Two high-oleic-acid soybean mutants, M23 and KK21, have disrupted microsomal omega-6 fatty acid desaturase, encoded by *GmFAD2-1a*

Toyoaki Anai*1), Tomoko Yamada*1), Rumiko Hidestima*1), Takehito Kinoshita*1,2), Shaikh M. Rahman*1,3) and Yutaka Takagi*1)

1) *Laboratory of Plant Genetics and Breeding, Faculty of Agriculture, Saga University, 1 Honjo-y-machi, Saga 840-8502, Japan*
2) *Present address: Saga Prefectural Agricultural Research Center, 1088 Kawase-machi, Saga 840-2205, Japan*
3) *Present address: National Agricultural Research Center for Western Region, 6-12-1 Nishi-fukatsu, Fukuyma 721-8514, Japan*

Elevating the oleic acid content of soybean (*Glycine max* (L.) Merr.) is a major focus of breeding programs. Previously, we created two high-oleic-acid soybean mutants, M23 and KK21, by X-irradiation. We expected them to have modifications in genes encoding microsomal omega-6 fatty acid desaturase. The objectives of this study were to evaluate which members of the *GmFAD2* gene family contribute to oleic acid production during seed maturation, to characterize the mutant genes, and to establish molecular markers for breeding of high-oleic-acid soybeans. Three *GmFAD2* genes were expressed in developing seeds; the gene products of *GmFAD2-1a* and *GmFAD2-1b* were more active than that of *GmFAD2-2a* during seed development. We identified different nucleotide modifications in *GmFAD2-1a* in M23 and KK21. Using nuclease-cleaved DNA fragment-length polymorphisms, we developed a novel molecular marker to distinguish between KK21 mutant and wild-type alleles. This information could be useful for improving soybean oil quality by using the mutant genes from M23 or KK21, and for screening novel high-oleic-acid soybean mutants.

**Key Words:** oleic acid, microsomal omega-6 fatty acid desaturase, *Glycine max*, mutant, *FAD2*, CEL 1, molecular marker.

**Introduction**

Soybean is an important oil crop, producing more than 30 million tonnes of oil annually worldwide (FAO 2007). The alteration of fatty acid composition is a major interest in soybean improvement, because the composition directly influences oil quality. In general, a high level of saturated fatty acids provides a high degree of stability, but it is nutritionally undesirable because of these acids’ capacity to raise blood levels of low-density lipoprotein cholesterol (Cox et al. 1995). Unfortunately, oil rich in polyunsaturated fatty acid is easily oxidized, and unwanted *trans*-fatty acids are formed during industrial hydrogenation (Ascherio and Willett 1997). In contrast, oil rich in monounsaturated fatty acids (e.g., oleic acid) is relatively resistant to oxidation, and can lower blood cholesterol (Grundy 1986) and protect endothelial cells (Toborek et al. 2002). Oleic acid–rich soybean oil is also a useful source of some industrial products (Jaworski and Cahoon 2003). In general, conventional soybean oil contains about 24% monounsaturated oleic acid and 54% polyunsaturated linoleic acid (Cherrak et al. 2003), however, novel soybean cultivar with higher-oleic acid and lower-linoleic acid is desired.

Microsomal omega-6 fatty acid desaturase (FAD2) catalyzes the conversion of oleic acid to linoleic acid in developing seeds. Thus, *FAD2* genes are a primary target of genetic manipulation in several oil crops (Kinney 1996, Stoutjesdijk et al. 2000, Buhr et al. 2002, Liu et al. 2002). Transgenic technologies have produced high-oleic-acid oil crops, but genetically modified organisms are not yet fully acceptable to consumers.

We isolated a high-oleic-acid mutant, M23 (50.4% content), from the *M*2 generation of an X-irradiated soybean, ‘Bay’, which has a normal oleic acid content (22.5%) (Rahman et al. 1994). The high-oleic-acid character is controlled by a recessive mutant allele, *ol* (Takagi and Rahman 1996). M23 has a deletion in an unknown member of the *GmFAD2-l* gene family (Kinoshita et al. 1998). We later isolated another high-oleic-acid mutant, KK21, from a different *M*2 population of X-irradiated Bay (unpublished data). Since a hybridization experiment we conducted showed that KK21 and M23 have different alleles at the same locus (unpublished data), we suspect that both lines contain lesions in the same gene.

To develop novel high-oleic-acid soybean cultivars by marker-assisted breeding, it is necessary to characterize the mutant genes responsible for these high-oleic-acid mutants and to develop an assay system for determining the mutant genes in M23 and KK21. Here, we characterize the molecular nature of these two mutants. This information will be useful for incorporating these mutant genes into other lines.
Materials and Methods

Plant materials
We grew soybean (Glycine max (L.) Merr.) cultivar Bay and Bay-derived high-oleic-acid mutants M23 and KK21 in a greenhouse at Saga University under natural light. We collected root nodules, stems, green leaves, pods, and seeds at different stages of development, quickly froze them in liquid nitrogen, and stored them at -80°C until use.

Nucleic acid preparation
Total DNA was extracted from green leaves by the CTAB method (Murray and Thompson 1980) for PCR and Southern blot analyses. Total RNA was prepared from all tissues by the modified acidic phenol-SDS method (Anai et al. 2003) and used for RT-PCR analysis.

Semi-quantitative RT-PCR analysis
For semi-quantitative gene expression analysis, first-strand cDNA was synthesized from 1 μg total RNA by oligo(dT) priming with a modified M-MLV reverse transcriptase (ReverTra Ace, ToyoBo). The first-strand cDNA concentrations were equalized against actin cDNA as a control with an actin-specific primer set (Soy-Actin-F, 5'-gtttgggatgcagggag ataatttgtaa-3'; Soy-Actin-R, 5'-aatctcagctctttttgccg-3'). Then the cDNAs were used as templates for PCR to amplify part of GmFAD2-1a, GmFAD2-1b, GmFAD2-2a, and GmFAD2-2b with sequence-specific primer sets (GmFAD2-1a-RT-F1, 5'-acaccttcgctaagctttctac-3'; GmFAD2-1b-RT-F1, 5'- accttcgctaagctttctac-3'; GmFAD2-2a-RT-R1, 5'-tctctcggctctcagagtgtttgtttctcag-3'; GmFAD2-2b-RT-R1, 5'-tctctcggctctcagagtgtttgtttctcag-3'; GmFAD2-2a/b-RT-R1, 5'-tctctcggctctcagagtgtttgtttctcag-3'; GmFAD2-2b-RT-R1, 5'-tctctcggctctcagagtgtttgtttctcag-3'). The specificity of these primer sets was confirmed by direct sequencing of amplified DNA fragments. The GmFAD2 gene family members were previously cloned by us from soybean cultivar Bay (acc. nos. AB188250, AB188251, AB188252, AB188253). Amplification was carried out under standard reaction conditions with Blend Taq DNA polymerase (ToyoBo) in a thermal cycler (iCycler, Bio-Rad) programmed at 94°C/2 min; 30 cycles of 94°C/30 s, 65°C/30 s, 72°C/1 min; and a final 72°C/5 min. The RT-PCR products were separated by 1% agarose gel electrophoresis.

Southern blot analysis
Total DNA (3 μg) was digested with EcoRI and electrophoresed in 1% agarose gel. The separated DNA fragments were transferred to a nylon membrane and detected with GmFAD2-1a cRNA probe. Hybridization and detection were carried out as described (Anai et al. 2005).

PCR analysis
Two seed-specific FAD2 genes (GmFAD2-1a and GmFAD2-1b) were amplified with three sequence-specific primer sets (GmFAD2-1-Full-F3, 5'-attgataacaacctcgcttctcaca-3', and GmFAD2-1-SEQ3R, 5'-atttgggatggtgagcagaggaa-3', for 5'-region of GmFAD2-1a; GmFAD2-1a-RT-F1 and GmFAD2-1-Full-R1, 5'-aaattacaccaagtcatagcg-3', for 3'-region of GmFAD2-1a; GmFAD2-1b-Full-F2, 5'- caacatctctctacatctcctctct-3', and GmFAD2-1b-Full-R1, 5'-tgacacaaagtcatagcg-3', for GmFAD2-1b) from genomic DNA, prepared from each mutant line as a template. Amplification conditions were the same as in the RT-PCR experiment, except for elongation (72°C/90 s).

DNA sequencing
For nucleotide sequencing, the open reading frame (ORF) of each gene was amplified with a sequence-specific primer set (GmFAD2-1-Full-F3 and GmFAD2-1-Full-R1 for GmFAD2-1a; GmFAD2-1b-Full-F2 and GmFAD2-1b-Full-R1 for GmFAD2-1b; GmFAD2-2a-Full-F1, 5'-agattttggtgtgagcagaggaa-3', and GmFAD2-2b/a-Full-R1, 5'-aaattacaccaagtcatagcg-3', for GmFAD2-2a; GmFAD2-2b/a-Full-R1, 5'-aaattacaccaagtcatagcg-3', for GmFAD2-2b) under the same thermal cycle conditions as for RT-PCR analysis, except for extension (72°C/2 min). The amplified DNA fragments were separated by agarose gel electrophoresis and purified by a QIAEX II Gel Extraction Kit (Qiagen). Purified PCR products were sequenced with a Big Dye Terminator Cycle Sequencing Kit v. 3.1 (Applied Biosystems), and were then further purified with Sephadex G-50 spin columns (Amersham Biosciences). Sequencing was performed by an automatic sequencer (Model 310, Applied Biosystems).

Yeast transformation and fatty acid analysis
Three cDNA fragments (GmFAD2-1a, GmFAD2-1b, and GmFAD2-2a) derived from Bay and a fragment (GmFAD2-1a) derived from high-oleic-acid mutant KK21 were cloned into the yeast expression vector pYES2/CT (Invitrogen). Yeast cells (INVSc1, Invitrogen) harboring the vector were grown in a synthetic liquid medium lacking uracil (SC-Ura) supplemented with 1% (w/v) raffinose and 2% (w/v) glucose for 72 h at 20°C. The harvested yeast cells were resuspended in 0.5 M sulfuric acid in methanol containing 2% (v/v) dimethoxypropane and incubated for 1 h at 80°C (Zank et al. 2002). Fatty acid methyl esters were extracted with n-hexane and separated under isothermal conditions at 180°C in a Yanaco G6800 series gas chromatograph equipped with a 25-m × 0.25-mm Quadrex 23 bonded fused silica capillary column and a flame ionization detector. Each peak was identified by comparison with the retention time of standard fatty acid methyl esters.

Mismatch detection with CEL 1 nuclease
GmFAD2-1a sequences of KK21 and Bay were amplified with sequence-specific primers (GmFAD2-1-Full-F3 and GmFAD2-1-Full-R1). Both amplified DNA fragments (~1.2 kbp) were used for CEL 1-mediated mismatch detection assay (Olejukowsk et al. 1998). Cleaved DNA fragments were
Molecular characterization of high-oleic soybean mutants

449

Fig. 1. Tissue-specific expression of GmFAD2 genes in normal soybean cultivar Bay. Total RNA was prepared from root nodules, stems, green leaves, pods, and seeds at different stages of development. RT-PCR was carried out with gene-specific primer sets.

Fig. 2. (A) Southern blot and (B) PCR analysis of GmFAD2 genes in normal soybean cultivar Bay and two high-oleic-acid mutants. In A, EcoRI-digested DNA fragments were hybridized with GmFAD2-1a cRNA as a probe. In B, each DNA fragment was amplified with gene-specific primer sets.

separated by 2% agarose gel electrophoresis and stained with ethidium bromide. CEL I was purified from 10 kg of celery stalks as reported (Yang et al. 2000) with several modifications: briefly, CEL I was sequentially purified from celery juice with ammonium sulfate precipitation, ConA-Sepharose chromatography, phosphocellulose P-11 chromatography, and Q-Sepharose FF chromatography.

Results

Characterization of GmFAD2 gene family expression

Four GmFAD2 gene family members (GmFAD2-1a, GmFAD2-1b, GmFAD2-2a, and GmFAD2-2b) have been identified in soybean cultivar Bay and have been sequenced. Their tissue-specific expressions were quantified by RT-PCR using gene-specific primers. GmFAD2-1a and GmFAD2-1b were predominantly expressed in developing seed, whereas GmFAD2-2a was moderately expressed in various tissues except for green leaf (Fig. 1). In contrast, GmFAD2-2b transcript was not detected in any tissues (data not shown). Besides the similarity in the expression profiles of GmFAD2-1a and GmFAD2-1b, the ORFs of these genes are 95.4% identical at the nucleotide level. This result suggests that GmFAD2-1a and GmFAD2-1b are duplicate genes that encode major isozymes of FAD2 during the biosynthesis of storage lipids in developing soybean seeds.

Identification of mutant genes in high-oleic-acid mutants M23 and KK21

Previously, we designated the high-oleic-acid mutant allele of M23 ol (Takagi and Rahman 1996). In addition, our unpublished hybridization experiment showed that the mutant gene of KK21 is allelic to that of M23. The radiation-treatment frequently induces relatively large lesions in their genome. To identify relatively large lesions on GmFAD2-1a or GmFAD2-1b, we analyzed DNA samples prepared from the mutants and Bay by Southern blot analysis with a DIG-labeled full-length cRNA probe for GmFAD2-1a. In Bay and KK21, we detected three hybridization signals (approx. 4.5, 2.5 and 1.9 kbp), but only the smaller two in M23 (Fig. 2A). However, because of their high similarity, it is impossible to distinguish GmFAD2-1a and GmFAD2-1b by Southern blot analysis. To elucidate which gene was disrupted in M23, we used PCR with gene-specific primers (Fig. 2B). Four primer sets successfully amplified the specific DNA fragments in Bay and KK21. By contrast, no DNA fragments were amplified with either GmFAD2-1a-specific primer set (GmFAD2-1-Full-F3 and GmFAD2-1-SEQ3R for 5’-region; GmFAD2-1a-RT-F1 and GmFAD2-1-Full-R1
for 3'-region) in M23. These results clearly indicate that the ORF of GmFAD2-1a was completely missing in M23. To elucidate the molecular consequence of the high-oleic-acid mutation in KK21, we analyzed the nucleotide sequences of GmFAD2 gene family members. Comparison of nucleotide sequences with Bay indicated that KK21 has a single nucleotide deletion 232 bases downstream from the ATG initiation codon of GmFAD2-1a. This deletion results in a frameshift mutation and an abnormal GmFAD2-1a gene product. In contrast, the nucleotide sequences of both GmFAD2-1b and GmFAD2-2a in these mutants are completely identical to that of the original cultivar Bay. Therefore, GmFAD2-1a was mutated in both mutant lines.

Evaluation of GmFAD2 enzymatic activities in yeast

Since baker's yeast, Saccharomyces cerevisiae, has no enzymes that produce polyunsaturated fatty acids, it is useful for evaluating plant FAD activity by heterologous expression (Anai et al. 2005). A non-yeast linoleic acid (arrowhead in Fig. 3) with a retention time of 7.4 min was observed in three recombinant yeast strains carrying the wild-type GmFAD2 family genes ([GmFAD2-1a (Fig. 3B), GmFAD2-1b (Fig. 3C), and GmFAD2-2a (Fig. 3D)]. This result clearly demonstrates that these enzymes have the FAD2 activity. Additionally, an extra peak with a retention time of 5.0 min (asterisk in Fig. 3) appeared only in the chromatogram of the GmFAD2-1b-expressing yeast strain. Because the rapeseeded FAD6 gene encoding plastidal ω-6 fatty acid desaturase could utilize at least three substrates [16:1 (7e), 16:1 (9c) and 18:1 (9c)] (Hitz et al. 1994), we deduced that the extra peak is a methyl ester of 16:2 fatty acid produced from palmitoleic acid [16:1 (9c)]. Our results also suggest that the substrate-specificity differs between GmFAD2-1b and another gene products. Unlike the yeast strain expressing wild-type GmFAD2-1a, the yeast strain expressing the KK21 mutation did not show the linoleic acid peak (Fig. 3E). This result provides direct support for a comparable oleic acid content in KK21 (47.2% in 2007) as in M23 (48.4% in 2007), which contains a null mutant of GmFAD2-1a (Table 1).

Detection of the single nucleotide deletion in GmFAD2-1a in KK21

To develop molecular markers which distinguish between wild-type and KK21 mutant alleles, we tried to detect mutated nucleotide sequences by using the mismatch-specific nuclease CEL1. CEL1 is a highly specific nuclease which recognizes and digests any mismatch- or indel-containing heteroduplex DNA fragments (Oleykowski et al., 1998). After denaturing-annealing treatment, the PCR products of the wild-type and KK21-derived GmFAD2-1a were digested singly or together with CEL1. Two digested DNA fragments (approx. 0.9 and 0.3 kbp) were observed only in the mixture lane after electrophoresis (Fig. 4). Identifying the genetic lesion underlying the high-oleic-acid trait of KK21 should provide a simple and robust tool for selecting this mutant allele in early generations of segregating progeny.

Discussion

We characterized FAD2 gene family members in soybean by tissue-specific RT-PCR analysis and heterologous expression analysis of recombinant proteins in yeast. Our RT-PCR results strongly suggest that GmFAD2-1a and GmFAD2-1b encode major FAD2. Because these genes are expressed mainly in the developing soybean seeds, they make attractive targets for the improvement of oleic acid content. Recently, Schlueter et al. (2007) reported the structural and functional diversity of the soybean FAD2 family and identified a FAD2 gene whose transcripts accumulate under cold conditions. They also reported that both FAD2-1 genes accumulated in developing seeds, as we observed. Because the expression of FAD2 genes is regulated in a complicated manner (e.g. developmental stage, temperature, etc.), the genes might physiologically affect some agronomic

Table 1. Fatty acid composition of Bay and high-oleic-acid mutant soybean lines in 2007

<table>
<thead>
<tr>
<th>Line</th>
<th>Palmitic acid (%)</th>
<th>Stearic acid (%)</th>
<th>Oleic acid (%)</th>
<th>Linoleic acid (%)</th>
<th>α-Linolenic acid (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bay</td>
<td>11.3±2.5</td>
<td>1.8±0.2</td>
<td>27.2±1.7</td>
<td>52.7±3.1</td>
<td>7.0±0.7</td>
</tr>
<tr>
<td>KK21</td>
<td>11.4±0.5</td>
<td>1.7±0.2</td>
<td>47.2±1.3</td>
<td>33.2±0.8</td>
<td>6.5±0.7</td>
</tr>
<tr>
<td>M23</td>
<td>11.7±0.7</td>
<td>1.4±0.1</td>
<td>48.4±0.9</td>
<td>30.9±0.6</td>
<td>6.6±0.3</td>
</tr>
</tbody>
</table>

Means±SD were obtained from three independent experiments.
characters like low-temperature tolerance in vegetative tissues. Thus, it is appropriate to choose developing seed-specific genes as targets for seed oil improvement by conventional breeding programs.

The biochemical characterization of GmFAD2 gene family products is an important aspect of uncovering their biological functions. Yeast expression assays revealed that at least three GmFAD2 genes are involved in the conversion of oleic acid into linoleic acid in developing soybean seeds. Furthermore, GmFAD2-1b-expressing yeast cells synthesized two non-yeast fatty acids, linoleic acid and a predicted 16:2 fatty acid derived from palmitoleic acid (16:1). The physiological functions of this unique enzymatic activity are yet unknown, but the difference in substrate specificity between GmFAD2-1b and the other members could provide us with clues to their molecular evolution. Broun et al. (1998) reported the structural similarity between Arabidopsis FAD2 gene products and another fatty acid modifier, hydroxylase, each of which is easily converted into the other with the displacement of a few amino acids. GmFAD2-1b might be an important example of an evolutionary intermediate of duplicate genes in the soybean paleopolyploid genome.

In addition, the results of expression level analysis (Fig. 1) and enzymatic activity analysis (Fig. 3) strongly suggested that the GmFAD2-1b gene contributes a substantial proportion of oleic acid production in soybean seed. We expect that extremely high-level oleic acid (over 70% or more) accumulation may be accomplished by the combination of GmFAD2-1a and GmFAD2-1b mutant genes. However, GmFAD2-1b mutant could not be obtained from this GCBased screening. This result may suggest that the GmFAD2-1b gene effect is unstable for some environmental conditions (e.g., temperature etc.).

Recently, large numbers of molecular markers based on genome sequences or expressed sequence tags have been developed in many commercially important crops. Several groups have independently developed different types of soybean molecular markers (Shoemaker and Specht 1995, Cregan et al. 1999, Maughan et al. 1996, Ferreira et al. 2000, Xia et al. 2007), and a draft genome sequence assembled from whole-genome shotgun data has been made available on the web site of DOE-JGI (http://www.phytozome.net/soybean.php). In their review, Varshney et al. (2006) indicated that ‘functional markers’ have some advantages over randomly developed markers. They defined functional markers showing phenotypic differences as ‘perfect markers’ (PMs), and suggested that marker quantitative trait locus (QTL) analysis was an efficient approach to developing novel PMs. Variant alleles of commercially important traits are limited, because QTL analysis normally needs the hybridized progeny of two cultivars with different phenotypes of some particular characters. However, here we developed a novel PM for high oleic acid content of soybean through the use of an X-ray-induced mutant with a single-base-pair deletion and CEL I nuclease digestion. Such mutant-based molecular marker development provides a robust method to obtain simultaneously both molecular-markers and novel genes linked with important phenotypes. Ionizing radiation, including X-ray, gamma-rays, and fast neutrons, generally induces larger (>5 kbp) deletions or chromosomal rearrangements (Shirley et al. 1992, Cecchini et al. 1998), which are often associated with undesirable agronomic traits. However, ionizing radiation can sometimes give small deletions, and is therefore quite useful to develop valuable mutants. The small deletion-based mutation is a preferable approach to generate loss-of-function mutants because this type of mutation frequently makes the gene product defective. Hence, the combination of a method for screening single-to several-base-pair deletions and a detection method compatible with short deletions is highly desirable for mutant-based PM development. Our results will enhance the development of genomics-assisted mutant breeding.

Acknowledgements

We thank Dr. A. Suzuki for helpful discussions and M. Haraguchi at Saga University for technical assistance. This study was partly supported by a grant from the Ministry of Agriculture, Forestry and Fisheries of Japan (Genomics for Agricultural Innovation).

Literature Cited


Buhr, T., S. Sato, F. Ebrahim, A. Xing, Y. Zhou, M. Mathiesen.