	9.08	8.74	92.6
	Summer 2007 (EII)	33.99	51.94	45.51	29.96-65.11	14.48	13.99	93.2
	Pooled (PE)	37.07	51.65	46.17	33.37-58.88	8.07	7.43	84.6
Linoleic acid	*Kharif* 2007 (EI)	38.03	28.98	30.41	19.72-40.56	11.73	11.24	91.9
	Summer 2007 (EII)	44.76	28.83	33.8	16.57-47.63	16.83	16.19	92.6
	Pooled (PE)	41.39	28.90	32.11	21.25-42.81	9.79	8.9	82.6
O/L ratio	*Kharif* 2007 (EI)	1.06	1.77	1.58	0.85-3.04	22.05	21.28	93.2
	Summer 2007 (EII)	0.76	1.80	1.43	0.63-3.93	37.07	35.78	93.1
	Pooled (PE)	0.91	1.79	1.51	0.78-2.98	18.78	17.08	82.7

PCV-phenotypic coefficient of variation, GCV-Genotypic coefficient of variation, h^2-broad sense heritability

Table 4: Correlation coefficients for protein content, oil content and oil quality traits in the mapping population pooled over two environments

Traits	Protein content	Oil content	Oleic acid	Linoleic acid	O/L ratio
Protein content	1.000				
Oil content	-0.294**	1.000			
Oleic acid	-0.302**	0.334**	1.000		
Linoleic acid	0.316**	-0.311**	-0.987**	1.000	
O/L ratio	-0.350**	0.298**	0.983**	-0.970**	1.000

**significant 1% level of probability respectively, O/L-Oleic/Linoleic acid, -indicates negative correlation

Table 5: Marker-trait association using single marker analysis (SMA)

Environments	*Kharif* 2007 (EI)		Summer 2007 (EII)	
Traits	**Marker**	**adjR^2**	**Marker**	**adjR^2**
Protein content (%)	TC1D12	5.42**	TC1D12	2.54*
	TC1B04	4.63**	TC6H03	9.78**
	TC3E2	5.87**	TC1B02	8.05**
	TC2C07	5.88**	PM36	4.31**
Oil content (%)	TC3A12	2.87*	IPAHM475	2.77*
	TC2B09	2.18*	pPGSSeq7H6	3.26*
	TC6E01	2.56*	TC3A12	2.44*
	TC2C07	6.43**	PM36	2.24*
	IPAHM103	5.72**	TC2B09	4.18*
			IPAHM103	6.98**
Oleic acid	PM137	2.24*	PM137	3.46*
	IPAHM176	2.88*	TC6H03	3.42*
	pPGSSeq11G7	2.48*	IPAHM395	3.54*
	IPAHM295	2.2*		
	IPAHM395	3.1*		
	TC5A07	3.6*		
Linoleic acid	TC3E05	5.2**	PM137	3.81*
	IPAHM176	2.60*	TC6H03	3.5*
	pPGSSeq 11G7	2.62*	IPAHM395	3.29*
	IPAHM295	2.61*		
	IPAHM395	3.59**		
	TC5A07	3.36**		
O/L Ratio	TC3E05	5.09**	PM137	2.89*
	IPAHM176	3.37*	TC11A04	2.37*
	pPGSSeq 11G7	3.09*	IPAHM176	2.33*
	IPAHM395	2.86*	TC6H03	3.22*
	TC5A07	4.07*	IPAHM395	2.17*
			pPGSSeq 17F06	2.74*
			TC2B09	2.53*

*, ** Significance at 5% and 1% level of probability respectively

Table 6: Features of QTLs identified through composite interval mapping (CIM)

Traits	Environments	LG	Marker interval	Position (cM)	LOD	R^2 (%)	Additive effect*
Protein content	EI	3	TC1D12-TC9B08	36	3.33	4.0	-0.594
		4	TC2E05-TC3E02	56	3.62	10.2	1.03
		7	IPAHM395-TC2C07	110	2.87	3.9	-0.609
		8	TC2B09-RN16F05	24	2.89	1.5	-0.552
	EII	1	TC6H03-TC11A04	12	3.42	10.7	-1.249
		6	pPGSSeq15C12-IPAHM105	28	3.04	7.1	1.053
	PE	3	TC1D12-TC9B08	36	2.99	8.9	-0.67
		4	TC2E05-TC3E02	52	2.95	6.2	0.684
Oil content	EI	3	IPAHM103-PM36	28	3.38	7.9	0.499
		4	TC2E05-TC3E02	42	3.01	1.5	-0.199
	EII	1	pPGSSeq7H6-IPAHM475	80	3.2	5.2	-0.434
		3	IPAHM103-PM36	28	3.53	9.1	0.408
		8	TC2B09-RN16F05	6	3.12	6.8	-0.28
	PE	1	pPGSSeq7H6-IPAHM475	62	3.49	2.3	-0.152
		3	IPAHM103-PM36	28	5.27	10.2	0.428
		4	TC2E05- TC3E02	42	3.47	2.2	-0.18
Oleic acid	EI	-	-	-	-	-	-
	EII	1	TC6H03-TC11A04	14	3.75	9.7	2.749
		7	TC5A07-IPAHM395	36	4.32	5.6	1.799
		8	TC3A12-PM433	4	3.6	7.2	1.885
	PE	1	PM137-TC6H03	4	2.74	6.1	1.423
		7	TC5A07-IPAHM395	24	2.65	3.8	1.363
		8	TC3A12-PM433	4	2.8	3.3	0.905
Linoleic acid	EI	-	-	-	-	-	-
	EII	1	TC6H03-TC11A04	14	3.04	9.0	-2.28
		7	TC5A07-IPAHM395	32	4.84	5.1	-1.665
		8	TC3A12-PM433	4	5.06	7.2	-1.641
	PE	1	PM137-TC6H03	4	3.04	6.8	-1.271
		7	TC5A07-IPAHM395	24	2.57	3.6	-1.119
		8	TC3A12-PM433	4	2.61	3.3	-0.759
O/L ratio	EI	-	-	-	-	-	-
	EII	1	TC6H03-TC11A04	14	3.48	6.8	0.192
	PE	1	PM137-TC6H03	4	2.54	5.1	0.11
		4	IPAHM171c-IPAHM352	58	2.89	1.9	-0.052
		8	TC3A12-PM433	4	2.57	1.4	0.05

*negative value indicates favorable allele from TG 26 and positive value indicates favorable allele from GPBD 4 parent, LG-Linkage group, LOD-Log of odds, R^2-Phenotypic variance, cM- centi Morgan, EI-*Kharif* 2007, EII-Summer 2007, PE-pooled across environments

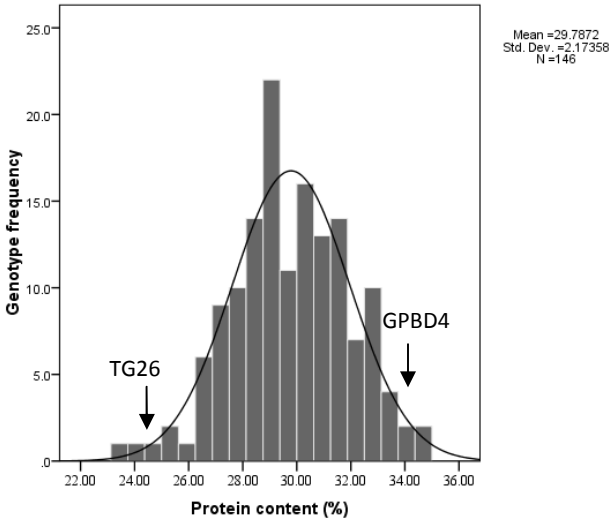
Figure legends

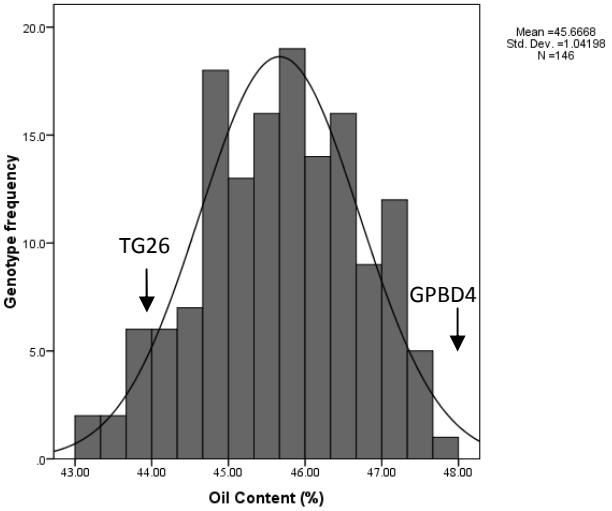
Figure 1: The normal distribution curve for protein content, oil content, oleic acid, linoleic acid and O/L ration in 146 RILs of TG 26 x GPBD 4 population pooled over two environments (EI and EII). Arrows represent the position of parental types (TG26 and GPBD4).

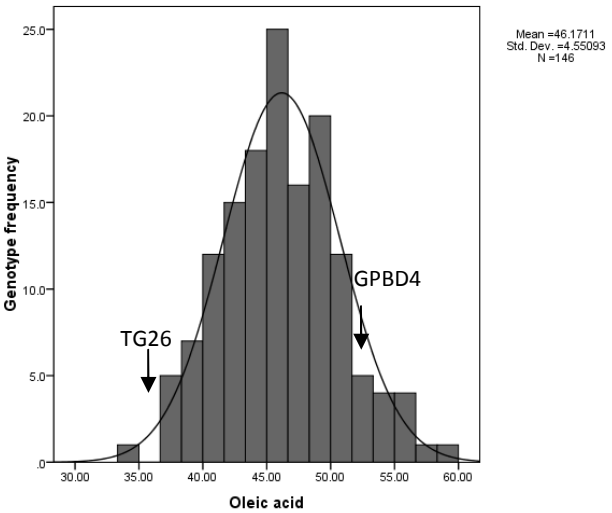
Figure 2: Linkage map depicting QTLs for protein content, oil content, oleic acid, linoleic acid and O/L ratio identified in individual environments (EI and EII) in TG26 x GPBD 4 mapping population. Vertical bars adjacent to corresponding linkage groups (LG) with annotations of traits name and respective environments in brackets represent QTLs.

Figure 3: QTLs identified across environments (PE) for protein content, oil content, oleic acid, linoleic acid and O/L ratio in TG26 x GPBD4 mapping population of groundnut. Vertical bars with diagonal lines inside indicates the QTLs identified across environments.

Figure 4. Comparison of LG3 of TG26 x GPBD4 population and LG19 of composite linkage map of Hong *et al*., 2010 and LG3 of TG26 x GPBD4 aligned through common markers of LG_AhIV of TAG 24 x ICGV86039 tetraploid reference map of Varshney *et al*., 2009


Mean =29.7872
Std. Dev. =2.17358
N =146
Genotype frequency
25.0
20.0
15.0
10.0
5.0
.0
22.00
24.00
26.00
28.00
30.00
32.00
34.00
36.00
Protein content (%)
TG26
GPBD4


Mean =45.6668
Std. Dev. =1.04198
N =146
Genotype frequency
20.0
15.0
10.0
5.0
.0
43.00
44.00
45.00
46.00
47.00
48.00
Oil Content (%)
TG26
GPBD4


Mean =46.1711
Std. Dev. =4.55093
N =146
Genotype frequency
25.0
20.0
15.0
10.0
5.0
.0
30.00
35.00
40.00
45.00
50.00
55.00
60.00
Oleic acid
TG26
GPBD4

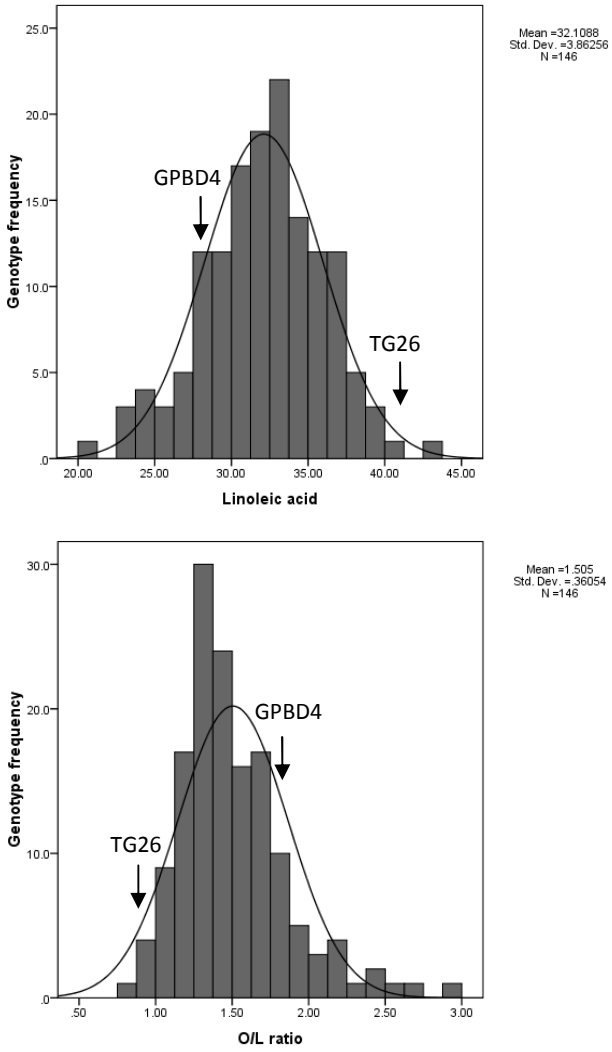

Mean =32.1088
Std. Dev. =3.86256
N =146
GPBD4
TG26
Genotype frequency
Linoleic acid
Mean =1.505
Std. Dev. =.36054
N =146
GPBD4
TG26
Genotype frequency
O/L ratio

Figure 1:

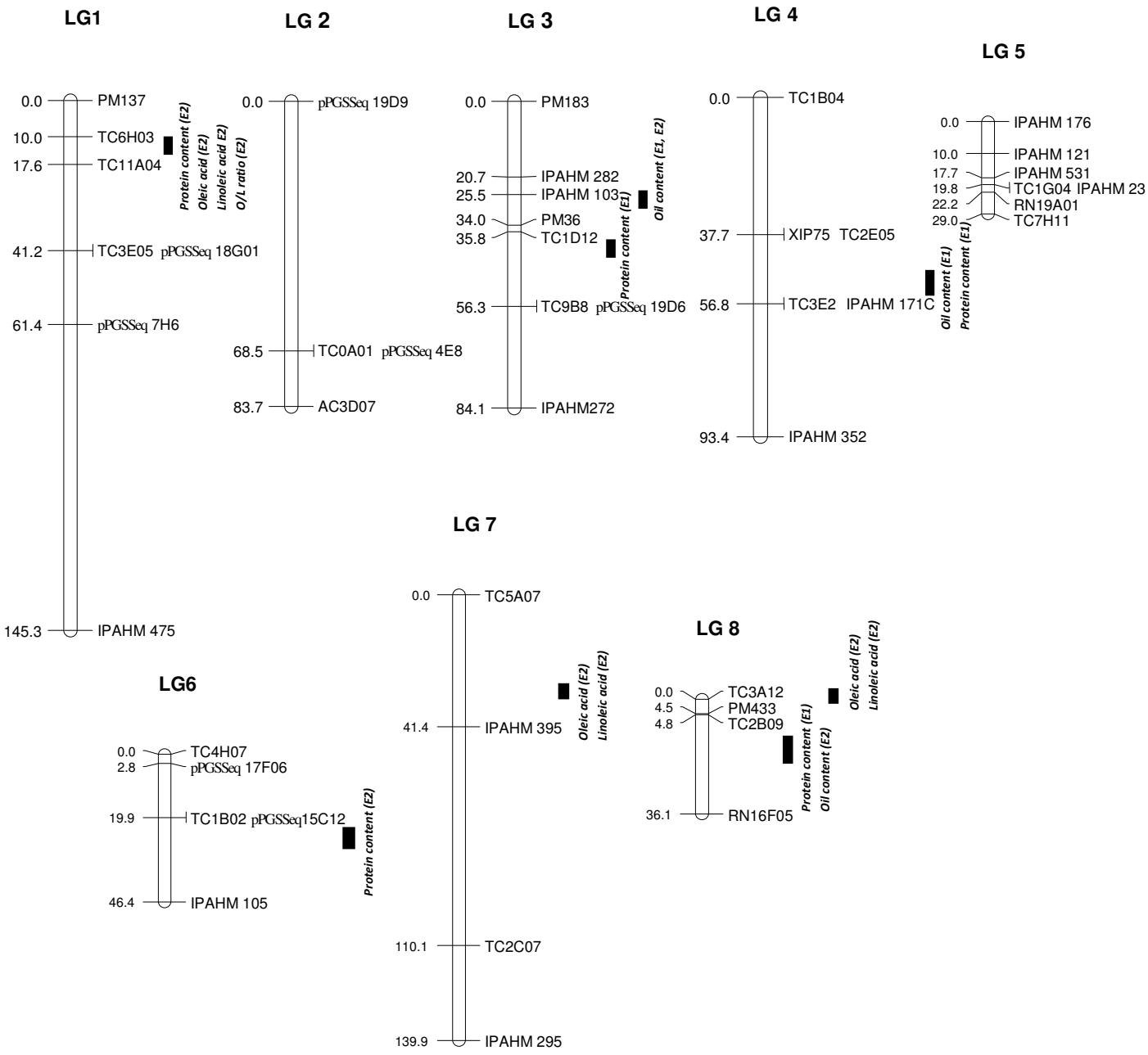

LG1
0.0 PM137
10.0 TC6H03
17.6 TC11A04
41.2 TC3E05 pPGSSeq 18G01
61.4 pPGSSeq 7H6
145.3 IPAHM 475
Protein content (E2)
Oleic acid (E2)
Linoleic acid E2)
O/L ratio (E2)
LG 2
0.0 pPGSSeq 19D9
68.5 TC0A01 pPGSSeq 4E8
83.7 AC3D07
LG 3
0.0 PM183
20.7 IPAHM 282
25.5 IPAHM 103
34.0 PM36
35.8 TC1D12
56.3 TC9B8 pPGSSeq 19D6
84.1 IPAHM272
Oil content (E1, E2)
Protein content (E1)
LG 4
0.0 TC1B04
37.7 XIP75 TC2E05
56.8 TC3E2 IPAHM 171C
93.4 IPAHM 352
LG 5
0.0 IPAHM 176
10.0 IPAHM 121
17.7 IPAHM 531
19.8 TC1G04 IPAHM 23
22.2 RN19A01
29.0 TC7H11
Oil content (E1)
Protein content (E1)
LG 7
0.0 TC5A07
41.4 IPAHM 395
110.1 TC2C07
139.9 IPAHM 295
Oleic acid (E2)
Linoleic acid (E2)
LG 8
0.0 TC3A12
4.5 PM433
4.8 TC2B09
36.1 RN16F05
Protein content (E1)
Oil content (E2)
Oleic acid (E2)
Linoleic acid (E2)
LG6
0.0 TC4H07
2.8 pPGSSeq 17F06
19.9 TC1B02 pPGSSeq15C12
46.4 IPAHM 105
Protein content (E2)

Figure 2:

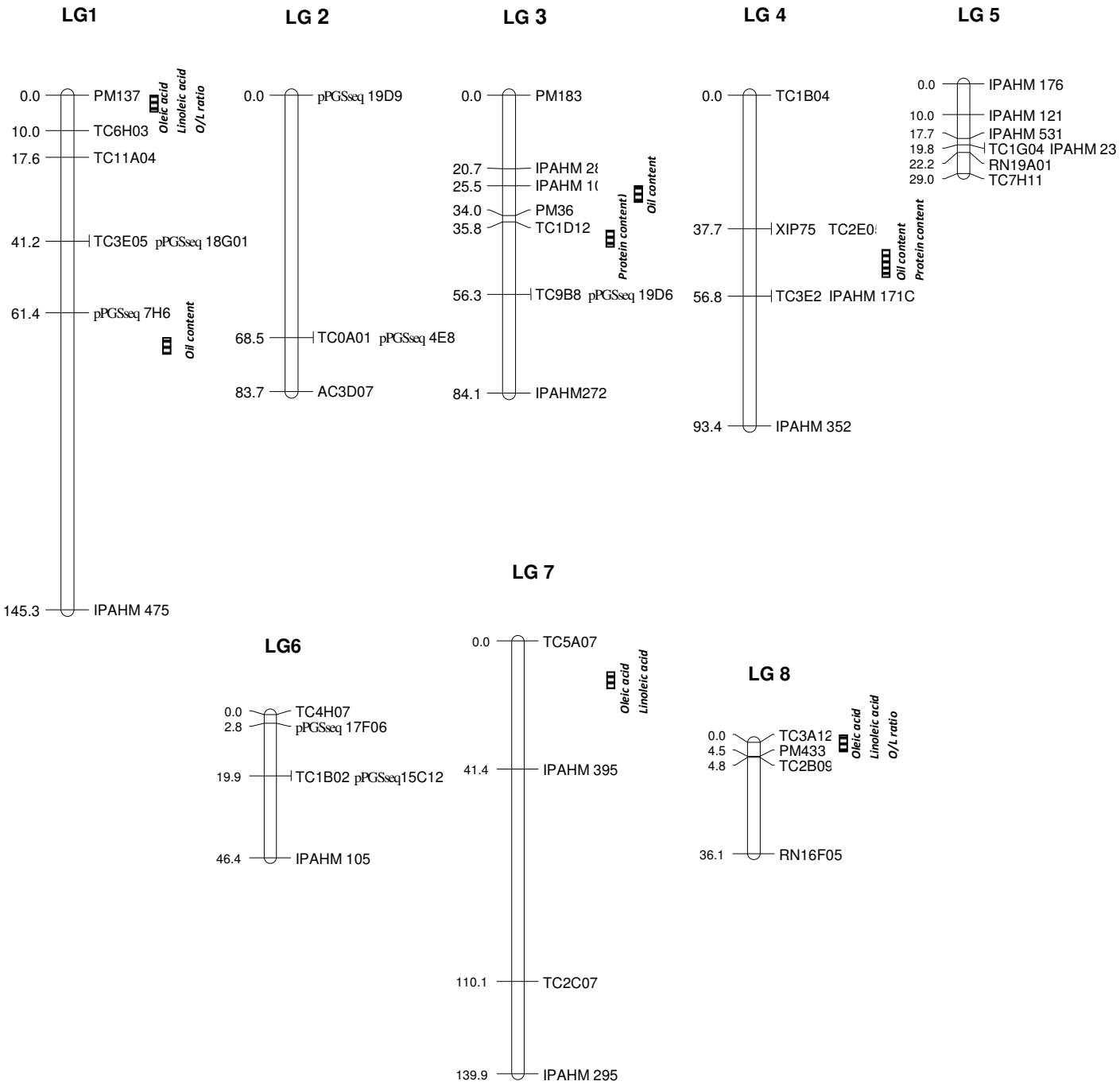

LG1
0.0 PM137
10.0 TC6H03
17.6 TC11A04
41.2 TC3E05 pPGSseq 18G01
61.4 pPGSseq 7H6
145.3 IPAHM 475
Oleic acid
Linoleic acid
O/L ratio
Oil content
LG 2
0.0 pPGSseq 19D9
68.5 TC0A01 pPGSseq 4E8
83.7 AC3D07
LG 3
0.0 PM183
20.7 IPAHM 2
25.5 IPAHM 1
34.0 PM36
35.8 TC1D12
56.3 TC9B8 pPGSseq 19D6
84.1 IPAHM272
Protein content)
Oil content
LG 4
0.0 TC1B04
37.7 XIP75 TC2E0
56.8 TC3E2 IPAHM 171C
93.4 IPAHM 352
Oil content
Protein content
LG 5
0.0 IPAHM 176
10.0 IPAHM 121
17.7 IPAHM 531
19.8 TC1G04 IPAHM 23
22.2 RN19A01
29.0 TC7H11
LG 7
0.0 TC5A07
41.4 IPAHM 395
110.1 TC2C07
139.9 IPAHM 295
Oleic acid
Linoleic acid
LG6
0.0 TC4H07
2.8 pPGSseq 17F06
19.9 TC1B02 pPGSseq15C12
46.4 IPAHM 105
LG 8
0.0 TC3A12
4.5 PM433
4.8 TC2B09
36.1 RN16F05
Oleic acid
Linoleic acid
O/L ratio

Figure 3:

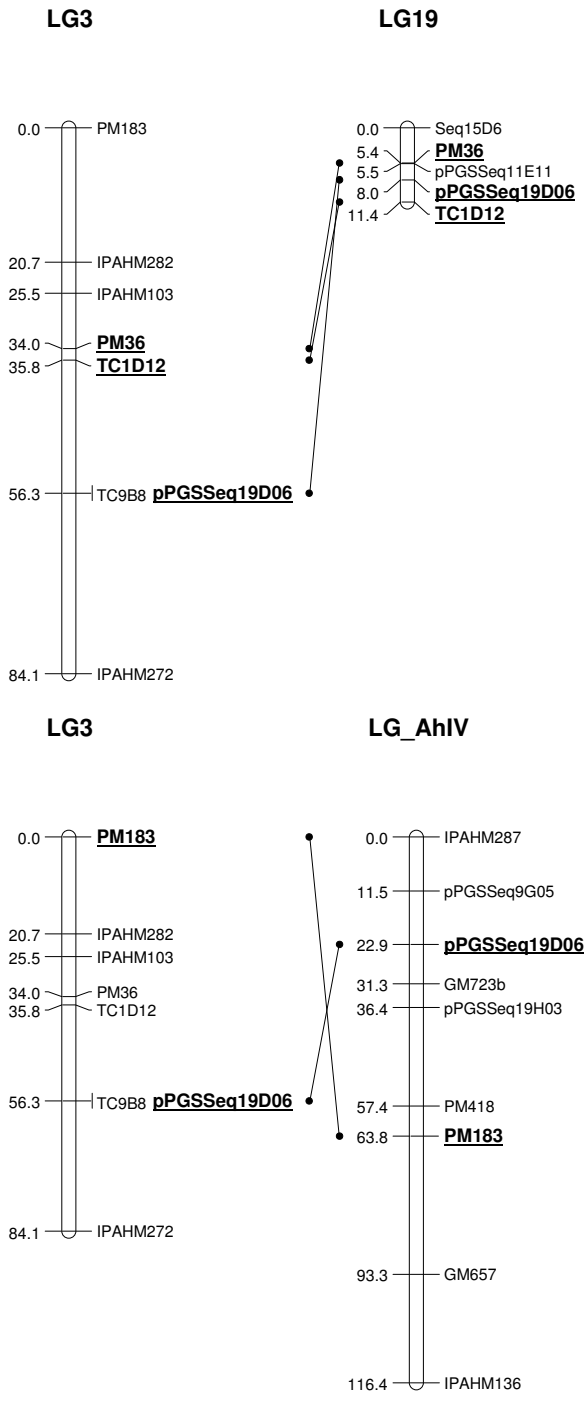

LG3
LG19
0.0 PM183
20.7 IPAHM282
25.5 IPAHM103
34.0 PM36
35.8 TC1D12
56.3 TC9B8 pPGSSeq19D06
84.1 IPAHM272
0.0 Seq15D6
5.4 PM36
5.5 pPGSSeq11E11
8.0 pPGSSeq19D06
11.4 TC1D12
LG3
LG_AhIV
0.0 PM183
20.7 IPAHM282
25.5 IPAHM103
34.0 PM36
35.8 TC1D12
56.3 TC9B8 pPGSSeq19D06
84.1 IPAHM272
0.0 IPAHM287
11.5 pPGSSeq9G05
22.9 pPGSSeq19D06
31.3 GM723b
36.4 pPGSSeq19H03
57.4 PM418
63.8 PM183
93.3 GM657
116.4 IPAHM136

Figure 4: