Identification of an effective genotyping assay for marker assisted selection of high oleic acid content trait in groundnut (*Arachis hypogaea* L.)

S. R. Mythili¹, N. Manivannan¹*, T. Kalaimagal², V. Paranidharan³ and P. Janaki⁴

¹Centre of Excellence in Molecular Breeding,  
²Department of Oilseeds, Centre for Plant Breeding and Genetics,  
³Department of Plant Pathology, Centre for Plant Protection Studies,  
⁴Department of Soil Science and Agricultural Chemistry,  
Tamil Nadu Agricultural University, Coimbatore - 641003, India  
*E-Mail: nm68@tnau.ac.in

**Abstract**  
The substitution and insertion mutation in the fatty acid desaturase gene is responsible for increased oleic acid in groundnut. It can be identified by various genotyping assays with allele specific primers. Present study compares the efficiency of three earlier reported genotyping assays to identify the effective and accurate method. A high yielding groundnut variety GJG 33 was crossed with high oleic acid content variety Girnar 4. Both parents and 23 BC<sub>2</sub>F<sub>2</sub> plants were used for genotyping. Allele specific primers assay proposed by Chen et al. (2010), the CAPS marker assay proposed by Chu et al. (2007 and 2009) and the primers assay proposed by Yu et al. (2013) were used for comparisons. The primer pairs reported by Yu et al. (2013) was found to be efficient in terms of time and cost. Hence these primers may be recommended for utilization in marker assisted breeding to evolve high oleic acid content groundnut genotypes.

**Keywords:** Groundnut; high oleic acid content; mutant allele; wild allele; Marker assisted selection.

**INTRODUCTION**  
Groundnut called as the “poor man's almond” contains 44 – 56 % of protein, 22 -30 % of oil, vitamins and minerals (Mukri et al., 2014). Nearly 90 % of groundnut oil is composed of the three major fatty acids such as oleic, linoleic and palmitic acid (Gangadhara and Nadaf, 2018). High oleic groundnut oil provides an increased shelf life of the oil combined with abundant health benefits. It helps reduce the serum cholesterol, guards against atherosclerosis, minimizes the occurrence of Coronary Heart Disease (Fraser et al., 1992). It also prevents the risk of cancer, reduces blood glucose level (Vassiliou et al., 2009). The ∆¹² fatty acid desaturase enzyme catalyses the conversion of oleic acid to linoleic acid. This enzyme is coded by the two homoeologous genes (FAD2A and FAD2B) located in the A and B genomes of tetraploid groundnut. Mutations within the FAD gene have resulted in the reduction of desaturase activity which in turn increases the accumulation of oleic acid content. The high oleic acid trait is governed by two recessive mutations. The key mutations within the fatty acid desaturase gene are 1bp substitution of G:C to A:T in A genome at position 448 after the start codon lead to missense amino-acid substitution. Another mutation 1bp insertion of A:T at position 442 after the start codon has resulted in the frame shift in the B genome (Jung et al., 2000; Lopez et al., 2000).
A simple allele specific PCR method was developed by Chen et al. (2010) to detect these mutations. In this method, a common forward primer with different allele specific reverse primers was used. The F435SUB-R primer was used to amplify the substitution mutation in the A genome at 203bp. Primer F435INS-R was utilized to detect the insertion mutation in the B genome at 195bp. This assay detects the presence of mutant alleles irrespective of their heterozygosity. CAPS marker assay for A and B genomes were developed by Chu et al. (2007) and Chu et al. (2009) respectively to differentiate homozygous and heterozygous plants with mutant alleles. Another allele specific genotyping method was developed by Yu et al. (2013) in which four reactions with four different primer combinations were used to identify the mutant lines. In this assay, reactions I and III are useful to identify wild alleles in A and B genomes respectively. Likewise, reactions II and IV are useful to identify mutant alleles in A and B genomes respectively. In this study, the efficacy of these three genotyping assays were compared to identify the effective and accurate method for utilization in marker assisted selection of high oleic lines in groundnut.

**MATERIALS AND METHODS**

Plant Material: A high yielding *Virginia* bunch groundnut variety with high oleic acid (70-80%) content Girnar 4 was used as the donor parent which was released from the ICAR-Directorate of Groundnut Research, Junagadh, India. A high yielding *Spanish* bunch groundnut variety with high oil content (50-52%), GJG 33 was used as a recipient parent. It was released by the Junagadh Agricultural University, Junagadh, Gujarat, India. Crosses were made to transfer the high oleic acid trait from Girnar 4 to GJG 33 during Dec 2020-March 2021. F$_1$ s were evaluated during July-Oct 2021 and the identified F$_1$ s were backcrossed with recurrent parent GJG 33. The BC$_1$F$_1$ and BC$_2$F$_1$ were evaluated during Dec 2021-March 2022 and May-August 2022 respectively. BC$_2$F$_1$ was evaluated in Dec 2022-April 2023. Initially, 23 progenies of BC$_2$F$_2$ population were used for genotyping. Three progenies had wild alleles for both A and B genomes in allele specific primer assay. Hence, the remaining 20 progenies of the BC$_2$F$_2$ segregating population of the cross GJG 33 x Girnar 4 were further used to test the efficacy of the previously reported high oleic acid marker systems.

Allele-Specific PCR Assay (Chen et al., 2010): A 10 µL of PCR assay consists of the following: 2µL of diluted genomic DNA template (10 ng/µL), 0.5µL of Forward primer (10 pmol/µL), 0.5µL of the Reverse primer (10 pmol/µL), 4.0µL of 2X smart Prime PCR Master Mix and 3µL of sterile water. The PCR was carried out with a thermocycler (Make: Applied Biosystems, USA; Model: Veriti). The PCR program for all primers with initial denaturing at 94°C for 4 min, followed by 30 cycles of 94°C for 30 sec, annealing at 55°C for 45 sec and 72°C for 1 min, then extending at 72°C for 20 min and final step of storing at 4°C. The amplified products were separated on 2% agarose gel and documented after a run time of 1 hour at 120V. This assay uses a single forward primer with three reverse primers, IC-R for internal control, SUB-R for A genome and INS-R for B genome (Table 1).

CAPS marker assay (Chu et al., 2007 and 2009): Cleaved Amplified Polymorphic Sequence (CAPS) markers designed to differentiate the mutant and wild alleles in the A (Chu et al., 2007) and B genomes (Chu et al., 2009) were used. The PCR reaction volume of 10µL was made with 2µL of template DNA (10ng/µL), 0.5µL of Forward primer (10 pmol/µL), 0.5 µL of the Reverse primer (10 pmol/µL), 3.5µL of 2X smart Prime PCR Master Mix and 3.5µL of sterile water. The PCR was done with denaturing at 94°C for 5 min, followed by 32 cycles of 94°C for 30 sec, 50°C annealing temperature for A genome (51°C for B genome) for 30 sec, 72°C for 45 sec, final extension at 72°C for 7 min and sample maintained at 4°C.

Restriction digestion of the CAPS PCR product was done with the respective enzymes. The reaction volume is made with 3µL of the amplified PCR product, 0.25µL or 0.25 U of the enzyme Hpy99I (New England Biolabs, Ipswich, MA) (0.2µL or 0.2 U in the case of Hpy188I), 1µL of the reaction buffer (10X), 5.75µL of sterile water. The digestion was done at 37°C for 4 hrs in case of Hpy99I and 16 hrs for Hpy188I. The digested product was separated on 2% agarose gel (Table 2).

Yu et al. (2013) assay: In this genotyping assay, the PCR mixture consists of 0.6µL of template DNA (10ng/µL), 0.4µL of two forward primers each depending upon the reaction, 0.1µL of the reverse primer, 5µL of 2X smart Prime PCR Master Mix and 3.5 µL of sterile water. The amplified product was separated on 2% agarose gel run at 120V for 1 hour. Reaction I, Reaction II, Reaction III and Reaction IV were designed to detect the FAD2A wild allele (Ol$_1$), FAD2A mutant allele (ol$_1$), FAD2B wild allele (Ol$_2$), and FAD2B mutant allele (ol$_2$), respectively.

Table 1. Primers used by Chen et al. (2010) to identify FAD2A and FAD2B mutations

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence 5’ to 3’</th>
<th>Expected product size(bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F435-F</td>
<td>ATCCAAAGGCTGATTCTCAC</td>
<td></td>
</tr>
<tr>
<td>F435IC-R</td>
<td>CTCCTGTTGATGGATCTGTT</td>
<td>250</td>
</tr>
<tr>
<td>F435SUB-R</td>
<td>TGGAACAAACCTTCGTT</td>
<td>203</td>
</tr>
<tr>
<td>F435INS-R</td>
<td>AACACTCTGTCGGGTTCT</td>
<td>195</td>
</tr>
</tbody>
</table>

https://doi.org/10.37992/2023.1404.147
Mythili et al.,

Table 2. Details of CAPS marker and restriction enzyme used for FAD2A (Chu et al., 2007) and FAD2B (Chu et al., 2009) mutations

<table>
<thead>
<tr>
<th>Target gene</th>
<th>Primer</th>
<th>Sequence (5’-3’)</th>
<th>Restriction Enzyme</th>
<th>Expected amplicon size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ahFAD2A</td>
<td>af19: forward</td>
<td>GATTACTGATTATGGACTTT</td>
<td>Hpy99I</td>
<td>826</td>
</tr>
<tr>
<td></td>
<td>1056: reverse</td>
<td>CCAACCAACCTTTCAGAG</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ahFAD2B</td>
<td>bf19: forward</td>
<td>CAGAACATTTAGCTTGT</td>
<td>Hpy188I</td>
<td>1230</td>
</tr>
<tr>
<td></td>
<td>R1FAD: reverse</td>
<td>CTCTGACTATGCATCAG</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

(From Table 3). These 4 reactions were amplified with the same PCR program of denaturing at 94°C for 1 min, followed by 30 cycles of 94°C for 30 s, annealing at 53°C for 30 sec and 72°C for 90 sec, and a final extension of 72°C for 5 min.

RESULTS AND DISCUSSION

Breeding for high oleic acid content is an important objective in the groundnut breeding programme. Though many groundnut varieties are available in US and Australia with high oleic content, but there were no high oleic groundnut varieties available for cultivation in India till recently. Two Virginia bunch groundnut varieties viz., Girnar 4 and Girnar 5 were released first time in India by ICAR-Directorate of Groundnut Research, Junagadh in 2020 (http://variety.dgr.org.in). Phenotyping for high oleic acid content in groundnut involves oil extraction from the samples and estimation of fatty acid profile through GCMS. These methods involve much time and are tedious. Moreover, it leads to the destruction of seed samples which is a problem in the early stage of the groundnut breeding programme. Marker assisted selection for high oleic acid content in groundnut is an important tool in the high oleic breeding programme. It helps the breeder to select the plants without the destruction of seed samples and also the estimation is relatively easier than the conventional phenotypic method.

In the present manuscript, we compare efficiency of three types of molecular marker systems available to identify the high oleic acid content genotypes in groundnut.

Table 3. Primer combinations to identify the mutant and wild type of FAD2A and FAD2B (Yu et al., 2013)

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’ to 3’)</th>
<th>Annealing temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAD2A-F</td>
<td>GATTACTGATTATGGACTTTG</td>
<td>-</td>
</tr>
<tr>
<td>FAD2A-G</td>
<td>GTTTTGAGACAACACCTTCTC</td>
<td>53</td>
</tr>
<tr>
<td>FAD2A-A</td>
<td>GTTTTGAGACAACACCTTCTT</td>
<td>53</td>
</tr>
<tr>
<td>FAD2B-F</td>
<td>CAGAACACTAGTTGATGAGT</td>
<td>53</td>
</tr>
<tr>
<td>FAD2B-C</td>
<td>AACACTTCGCGCGCTT</td>
<td>53</td>
</tr>
<tr>
<td>FAD2B-A</td>
<td>AACACTTCGCGCGCTT</td>
<td>53</td>
</tr>
<tr>
<td>FAD2-R</td>
<td>CTCTGACTATGCATCAGAACCTT</td>
<td>-</td>
</tr>
</tbody>
</table>

Note: Reaction I: FAD2A-F + FAD2A-G + FAD2B-R; Reaction II: FAD2A-F + FAD2A-A + FAD2-R; Reaction III: FAD2B-F + FAD2B-C + FAD2-R; Reaction IV: FAD2B-F + FAD2B-A + FAD2-R

Allele Specific PCR markers: Positive plants in allele specific primers assay indicated the presence of mutant alleles of FAD2A (Fig. 1a) and FAD2B (Fig. 1b). However, this system cannot indicate whether the identified positive plants are homozygous or heterozygous. In Fig. 1a, except for plants 6, 12 and 18, all the progenies along with the donor parent P2 (Girnar 4) show positive for substitution mutation in the A genome at 203 bp. Similarly, in Fig 1b, all the progenies along with the donor parent Girnar 4 except for progenies 6, 12 and 18 show positive for insertion mutation in the B genome at 195 bp. The progenies 6, 12 and 18 have only wild alleles in both A and B genomes. Hence only the internal control band at 250 bp gets amplified. These progenies were not selected for further studies.

The SUB and INS primers designed by Chen et al. (2010) show amplification for the presence of the mutant alleles in both the A and B genomes irrespective of the heterozygosity. The main disadvantage of this allele specific PCR assay is that genotypes of Ol1Ol1/Ol2Ol2 and ol1ol1/Ol2Ol2 or Ol1Ol2/ol1ol2 and Ol1Ol2/ol1ol2 could not be differentiated. Chen et al. (2010) suggested using the control primer (amplification at 250 bp) along with SUB and INS primers indicating the PCR success. It also differentiated the PCR failure from the absence of mutant alleles. With this, the negative lines having the wild alleles of A and B genomes can be rejected at the early stage of screening. The time involved for the PCR program and the agarose gel electrophoresis run is about 2 hours 48 min.
Fig. 1a. Allele specific primer assay (Chen et al., 2010) for 23 BC2F2 progenies with parents
The lane L indicates the 100 bp ladder. P1 (Parent 1)-GJG 33; P2 (Parent 2)-Girnar 4; 1-23 are progenies of segregating population. Positive plants with substitution allele (mutation in FAD2A) had specific bands at 203 bp. The internal control band at 250 bp. wt- wild type; m- mutant

Fig. 1b Allele specific primer assay (Chen et al., 2010) for 23 BC2F2 progenies with parents
The lane L indicates the 100 bp ladder. P1 (Parent 1)-GJG 33; P2 (Parent 2)-Girnar 4; 1-23 are progenies of segregating population. Positive plants with insertion allele (mutation in FAD2B) had specific bands at 195 bp. The internal control band at 250 bp. wt- wild type; m- mutant

per mutant allele. Hence the time taken for genotyping of both alleles is approximately 5 hours and 36 min. Though this marker system helps to identify plants with mutant alleles, it could not differentiate the homozygote from the heterozygote. Hence, this marker system can be useful in the early stage of marker assisted selection and marker assisted backcross breeding programme. To identify the homozygotes in both the genome, the identified positive plants in this assay is subjected to CAPS marker assay.

CAPS marker assay (Chu et al., 2007 and 2009): In CAPS marker assay, the PCR product amplification was obtained at 826 bp for FAD2A (Fig 2a) and 1230 bp for FAD2B (Fig 2b). The PCR products were subjected to restriction digestion of FAD2A and FAD2B with the enzyme Hpy99I for 4 hrs and Hpy188i for 16 hrs respectively. The restriction enzyme Hpy99I was designed to identify the wild type. The PCR product of the progenies with wild alleles was identified by the enzyme and cut into two fragments, 598 bp and 228 bp. In the ahFAD2A mutant lines, due to the G to A transition, the sequence becomes unrecognizable for the enzyme and hence no digestion products were observed. In Fig 2c, the parent P2 (Girnar 4) had an undigested band at 826 bp and hence it is considered homozygous for the ahFAD2A mutant allele. Similarly, the progenies 4, 5, 7, 8, 9, 11, 15, 16, 17, 19, 20, 21, 22 and 23 were homozygous (ol/ol) for mutant allele of A genome. The parent P1 (GJG 33) is homozygous for the wild allele of the A genome and should have bands at 598 and 228 bp only. But due to improper digestion, a band at 826 bp is also visible in the case of parent P1. The progenies 1, 2, 3, 10, 13 and 14 were heterozygous for A genome (Ol/ol) and it has a banding pattern of the mutant (826 bp) and the wild allele (598 bp and 228 bp).

In the case of the B genome, the enzyme Hpy188I, produces restriction fragment bands at 736, 263, 171, 32 and 12 bp (last two bands were not visible on agarose gel) for wild allele. The mutant allele had restriction fragment bands at 505, 263, 231, 171, 32 and 12 bp (last two bands were not visible on agarose gel). In Fig 2d, the recurrent parent P1 (GJG 33), had bands at 736, 263 and 171 bp and hence considered homozygous for wild allele of the B genome. The donor parent P2
(Girnar 4) had bands at 505, 263, 231 and 171 bp and hence considered homozygous for mutant allele of the B genome. The progenies 4, 5, 7, 9, 16, 19, 20, 21 and 23 were homozygous (ol/ol) for ahFAD2B genome as these progenies had bands at 505, 263, 231 and 171 bp. The progenies 1, 2, 3, 8, 10, 11, 13, 14, 15, 17 and 22 were heterozygous for the ahFAD2B allele (Ol/ol). It has the banding pattern of both wild and mutant alleles (736, 505, 263, 231, 171 bp). Based on the results of CAPS markers of both genomes, following results were obtained: i) the
progenies 1, 2, 3, 10, 13 and 14 were heterozygous for both A and B genomes \((\text{Ol}_{ol}/\text{Ol}_{ol})\), ii) the progenies 4, 5, 7, 9, 16, 19, 20, 21 and 23 were homozygous for both A and B genomes \((\text{ol}_{ol}/\text{ol}_{ol})\) and iii) the progenies 8, 11, 15, 17 and 22 were homozygous for A and heterozygous for B genomes \((\text{ol}_{ol}/\text{Ol}_{ol})\).

The time taken for the CAPS marker PCR program is 1 hour 28 min for the A genome and 1 hour 27 min for the B genome. The confirmation of CAPS marker amplification on the agarose gel takes about 1 hour. It is followed by digestion of the PCR product with the restriction enzymes for about 4 hrs and 16 hrs under incubation for A and B genomes respectively. The restricted product is visualized under agarose gel after 1 hour of electrophoresis run time.

The time involved to identify the homozygous line for the A genome alone is approximately 10 hours 16 min and for the B genome alone is 22 hours 15 min. Hence time taken for genotyping of both alleles is approximately 32 hours and 31 min. In the case of restriction digestion of ahFAD2A in parent P1(GJG 33), a small amount of the product is left undigested even if it is digested with increased enzyme concentration. Hence, the undigested band is visible at 826 bp. Chu et al. (2007) also reported a similar situation for homozygous parents for the wild allele in the A genome. In the CAPS marker system, homozygote plants can be differentiated from heterozygotes. This system can be used in all the stages of marker assisted selection and marker assisted backcross breeding programme. However, this system involves additional costs for the restriction enzymes. Hence, many researchers (Janila et al., 2016; Bera et al., 2018; Bera et al., 2019) used the allele specific primer (Chen et al., 2010) system for the initial stages of the breeding programme and the CAPS markers in the later stage to identify homozygotes. The combination of both systems is good for both time and cost in high oleic groundnut breeding programme.

Yu et al. (2013) assay: The Reaction I primers amplified the wild allele in the A genome \((\text{Ol}_{.\text{ol}}, \text{and Ol}_{.\text{ol}})\) at 557 bp (Fig 3a). The progenies 1, 2, 3, 10, 13 and 14 had wild allele for the A genome. Reaction II primers amplified the mutant allele of the A genome \((\text{Ol}_{.\text{ol}}, \text{and Ol}_{.\text{ol}})\) at 550 bp (Fig 3b). All the 20 progenies (except 6, 12 and 18) had the mutant allele for the A genome. Reaction III primers amplified the wild allele of the B genome \((\text{Ol}_{.\text{ol}}, \text{and Ol}_{.\text{ol}})\) at 550 bp (Fig 3d). All the progenies (except for 6, 12 and 18) had the mutant allele of the B genome. With Reactions I and II, the progenies 4, 5, 7, 8, 9, 11, 15, 16, 17, 19, 20, 21, 22 and 23 were identified as homozygous \((\text{ol}_{.\text{ol}}, \text{for}}\) A genome. The progenies 1, 2, 3, 10, 13 and 14 were identified as heterozygous \((\text{Ol}_{.\text{ol}}, \text{for}}\) A genome. Similarly, with Reactions III and IV, the progenies 4, 5, 7, 9, 16, 19, 20, 21 and 23 were identified as homozygous \((\text{ol}_{.\text{ol}}, \text{for}}\) B genome. The progenies 1, 2, 3, 8, 10, 11, 13, 14, 15, 17 and 22 were identified as heterozygous \((\text{ol}_{.\text{ol}}, \text{for}}\) B genome. Based on these results, it can be conclude that progenies 4, 5, 7, 8, 9, 11, 15, 16, 17, 19, 20, 21, 22, 23 for ahFAD2B are homozygous for mutant alleles. The progenies 4, 5, 7, 9, 16, 19, 20, 21, 23 were homozygous for both A and B genomes. The forward primer (FAD2A-F or FAD2B-F) and the reverse primer (FAD2-R) helps in the amplification of the internal reference band at 1200bp which is an indication of PCR success. The time taken for this assay is approximately 2 hours and 38 min. The total time taken to identify a line with homozygous mutant alleles for both A and B genomes is about 10 hours and 32 min. In the case of the identification of mutant alleles alone, the time taken for genotyping is approximately 5 hours and 16 min only. This system can be used at all the stages of marker assisted selection and backcross breeding programme. In the early stages of the breeding programme, Reactions II and IV can be used.
Fig. 3. Banding pattern of progenies of BC$_2$F$_2$ generation of the cross GJG 33 x Girnar 4 with assay of Yu et al. (2013)
L- 100 bp ladder; P1(Parent1)-GJG 33; P2 (Parent2)-Girnar 4; wt- wild type; m- mutant

Table 4. Time duration for different assay systems used

<table>
<thead>
<tr>
<th>Assay system</th>
<th>Total time duration taken for genotyping both the alleles to identify the homozygotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allele specific PCR assay + CAPS marker assay</td>
<td>38 hours and 7 min</td>
</tr>
<tr>
<td>Yu et al. (2013) assay</td>
<td>10 hours and 32 min</td>
</tr>
</tbody>
</table>

to identify plants with mutant alleles. In the later stage of the breeding programme where the identification of homozygote for mutant allele is required, all the reactions I, II, III and IV can be used. The time taken by different assay systems to identify the homozygous alleles in both the genome is given in Table 4.

Marker assisted breeding program aims to identify the appropriate progeny with the target trait under homozygous condition. On comparing the three genotyping assays, the reaction primers system reported by Yu et al. (2013) was found as effective markers in terms of time as well as cost for marker assisted selection of high oleic acid breeding lines in groundnut.

ACKNOWLEDGEMENT
This work was supported by the Science and Engineering Research Board, Department of Science and Technology, Government of India [Grant Number: CRG/2020/000127] to the corresponding author.

REFERENCES

Bera, S.K., Kamdar, J.H., Kasundra, S.V., Patel, S.V., Jasani, M.D., Maurya, A.K., Dash, P., Chandrashekar, A.B., Rani, K., Manivannan, N. and Janila, P. 2019. Steady expression of high oleic acid in peanut bred by marker-assisted backcrossing for fatty acid desaturase mutant alleles and its effect on seed germination along with other seedling traits. PloS one, 14(12):0226252. [Cross Ref]


Fraser, G.E., Sabate, J., Beeson, W.L. and Strahan, T.M. 1992. A possible protective effect of nut consumption on risk of coronary heart disease: the Adventist Health Study. *Archives of Internal medicine, 152*(7):1416-1424. [Cross Ref]


