Genes & Genomics (2025) 47:307–320
https://doi.org/10.1007/s13258-024-01608-5

Online ISSN 2092-9293
Print ISSN 1976-9571

RESEARCH ARTICLE

Identification of SNPs associated with fatty acid contents in mutant soybean lines by a genome-wide association study

Jeong Woo Lee[1,3] · Jung Min Kim[1] · Dae June Kim[1,3] · Ji Su Seo[2] · Bo-Keun Ha[3] · Soon-Jae Kwon[1]

Received: 24 September 2024 / Accepted: 6 December 2024 / Published online: 23 December 2024

Abstract

Background Vegetable oils are primarily composed of unsaturated fatty acids. Soybean [*Glycine max* (L.) Merr.] oil, accounting for 28% of the global production of vegetable oil, contains mainly two saturated fatty acids (palmitic acid and stearic acid) and three unsaturated fatty acids (oleic acid, linoleic acid, and linolenic acid) in seeds.

Objective The five fatty acids determine soybean oil quality. We aimed to identify genetic relationship between genomics and fatty acid contents in soybean mutant pool.

Methods This study used a mutant diversity pool (MDP) comprising 192 soybean lines. A genome-wide association study (GWAS) was conducted with the diverse fatty acid contents in MDP and 17,631 filtered SNPs from genotyping-by-sequencing (GBS).

Results The GWAS revealed nine significant SNPs within intragenic regions that were associated with fatty acid composition. These SNPs corresponded to six genes (*Glyma.03g042500*, *Glyma.07g069200*, *Glyma.13g150200*, *Glyma.14g223100*, *Glyma.15g084700*, and *Glyma.15g274000*), of which three (*Glyma.03g042500*, *Glyma.13g150200*, and *Glyma.15g274000*) were predicted to be candidate genes influencing oleic acid, linoleic acid, and linolenic acid contents. Analyses of SNP allelic effects revealed the largest and smallest significant differences in fatty acid contents were 5.53% (linolenic acid) and 0.4% (stearic acid), respectively.

Conclusion The present study detected significant phenotypic variations and genetic associations underlying the fatty acid composition of soybean seeds in MDP lines. The mutant seeds differed from the original cultivars in terms of fatty acids composition, with the allelic effects of significant SNPs influencing the fatty acid content in seeds. These findings may be useful for enhancing breeding strategies to optimize soybean oil quality for various uses.

Keywords Soybean · Fatty acid · Genotyping-by-sequencing · Genome-wide association study · SNP

Introduction

In contrast to animal fats, which have large amounts of saturated fatty acids, vegetable oils are primarily composed of unsaturated fatty acids, with some exceptions (Phuah et al. 2022). Oils extracted from seeds or other fruit parts are called vegetable oils. The global production of vegetable oil exceeds 200 million tons per year and oil crops are grown on approximately 330 million hectares (FAO 2019). The most widely produced vegetable oil is palm oil, which accounts for about 37% of global vegetable oil production (more than 75.7 million tons per year), followed by soybean oil extracted from seeds, accounting for 28% of global vegetable oil production (approximately 60 million tons per year) (Shahbandeh 2024; Chiriacò et al. 2024). Soybean oil (20% of seeds) contains 16% saturated fat, 23% monounsaturated

Jeong Woo Lee and Jung Min Kim have contributed equally to this work.

✉ Bo-Keun Ha
bkha@chonnam.ac.kr

✉ Soon-Jae Kwon
soonjaekwon@kaeri.re.kr

[1] Advanced Radiation Technology Institute, Korea Atomic Energy Research Institute, Jeongeup 56212, Republic of Korea

[2] College of Pharmacy, Chungnam National University, Daejeon 34134, Republic of Korea

[3] Division of Plant Biotechnology, Colleage of Agriculture and Life Sciences, Chonnam National University, Gwangju 61186, Republic of Korea

fat, and 58% polyunsaturated fat (Poth 2000). Almost 100% of soybean oil consists of the following five fatty acids: two saturated fatty acids (palmitic acid and stearic acid) and three unsaturated fatty acids (oleic acid, linoleic acid, and linolenic acid) (Ivanov et al. 2010). The composition of these five fatty acids is an important determinant of soybean oil quality. The three unsaturated fatty acids have the same 18 carbon atoms. As polyunsaturated fatty acids, linoleic acid (18:2), which has two carbon double bonds, and linolenic acid (18:3), which contains three carbon double bonds, represent approximately 53% and 8% of soybean oil, respectively. With identical carbon atoms, another unsaturated fatty acid, oleic acid (18:1), which has one carbon double bond, is a monounsaturated fatty acid that accounts for about 23% of soybean oil. As saturated fatty acids, palmitic acid (16:0), which has 16 carbon atoms, and stearic acid (18:0), which contains 18 carbon atoms, represent approximately 4% and 12% of soybean oil, respectively (Lee et al. 2007). Modulating the contents of these five fatty acids in soybean oil is an objective of soybean breeding programs (Wilson 2004).

Modifying the contents of saturated fatty acids (palmitic acid and stearic acid) and unsaturated fatty acids (oleic acid, linoleic acid, and linolenic acid) can affect human health as well as biofuel production (Vega et al. 2024; Meydani et al. 1991). Considering the potential effects of unsaturated fatty acids in vegetable oils, soybean oil, which has a relatively low saturated fat content, may positively affect overall health and decrease the risk of coronary heart disease (Messina et al. 2021). Among unsaturated fatty acids, oleic acid is important because of its health effects and oxidative stability. A diet rich in oleic acid-containing edible oils is associated with a decrease in cholesterol levels and a reduced risk of arteriosclerosis and coronary heart disease (FDA 2018; Wardlaw and Snook 1990). Another unsaturated fatty acid, linolenic acid, has many double bonds, making it susceptible to oxidation, which results in off-flavors (Liu and White 1992). Although hydrogenation can improve the oxidative stability and taste of oil, it leads to the formation of trans fats (Romero et al. 2000; Liu et al. 2007). Because of oxidative stability, oils with high oleic acid contents and low linolenic acid contents must have low rancidity levels and a long shelf life, especially high-quality soybean oil used for preparing food (Napolitano et al. 2018). Additionally, the oxidative stability of oleic acid is important because the esterification of oleic acid is reportedly involved in biodiesel synthesis (Narkhede and Patel 2013). Modifying the fatty acid composition and content in soybean oil using biotechnology-related breeding methods may lead to increased interest in biodiesel derived from commodity crops (Kinney and Clemente 2005).

Quantitative trait loci (QTLs) associated with the five fatty acids in soybean oil have been detected in recombinant inbred lines. Moreover, genotypic associations with soybean fatty acid traits have been revealed on the basis of genome-wide association study (GWAS) data. Notably, QTLs for palmitic acid and stearic acid were identified on chromosomes 2, 14, and 16, whereas QTLs for four fatty acids (palmitic acid, oleic acid, linoleic acid, and linolenic acid) were detected on chromosome 19 (Panthee et al. 2006; Hyten et al. 2004). In another study, QTLs for five fatty acids were identified on chromosomes 3, 13, 14, 15, 16, and 18 (Akond et al. 2014). In terms of saturated fatty acids (palmitic acid and stearic acid), the results of a GWAS involving soybean seeds revealed single nucleotide polymorphisms (SNPs) on chromosomes 5, 6, 11, and 14 and five genes (*Glyma.05g012300*, *Glyma.05g011100*, *Glyma.06g214800*, *Glyma.11g226900*, and *Glyma.14g121400*) were identified as candidate genes (Sung et al. 2021). Furthermore, 149 SNPs associated with unsaturated fatty acids (oleic acid, linoleic acid, and linolenic acid) in soybean seeds were identified according to GWAS results (Zhao et al. 2019b). On the basis of GWAS data, 12 genes on chromosomes 7, 8, 13, 14, 16, 18, and 20 were predicted to be involved in unsaturated fatty acid synthesis. Another GWAS identified 37 genes related to oleic acid, with *Glyma.04g116500.1* (*GmWRI14*) detected as a candidate gene in three consecutive years (Di et al. 2021).

Some key fatty acid-related genes have been identified (Okuley et al. 1994; Bocianowski et al. 2012; Kong et al. 2020). Fatty acid desaturase 2 (FAD2) affects soybean seed development by converting oleic acid to linoleic acid (Schlueter et al. 2007; Lakhssassi et al. 2017). The soybean genome contains five *FAD2* gene family members at four different loci (Bocianowski et al. 2012; Pham et al. 2010); these genes have been designated as *FAD2-1* (*FAD2-1 A* and *FAD2-1B*) and *FAD2-2* (*FAD2-2 A*, *FAD2-2B*, and *FAD2-2 C*) (Zhang et al. 2014). Genes encoding fatty acid desaturase 3 (FAD3), which modulates the linolenic acid content, are present in at least three loci in the soybean genome (Wilcox and Cavins 1985; Fehr et al. 1992; Fehr and Hammond 2000; Thapa et al. 2018). Mutations to *FAD3A*, *FAD3B*, and *FAD3C* reportedly decrease the soybean oil linolenic acid content (Bilyeu et al. 2005, 2011; Anai et al. 2005). WRINKLED1 (WRI1) is a transcription factor that helps regulate fatty acid synthesis, thereby controlling the accumulation of oil in maturing Arabidopsis seeds (Baud et al. 2009; Liu et al. 2019). A comparison with the wild-type control revealed an 80% decrease in the accumulation of triacylglycerols, which are major components of Arabidopsis seed oil, in the *wri1* mutant (Focks and Benning 1998). Three distinct *WRI1* orthologs that are differentially expressed in an organ-specific manner were isolated and characterized in soybean (Chen et al. 2018).

Among the three *GmWRI1* homologs, *GmWRI1a* is highly expressed in developing soybean seeds; its overexpression increases the total fatty acid content in seeds. Similarly, the overexpression of *WRI1* orthologs in maize increases the fatty acid content (Shen et al. 2010; Zhang et al. 2019). AINTEGUMENTA-LIKE7 (AIL7) is another transcription factor that modulates fatty acid biosynthesis and triacylglycerol accumulation in the developing seeds of transgenic Arabidopsis lines (Singer et al. 2021). The overexpression of *AIL7* significantly alters the seed fatty acid composition and decreases the accumulation of lipids in seeds.

A study on fatty acids in soybean mutants has been conducted. More specifically, Hong et al. (2019) examined the linolenic acid content of mutant soybean populations generated via gamma irradiation. Four mutant lines with the highest linolenic acid contents (33.9–67.7% higher than in the original cultivars) were selected for further analyses, which detected the increased expression of *FAD* genes during seed development as well as nucleotide polymorphisms in *FAD* genes. In the current study, we used mutant diversity pool (MDP) soybean lines that were derived from gamma irradiation (Kim et al. 2020). A GWAS was conducted to detect associations between genotypes and phenotypes, with the objective of identifying genomic regions related to the five dominant fatty acids (palmitic acid, stearic acid, oleic acid, linoleic acid, and linolenic acid) in the seeds of MDP lines.

Materials and methods

Plant materials

The seeds of MDP lines used in this study were obtained from samples irradiated with gamma rays generated using a ^{60}Co gamma irradiator (150 TBq capacity; ACEL, Ottawa, ON, Canada) at the Korea Atomic Energy Research Institute (KAERI) (Kim et al. 2020). Among the 208 MDP lines that were generated in the M_{12} generation, 192 lines exhibiting non-redundant morphological traits were subjected to GBS analysis. Six representative Korean soybean cultivars (94Seori, Bangsa, Paldal, Danbaek, Daepung, and Hwangkeum) and one landrace (KAS360-22) were utilized to construct MDP lines. Specifically, the MDP lines were composed of 5 derived from 94Seori, 6 from Bangsa, 8 from Paldal, 61 from Danbaek, 54 from Daepung, 46 from Hwangkeum, and 12 from KAS360-22. Fatty acid contents in the 192 MDP lines were determined in 2022 at the Radiation Breeding Research Farm of KAERI.

Fatty acid analysis of soybean seeds

Fatty acids were extracted as described by Ryu et al. (2017) and Kim et al. (2021), with the following modifications for a gas chromatography-mass spectrometry (GC-MS) analysis (Ryu et al. 2017; Kim et al. 2021). A powdered freeze-dried seed sample (10 mg) was mixed with 1 mL *n*-hexane for a 12 h extraction, after which 0.1 mL 2 N potassium hydroxide in methanol was added. The mixture was centrifuged for 5 min at 3,000 × g and then the supernatant was filtered using a 0.45-µm syringe filter. The fatty acid composition was analyzed using a GC-MS instrument (Plus-2010; Shimadzu, Japan) equipped with an HP-88 capillary column (60 m × 0.25 mm × 0.25 m; J&W Scientific, Folsom, CA, USA). The GC-MS conditions were as follows: ionization voltage, 70 eV; mass scan range, 50–450 mass units; injector temperature, 230 °C; detector temperature, 230 °C; injection volume, 1 µL; split ratio, 1:30; carrier gas, helium; and flow rate, 1.7 mL/min. The column temperature program was as follows: isothermal temperature of 40 °C for 5 min, increase to 180 °C at 5 °C/min, and then decrease to 28 °C at 1 °C/min. The substances present in the extracts were identified according to their retention time and the information available in a mass spectral database (NIST. 62 Library).

Principal component analysis and genome-wide association study

The contents of five fatty acids (palmitic acid, stearic acid, oleic acid, linoleic acid, and linolenic acid) in soybean seeds were used to complete a PCA. PCs with a cumulative proportion of variance exceeding 0.8 were identified using the dplyr package in the R software. A PCA plot in three dimensions was constructed for PC1, PC2, and PC3 using the ggplot2 package.

A genotyping-by-sequencing approach involving 37,249 SNPs in the MDP lines was applied (Kim et al. 2022). The Tassel 5.2 software was used for the GWAS. A total of 17,631 SNPs were filtered from the raw SNPs in the 192 MDP lines, with a minor allele frequency of 0.07. Additionally, MLM was used and a suggestive threshold *p*-value of 1.0×10^{-4} was determined as the criterion for detecting significant SNPs for the five fatty acids (Chang et al. 2018; Hu et al. 2023). On the basis of GWAS results, Manhattan and quantile-quantile plots were generated using the qqman package in the R software. Information regarding the annotated genes associated with significant SNPs and QTLs related to soybean seed fatty acids was obtained from the SoyBase database (https://www.soybase.org/).

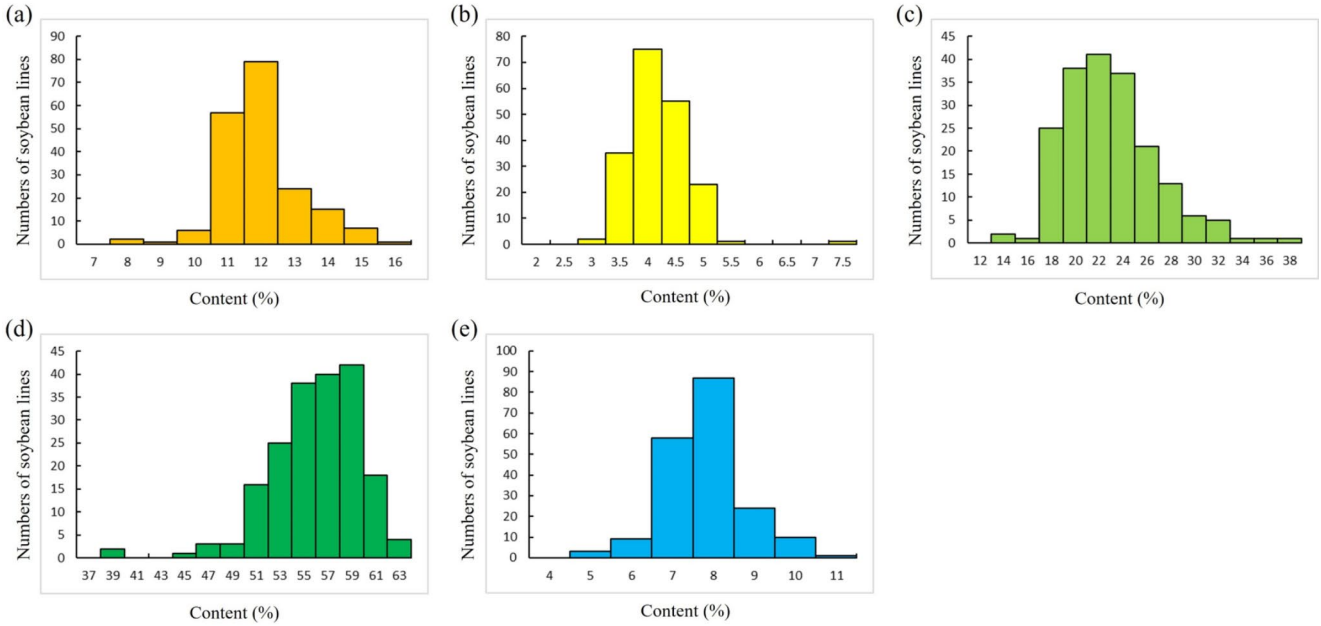

Fig. 1 Frequency distribution of fatty acid contents in soybean seeds. (**a**–**e**) correspond to palmitic acid, stearic acid, oleic acid, linoleic acid, and linolenic acid, respectively

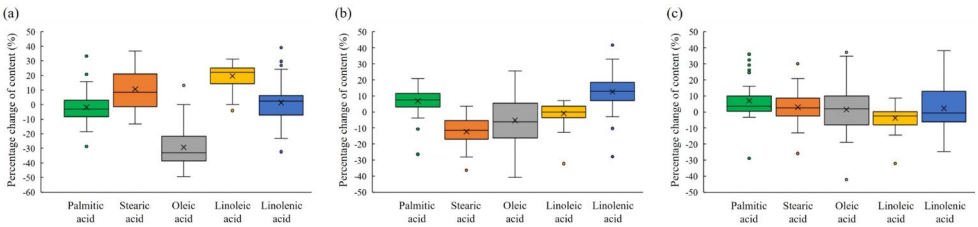

Fig. 2 Box plots of the percentage changes in fatty acid contents in 158 mutant lines. (a), (b), and (c) correspond to mutant lines derived from Danbaek, Daepung, and Hwangkeum, respectively

Results

Variations in the fatty acid composition in the mutant diversity pool

In the MDP lines, the average palmitic acid, stearic acid, oleic acid, linoleic acid, and linolenic acid contents were 11.52%, 3.93%, 21.89%, 55.01%, and 7.29%, respectively (Fig. 1a–e). Among the 192 MDP lines, 158 were derived from the Danbaek, Daepung, and Hwangkeum cultivars. The percentage change in the fatty acid content of these 158 mutant lines (relative to the fatty acid content in the corresponding wild-type cultivars) was calculated; the remaining 34 mutant lines were excluded from the analysis of percentage change because they represented a relatively small proportion of the mutants. The variations and percentage changes for these 158 mutant lines are presented in Fig. 2; Table 1. In the mutant lines derived from Danbaek, the highest and lowest increases were for linolenic acid (39.02%) and oleic acid (14.40%), respectively (Fig. 2a). The highest and lowest decreases were for oleic acid (−49.27%) and linoleic acid (−4.11%), respectively. Among the mutant lines derived from Daepung, the highest and lowest increases were for linolenic acid (41.65%) and stearic acid (3.59%), respectively (Fig. 2b), whereas the highest and lowest decreases were for oleic acid (−40.82%) and palmitic acid (−26.44%), respectively. Of the mutant lines derived from Hwangkeum, the highest increase and lowest decrease were for linolenic acid (38.26% and −24.61%, respectively) (Fig. 2c), whereas the lowest increase and highest decrease were for linoleic acid (8.58%) and oleic acid (−42.15%), respectively. In these 158 mutant lines, the highest and lowest increases in fatty acid contents (relative to the content in the original cultivars) were for linolenic acid (41.65%) and stearic acid (3.59%), respectively, in the mutant lines derived from Daepung. In contrast, the highest and lowest decreases were for oleic acid (−49.27%) and linoleic acid (−4.11%), respectively, in the mutant lines derived from Danbaek.

The results of the principal component analysis (PCA) of the five fatty acids indicated that the first three principal components (PCs) explained more than 80% of the variance, with PC1, PC2, and PC3 explaining 38.87%, 22.88%, and 18.71% of the variance, respectively. The features of

Table 1 Maximum and minimum percentage changes in fatty acid contents in 158 mutant lines

Mutant lines	Values	Palmitic acid	Stearic acid	Oleic acid	Linoleic acid	Linolenic acid
Danbaek [a] (*N* = 60)	Minimum	-28.62	-13.28	-49.27	-4.11	-32.3
	Maximum	33.27	36.86	14.4	31.23	39.02
	Mean	-1.85	10.57	-29.32	19.64	1.34
	STD	10.71	13.47	13.98	7.63	13.78
Daepung [b] (*N* = 53)	Minimum	-26.44	-36.39	-40.82	-32.27	-27.88
	Maximum	20.86	3.59	25.62	7.07	41.65
	Mean	6.89	-12.27	-5.22	-0.95	12.58
	STD	7.74	8.73	13.75	6.32	11.36
Hwangkeum [c] (*N* = 45)	Minimum	-28.96	-25.9	-42.15	-32.06	-24.61
	Maximum	37.59	29.94	37.14	8.58	38.26
	Mean	7.08	2.97	1.44	-3.83	2.22
	STD	12.43	9.64	15.25	6.5	14.24

STD : Standard deviation, N : Number of mutant lines

[a] Mutant lines from Danbaek cultivar

[b] Mutant lines from Daepung cultivar

[c] Mutant lines from Hwangkeum cultivar

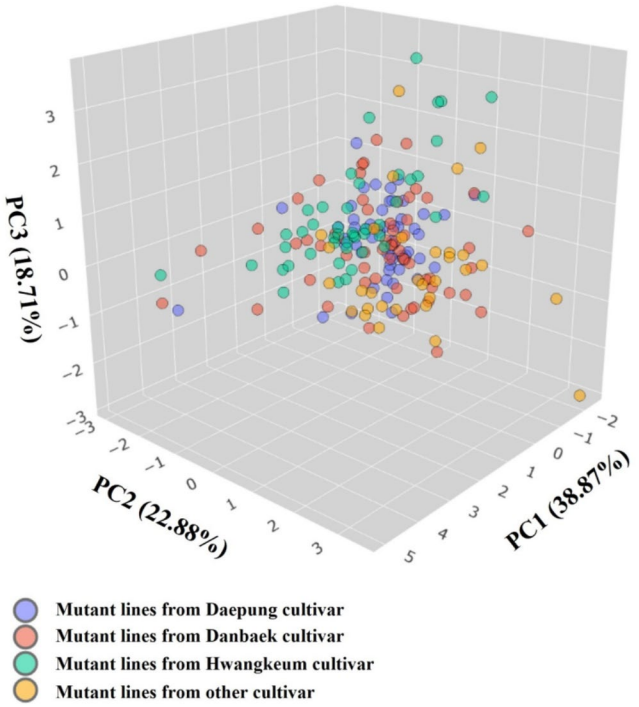

Fig. 3 PCA plot of the fatty acid contents in MDP lines

oleic acid, stearic acid, and linolenic acid were primarily represented by PC1, PC2, and PC3, respectively. A PCA plot of the first three PCs showed that mutant lines derived from the original cultivars were not clearly clustered on the basis of the contents of the five fatty acids (Fig. 3), reflecting the substantial variation in fatty acid contents among the MDP lines.

Correlation analysis of five fatty acid contents in the mutant diversity pool

Analyses of the contents of five fatty acids revealed positive and negative correlations (Fig. 4). Palmitic acid was positively correlated with stearic acid ($r = 0.21$, $p < 0.01$), linoleic acid ($r = 0.08$, not significant), and linolenic acid ($r = 0.18$, $p < 0.01$). Stearic acid and linoleic acid were positively correlated with linoleic acid ($r = 0.1$, not significant) and linolenic acid ($r = 0.25$, $p < 0.001$), respectively. Conversely, stearic acid was negatively correlated with linolenic acid ($r = -0.06$, not significant), while oleic acid was negatively correlated with the other fatty acids ($r = -0.66$ to -0.07). Of the saturated fatty acids, palmitic acid was significantly correlated with the other fatty acids, with the exception of linoleic acid, whereas stearic acid was not significantly correlated with three unsaturated fatty acids (oleic acid, linoleic acid, and linolenic acid). Among the unsaturated fatty acids, oleic acid, linoleic acid, and linolenic acid were significantly correlated with each other. Of the significant correlations, the highest negative correlation ($r = -0.66$, $p < 0.001$) was between oleic acid and linoleic acid. Moreover, the correlations between unsaturated fatty acids ($r = 0.25$, -0.66, and -0.33) were stronger than those between saturated fatty acids ($r = 0.21$). Consequently, the correlations between unsaturated fatty acids were more significant than those between saturated fatty acids.

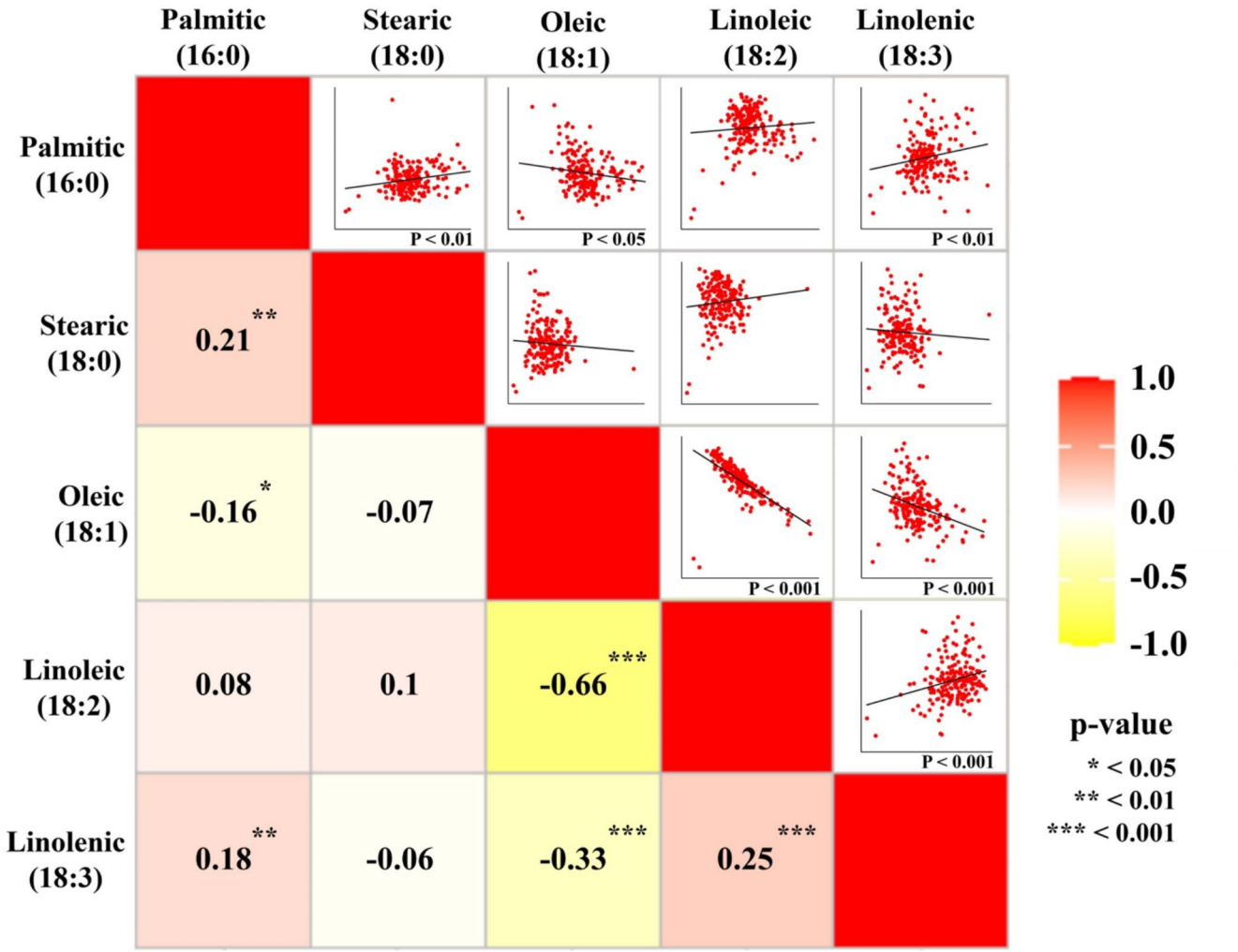

Fig. 4 Pairwise comparison matrix of fatty acids in soybean seeds on the basis of Pearson's correlation coefficient. Scatter plots of the correlation and the correlation values with significance levels are presented. Red and yellow represent positive and negative correlation coefficients, respectively. Color intensity reflects the strength of the correlation. Significance levels are indicated by asterisks: *, **, and *** represent significance at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively

Identification of candidate genes for fatty acids according to a genome-wide association study

A total of 17,631 SNP markers were filtered and used for a GWAS of the five fatty acid contents. According to the results of the GWAS using the mixed linear model (MLM), a suggestive threshold *p*-value of 1.0×10^{-4} was set as the criterion for detecting significant SNP positions in Manhattan plots (Fig. 5a–e). Among the SNPs above this threshold *p*-value, 17 significant SNPs with alleles that were associated with significant differences in the contents of the five fatty acids were selected (Table 2). Two significant SNPs for oleic acid were detected on chromosome 13 (Chr13_26366982 and Chr13_26386181). Two SNPs for palmitic acid were localized to chromosome 20 (Chr20_25980271 and Chr20_25980372). The SNPs on chromosome 13 were located in nonsynonymous and intergenic regions, whereas those on chromosome 20 were located in intergenic regions. For stearic acid, three significant SNPs, which were detected on chromosomes 7, 14, and 16, were located in nonsynonymous (Chr07_6246909), intronic (Chr14_48837019), and intergenic (Chr16_6564692) regions, respectively. One significant SNP for linolenic acid (Chr03_5370099) was detected in a synonymous region on chromosome 3. The highest number of significant SNPs was detected for linoleic acid, with 10 SNPs identified on chromosomes 1, 2, 4, 12, 13, and 15, among which four were located in intronic regions (Chr15_51160533, Chr15_51160537, Chr15_51160538, and Chr15_51160561), two were located in nonsynonymous and synonymous regions (Chr13_26366982 and Chr15_6500697, respectively), and four were located in intergenic regions (Chr01_11186648, Chr02_34309229, Chr04_46925662, and Chr12_13298963). One significant SNP (Chr13_26366982)

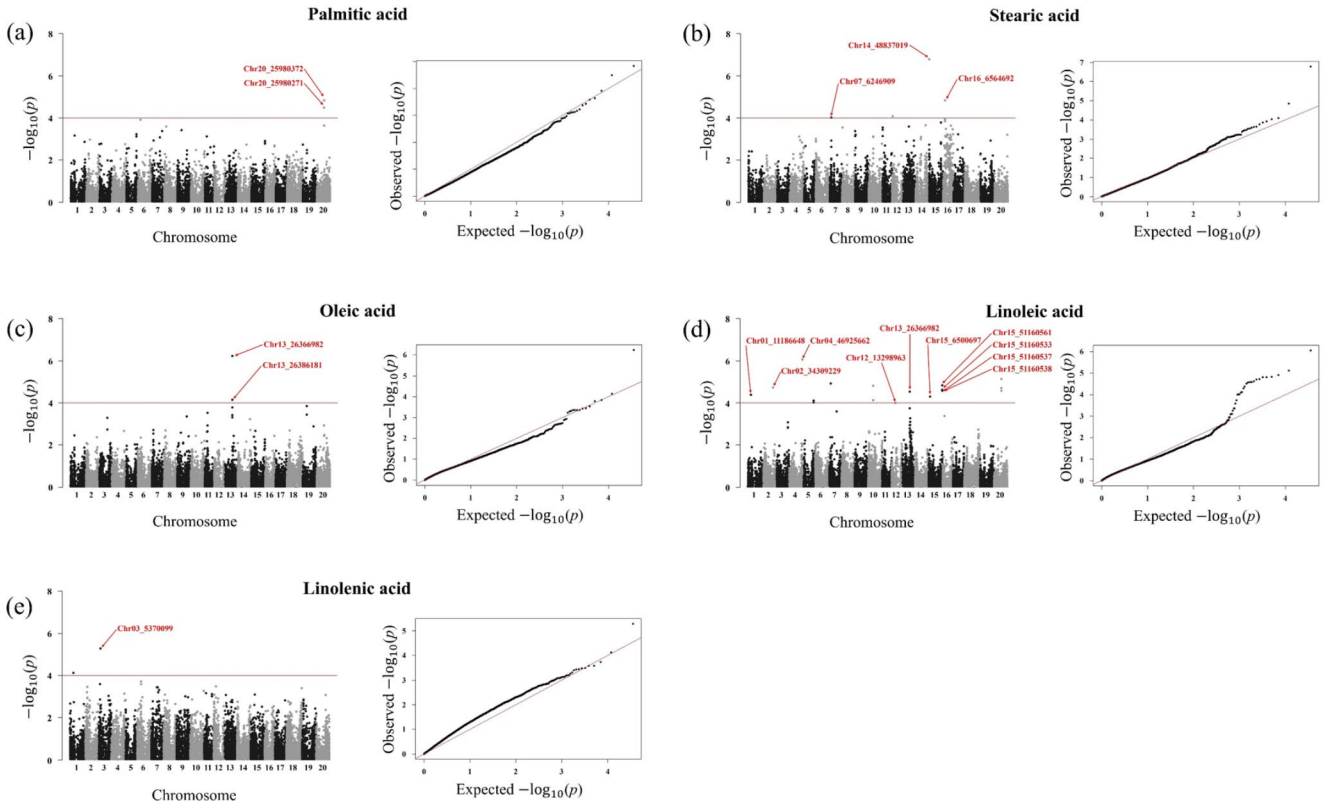

Fig. 5 Manhattan plots of genome-wide SNPs and quantile-quantile plots of *p*-values obtained from the MLM of GWAS. (a), (b), (c), (d), and (e) correspond to palmitic acid, stearic acid, oleic acid, linoleic acid, and linolenic acid, respectively

was common to oleic acid and linoleic acid. This SNP had a *p*-value (LOD) of 5.83E-07 (6.23) for oleic acid and 2.98E-05 (4.53) for linoleic acid. The highest and lowest suggestive threshold *p*-value (LOD) were 1.66E-07 (6.78) for Chr14_48837019 and 9.88E-05 (4.01) for Chr12_13298963, which were associated with stearic acid and linoleic acid, respectively.

Candidate genes related to soybean seed fatty acid contents were predicted. Table 3 lists nine significant SNPs located within intragenic regions along with the corresponding gene annotation information. Among these SNPs, four (Chr15_51160533, Chr15_51160537, Chr15_51160538, and Chr15_51160561) were associated with linoleic acid and corresponded to the gene *Glyma.15g274000*. Of the remaining five SNPs, Chr03_5370099, Chr07_6246909, Chr14_48837019, and Chr15_6500697 were associated with three fatty acids (stearic acid, linoleic acid, and linolenic acid) and corresponded to genes *Glyma.03g042500*, *Glyma.07g069200*, *Glyma.14g223100*, and *Glyma.15g084700*, respectively. The SNP (Chr13_26366982) common to oleic acid and linoleic acid corresponded to the gene *Glyma.13g150200*.

Allelic effects on fatty acid contents

The significant differences in fatty acid contents among the alleles of 17 significant SNPs are illustrated in the box plots in Figs. 6 and 7. The number of alleles for these 17 SNPs varied from two to three. In the box plots for saturated fatty acids (palmitic acid and stearic acid), three of the five SNPs (Chr14_48837019, Chr20_25980271, and Chr20_25980272) were compared between two alleles (one homozygous and one heterozygous), which revealed significant differences (Fig. 6a, b, and d). The remaining two SNPs (Chr07_6246909 and Chr16_6564692) were compared among three alleles (two homozygous and one heterozygous). One SNP (Chr16_6564692) had significant differences among all three alleles. For another SNP (Chr07_6246909), the AA allele differed significantly from the CC and CA alleles, but there was no significant difference between the CC and CA alleles (Fig. 6c and e).

In the box plots for unsaturated fatty acids (oleic acid, linoleic acid, and linolenic acid), seven SNPs (Chr01_11186648, Chr12_13298963, Chr15_6500697, Chr15_51160533, Chr15_51160537, Chr15_51160538, and Chr15_51160561) were compared between two alleles (one homozygous and one heterozygous), which detected significant differences (Fig. 7c, f, h, i, j, k, and l). Two

Table 2 Significant fatty acid-associated SNPs detected by GWAS

Trait	SNP ID	Chr.	Physical Position (bp)	Allele	Allele Frequency	*p*-Value (LOD)	SNP Location
Palmitic acid	Chr20_25980271	20	25,980,271	CC	11 (8.08%)	3.29E-05 (4.48)	Intergenic
				TC	125 (91.91%)		
	Chr20_25980372	20	25,980,372	AA	106 (80.30%)	1.48E-05 (4.83)	Intergenic
				TA	26 (19.69%)		
Stearic acid	Chr07_6246909	7	6,246,909	AA	4 (3.05%)	9.41E-05 (4.03)	Nonsynonymous
				CC	9 (6.87%)		
				CA	118 (90.07%)		
	Chr14_48837019	14	48,837,019	AA	67 (55.83%)	1.66E-07 (6.78)	Intron
				GA	53 (44.16%)		
	Chr16_6564692	16	6,564,692	CC	110 (88%)	1.44E-05 (4.84)	Intergenic
				TT	10 (8%)		
				CT	5 (4%)		
Oleic acid	Chr13_26366982	13	26,366,982	AA	27 (15.16%)	5.85E-07 (6.23)	Nonsynonymous
				GG	147 (82.58%)		
				AG	4 (2.24%)		
	Chr13_26386181	13	26,386,181	AA	98 (81.66%)	7.17E-05 (4.14)	Intergenic
				GG	22 (18.33%)		
Linoleic acid	Chr01_11186648	1	11,186,648	GG	113 (83.08%)	4.09E-05 (4.39)	Intergenic
				AG	23 (16.91%)		
	Chr02_34309229	2	34,309,229	CC	23 (15.97%)	1.84E-05 (4.73)	Intergenic
				TT	6 (4.16%)		
				CT	115 (79.86%)		
	Chr04_46925662	4	46,925,662	AA	18 (15.25%)	8.85E-07 (6.05)	Intergenic
				GG	100 (84.74%)		
	Chr12_13298963	12	13,298,963	CC	21 (18.42%)	9.88E-05 (4.01)	Intergenic
				TC	93 (81.57%)		
	Chr13_26366982	13	26,366,982	AA	27 (15.16%)	2.98E-05 (4.53)	Nonsynonymous
				GG	147 (82.58%)		
				AG	4 (2.24%)		
	Chr15_6500697	15	6,500,697	CC	16 (13.22%)	4.99E-05 (4.30)	Synonymous
				CT	105 (86.77%)		
	Chr15_51160533	15	51,160,533	TT	96 (66.21%)	2.40E-05 (4.62)	Intron
				CT	49 (33.79%)		
	Chr15_51160537	15	51,160,537	GG	105 (69.07%)	2.60E-05 (4.58)	Intron
				AG	47 (30.92%)		
	Chr15_51160538	15	51,160,538	GG	108 (71.52%)	2.65E-05 (4.58)	Intron
				TG	43 (28.47%)		
	Chr15_51160561	15	51,160,561	GG	111 (70.25%)	1.46E-05 (4.84)	Intron
				AG	47 (29.74%)		
Linolenic acid	Chr03_5370099	3	5,370,099	GG	24 (17.77%)	5.25E-06 (5.28)	Synonymous
				TT	106 (78.51%)		
				TG	5 (3.7%)		

other SNPs (Chr13_26386181 and Chr04_46925662) were compared between two homozygous alleles, which revealed significant differences (Fig. 7b and e). The SNP (Chr13_26366982) common to oleic acid and linoleic acid was compared among three alleles (two homozygous and one heterozygous). For both of these fatty acids, the AA allele differed significantly from the GG and AG alleles, which was in contrast to the lack of significant difference between the GG and AG alleles (Fig. 7a and g). For linoleic acid and linolenic acid, two SNPs (Chr02_34309229 and Chr03_5370099) were compared among three alleles (two homozygous and one heterozygous), with significant differences detected between one homozygous allele and the heterozygous allele (Fig. 7d and m).

Table 3 Genes corresponding to SNPs in intragenic regions detected by GWAS

Trait	SNP ID	Candidate gene	Start (bp)	Stop (bp)	Annotation
Stearic acid	Chr07_6246909	*Glyma.07g069200*	6,246,770	6,247,246	ARID/BRIGHT DNA-binding domain
	Chr14_48837019	*Glyma.14g223100*	48,827,645	48,842,663	Ubiquitin-like superfamily protein
Oleic acid	Chr13_26366982	*Glyma.13g150200*	26,365,922	26,371,905	KDO transferase
Linoleic acid	Chr13_26366982	*Glyma.13g150200*	26,365,922	26,371,905	KDO transferase
	Chr15_6500697	*Glyma.15g084700*	6,499,058	6,503,126	Pentatricopeptide repeat (PPR) superfamily protein
	Chr15_51160533 Chr15_51160537 Chr15_51160538 Chr15_51160561	*Glyma.15g274000*	51,156,044	51,170,334	Telomerase reverse transcriptase
Linolenic acid	Chr03_5370099	*Glyma.03g042500*	5,367,447	5,370,281	Iron regulatory protein 1

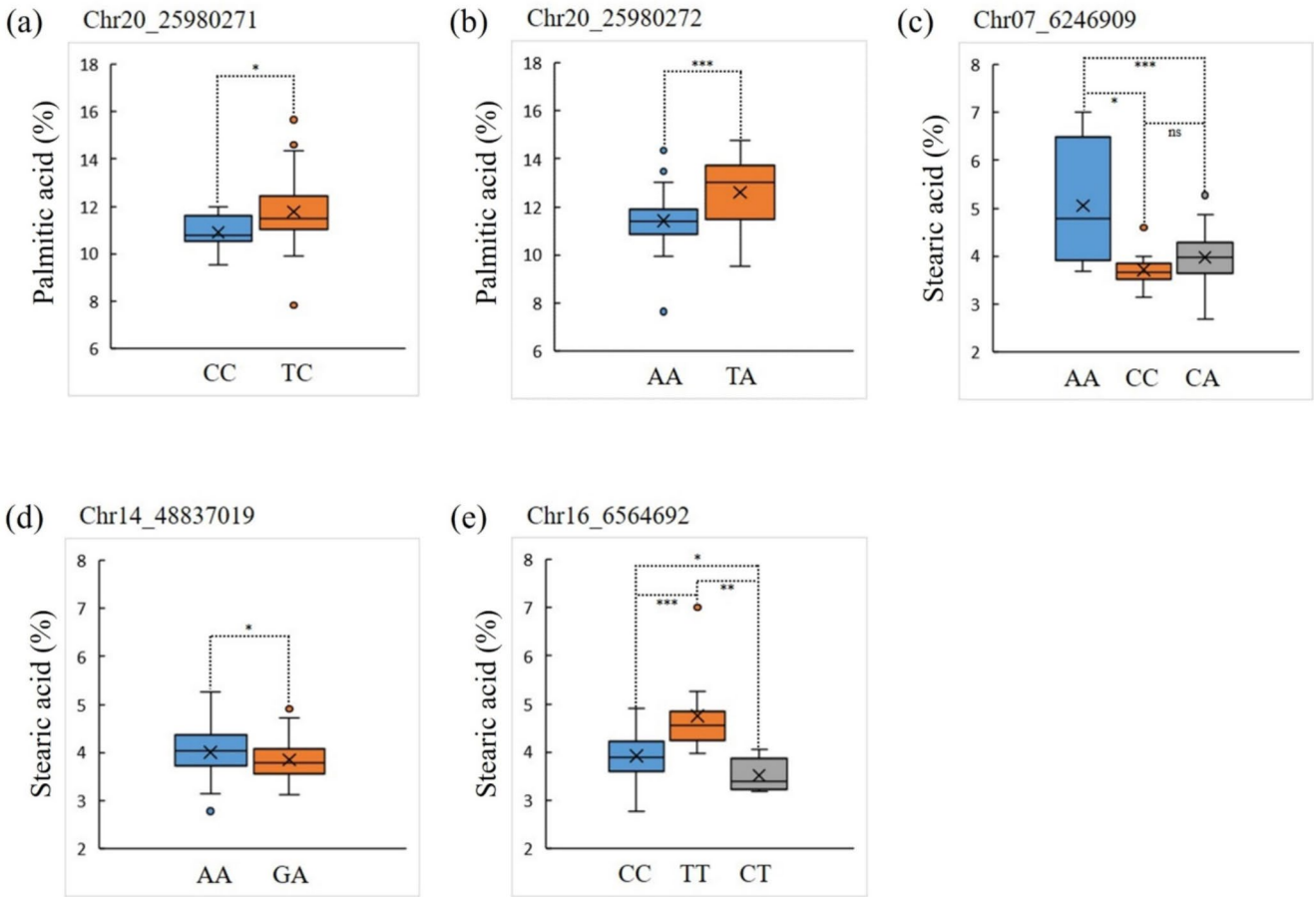

Fig. 6 Box plots of palmitic acid (**a** and **b**) and stearic acid (**c**–**e**) contents for the alleles of selected SNPs. Significance levels are indicated by asterisks: *, **, and *** represent significance at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively. ns indicates not significant

Discussion

The fatty acid composition of soybean is typically as follows: 12% palmitic acid, 4% stearic acid, 23% oleic acid, 53% linoleic acid, and 8% linolenic acid (Lee et al. 2007). A similar fatty acid composition was detected in the MDP lines used in this study (11.52% palmitic acid, 3.93% stearic acid, 21.89% oleic acid, 55.01% linoleic acid, and 7.29% linolenic acid). However, the contents of five fatty acids in the analyzed mutants differed to varying degrees from the corresponding contents in the original cultivars (Danbaek, Daepung, and Hwangkeum). The percentage changes in the

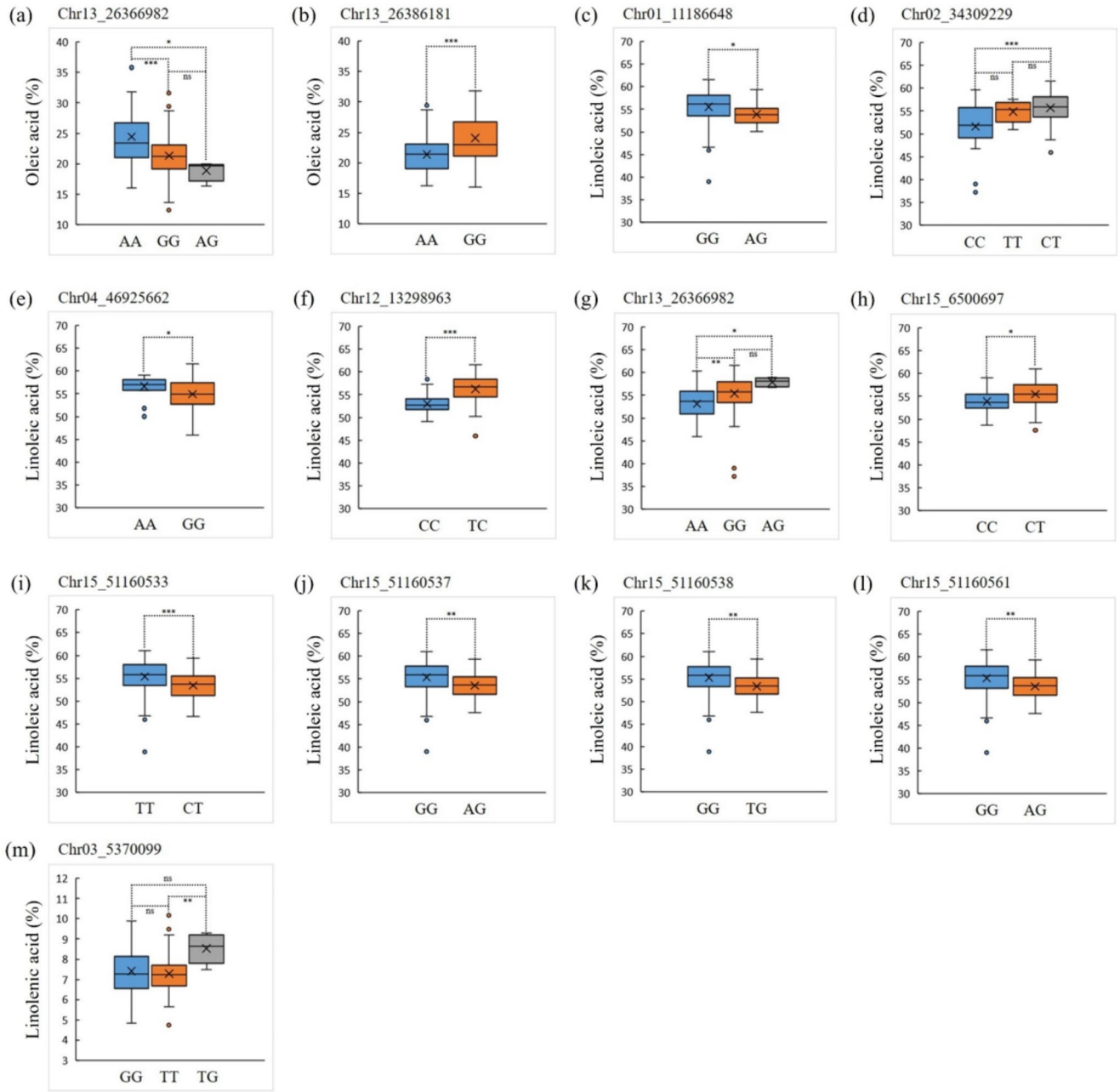

Fig. 7 Box plots of oleic acid (**a** and **b**), linoleic acid (**c**–**l**), and linolenic acid (**m**) contents for the alleles of selected SNPs. Significance levels are indicated by asterisks: *, **, and *** represent significance at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively. ns indicates not significant

fatty acid contents of the mutant lines ranged from − 26.44 to 37.59% for palmitic acid, − 13.28–36.86% for stearic acid, − 49.27–37.14% for oleic acid, − 32.27–31.23% for linoleic acid, and − 32.3–41.65% for linolenic acid (Fig. 2; Table 1). In an earlier study, Hong et al. (2019) observed that the linolenic acid content in seeds was 33.9–67.7% higher in mutant lines than in the original cultivars. Similar to these findings, in the present study, the seed fatty acid composition differed between the selected mutant lines and the original cultivars, suggesting selective breeding may optimize fatty acid profiles via genotype-specific effects.

The PCA results indicated that the features of oleic acid, stearic acid, and linolenic acid were reflected by the first three PCs, which accounted for more than 80% of the total variance. According to the PCA of fatty acid contents, the mutant soybean lines included in this study were widely distributed, indicative of the diversity among the MDP lines derived from different cultivars. Therefore, the MDP lines

used in this study may be useful for analyzing soybean seed fatty acid compositions on the basis of PCA.

In terms of the correlations among fatty acid contents, palmitic acid was significantly correlated with the other fatty acids (except for linoleic acid). In addition, the correlations between unsaturated fatty acids were significant. Among the unsaturated fatty acids, oleic acid and linoleic acid were the most highly correlated ($r = -0.66$). There was also a highly negative correlation between oleic acid and linolenic acid ($r = -0.33$), which was in contrast to the positive correlation between linoleic acid and linolenic acid ($r = 0.25$). Similarly, previous studies reported that the oleic acid content is negatively correlated with the linoleic acid content ($r = -0.96$) (Di et al. 2021; Wang et al. 2020). However, the correlation between palmitic acid and oleic acid was positive ($r = 0.42$) in an earlier study (Wang et al. 2020), unlike the negative correlation revealed in the current study ($r = -0.16$). Thus, the correlations among soybean fatty acid contents determined in this study may be relevant to breeders interested in modulating soybean seed fatty acid contents.

The GWAS conducted in the present study detected significant SNPs in novel regions and QTL regions associated with soybean seed fatty acid contents. To screen for potential candidate genes, we identified nine significant SNPs located within intragenic regions among the 17 significant SNPs associated with soybean seed fatty acids (except for palmitic acid) (Table 3). These nine SNPs corresponded to six genes (*Glyma.03g042500*, *Glyma.07g069200*, *Glyma.13g150200*, *Glyma.14g223100*, *Glyma.15g084700*, and *Glyma.15g274000*). Specifically, seven SNPs were related to unsaturated fatty acids (oleic acid, linoleic acid, and linolenic acid), whereas two SNPs were related to a saturated fatty acid (stearic acid). According to the QTLs for fatty acids obtained from the SoyBase database, one SNP (Chr03_5370099) related to linolenic acid, which corresponded to the gene *Glyma.03g042500*, was located within the reported QTL region (*Seed linolenic 14 − 6*) associated with linolenic acid (Fan et al. 2015). This gene encodes iron regulatory protein 1 (IRP1). Iron is essential for various plant physiological processes and may be important for enzymatic activities (Briat et al. 2010). Among the five examined fatty acids, oleic acid and linoleic acid were the most highly correlated and had a common SNP (Chr13_26366982). This SNP corresponded to the gene *Glyma.13g150200*, which encodes 3-deoxy-D-manno-oct-ulosonic acid (KDO) transferase, implying it is involved in transferring KDO into cell wall rhamnogalacturonan II, a complex polysaccharide component of pectin in the primary cell walls of plants (Séveno et al. 2010). Most significant SNPs were located within a single gene. Notably, four SNPs (Chr15_51160533, Chr15_51160537, Chr15_51160538, and Chr15_51160561) related to linoleic acid corresponded to a single gene (*Glyma.15g274000*) encoding telomerase reverse transcriptase (TERT). Earlier research indicated TERT is the catalytic subunit of telomerase and is necessary for enzyme activity (Heller-Uszynska et al. 2002; Oguchi et al. 1999). In tobacco, specific *TERT* transcription and telomerase activity may be enhanced in reproductive tissues (Jurečková et al. 2017). These three genes that were predicted to be related to oleic acid, linoleic acid, and linolenic acid contents may contribute to fatty acid synthesis in soybean seeds by influencing the iron content, a complex polysaccharide component of pectin in primary cell walls, and telomerase activity in reproductive tissues. The functional validation of these genes can provide insights into the genetic mechanisms underlying fatty acid biosynthesis, with possible implications for improving soybean seed fatty acid compositions through genetic modifications.

Box plot analyses of allelic effects for 17 significant SNPs detected significant differences in fatty acid contents among various SNP alleles. For the saturated fatty acids (palmitic acid and stearic acid), analyses of five SNPs on chromosomes 7, 14, 16, and 20 revealed significant differences between two alleles (one homozygous and one heterozygous) and among three alleles (two homozygous and one heterozygous). For the unsaturated fatty acids (oleic acid, linoleic acid, and linolenic acid), the comparison of the alleles of 12 SNPs on chromosome 1, 2, 3, 4, 12, 13, and 15 detected significant differences in fatty acid contents between two homozygous alleles, between one homozygous allele and one heterozygous allele, and among three alleles (two homozygous and one heterozygous). The largest and smallest significant differences in fatty acid contents due to SNP alleles were 5.53% (linolenic acid) and 0.4% (stearic acid), respectively (Figs. 6e and 7a). A previous genomic analysis conducted to identify SNPs associated with fatty acid components in soybean seeds showed that specific alleles contribute to decreases and increases in palmitic acid and oleic acid contents, respectively (Priolli et al. 2019). Additional genotyping is needed to verify the distinct allelic effects on the fatty acid contents of the MDP lines.

Zhao et al. (2019a) examined the saturated fatty acid (palmitic acid and stearic acid) contents in soybean seeds for a GWAS in three environments. The variations in the two fatty acid contents among the three tested environments were significant, but the skewness and kurtosis of the two fatty acid contents were less than ± 1. This reflects continuous variation and a near-normal distribution, indicative of the reliability of the analyses and the relative lack of extreme values or anomalous patterns. In another study, Zhao et al. (2019b) conducted a GWAS of unsaturated fatty acid (oleic acid, linoleic acid, and linolenic acid) contents in three environments. On the basis of the coefficient of

variation (1.1–4.12%), there were no significant differences in the contents of the three fatty acids among the three tested environments. These studies indicate that the fatty acid contents of soybean seeds in multiple environments were reliable and consistent. The present study analyzed the contents of five fatty acids in soybean seeds in a single environment, but the GWAS results for the five fatty acid contents were appropriate for predicting candidate genes because earlier studies showed that the variation in fatty acid contents in soybean seeds across various environments is either not significant or is continuous (Zhao et al. 2019a, b). However, Di et al. (2021) performed a GWAS for linoleic acid and detected a candidate gene that was correlated with the linoleic acid content in three environments. Therefore, additional analyses of the five fatty acid contents in the seeds of the MDP lines are required to determine whether the same or novel significant SNPs will be identified by GWAS in other environments.

Conclusions

The present study detected significant phenotypic variations and genetic associations underlying the fatty acid composition of soybean seeds in MDP lines. The mutant seeds and the seeds of the original cultivars varied regarding fatty acid composition. According to the PCA data, the features of oleic acid, stearic acid, and linolenic acid were indicated by the first three PCs. Additionally, the mutant soybean lines were not clustered regardless of the original cultivars. Palmitic acid was significantly correlated with the other fatty acids (with the exception of linoleic acid). Moreover, significant correlations between unsaturated fatty acids were detected. The highest correlation was observed between oleic acid and linoleic acid ($r = -0.66$). Among the 17 significant SNPs, nine were located within intragenic regions and were associated with fatty acids in soybean seeds. These nine SNPs corresponded to six genes (*Glyma.03g042500*, *Glyma.07g069200*, *Glyma.13g150200*, *Glyma.14g223100*, *Glyma.15g084700*, and *Glyma.15g274000*). Three of these genes (*Glyma.03g042500*, *Glyma.13g150200*, and *Glyma.15g274000*) were designated as candidate genes related to oleic acid, linoleic acid, and linolenic acid contents. The allelic effects of these 17 significant SNPs on seed fatty acid contents were visualized in box plots. These findings may be exploited by soybean breeders to optimize fatty acid compositions and oil quality.

Acknowledgements The authors express gratitude to the editor and reviewers for constructive feedback and valuable suggestions that significantly enhanced the quality of this manuscript.

Author contributions Jeong Woo Lee and Jung Min Kim wrote the text and performed the experiment. Dae June Kim and Ji Su Seo helped to draft the manuscript. All authors edited and approved the manuscript for publication.

Funding This work was supported by the research program of the Korea Atomic Energy Research Institute (Project No. 523420-25).

Declarations

Conflict of interest The authors declare no conflicts of interest relevant to this research.

Ethical approval Not applicable.

Informed consent Not applicable.

References

Akond M, Liu S, Boney M, Kantartzi SK, Meksem K, Bellaloui N, Lightfoot DA, Kassem MA (2014) Identification of quantitative trait loci (QTL) underlying protein, oil, and five major fatty acids' contents in soybean. Am J Plant Sci. https://doi.org/10.4236/ajps.2014.51021

Anai T, Yamada T, Kinoshita T, Rahman SM, Takagi Y (2005) Identification of corresponding genes for three low-α-linolenic acid mutants and elucidation of their contribution to fatty acid biosynthesis in soybean seed. Plant Sci 168:1615–1623. https://doi.org/10.1016/j.plantsci.2005.02.016

Baud S, Wuilleme S, To A, Rochat C, Lepiniec L (2009) Role of WRINKLED1 in the transcriptional regulation of glycolytic and fatty acid biosynthetic genes in Arabidopsis. Plant J 60:933–947. https://doi.org/10.1111/j.1365-313X.2009.04011.x

Bilyeu K, Palavalli L, Sleper D, Beuselinck P (2005) Mutations in soybean microsomal omega-3 fatty acid desaturase genes reduce linolenic acid concentration in soybean seeds. Crop Sci 45:1830–1836. https://doi.org/10.2135/cropsci2004.0632

Bilyeu K, Gillman JD, LeRoy AR (2011) Novel FAD3 mutant allele combinations produce soybeans containing 1% linolenic acid in the seed oil. Crop Sci 51:259–264. https://doi.org/10.2135/cropsci2010.01.0044

Bocianowski J, Mikołajczyk K, Bartkowiak-Broda I (2012) Determination of fatty acid composition in seed oil of rapeseed (*Brassica napus* L.) by mutated alleles of the FAD3 desaturase genes. J Appl Genet 53:27–30. https://doi.org/10.1007/s13353-011-0062-0

Briat J-F, Duc C, Ravet K, Gaymard F (2010) Ferritins and iron storage in plants. Biochim Biophys Acta Gen Subj 1800:806–814. https://doi.org/10.1016/j.bbagen.2009.12.003

Chang F, Guo C, Sun F, Zhang J, Wang Z, Kong J, He Q, Sharmin RA, Zhao T (2018) Genome-wide association studies for dynamic plant height and number of nodes on the main stem in summer sowing soybeans. Front Plant Sci 9:1184. https://doi.org/10.3389/fpls.2018.01184

Chen L, Zheng Y, Dong Z, Meng F, Sun X, Fan X, Zhang Y, Wang M, Wang S (2018) Soybean (*Glycine max*) WRINKLED1 transcription factor, GmWRI1a, positively regulates seed oil accumulation. Mol Genet Genomics 293:401–415. https://doi.org/10.1007/s00438-017-1393-2

Chiriacò MV, Galli N, Santini M, Rulli MC (2024) Deforestation and greenhouse gas emissions could arise when replacing palm oil with other vegetable oils. Sci Total Environ 914:169486. https://doi.org/10.1016/j.scitotenv.2023.169486

Di Q, Piersanti A, Zhang Q, Miceli C, Li H, Liu X (2021) Genome-wide association study identifies candidate genes related to the

linoleic acid content in soybean seeds. Int J Mol Sci 23:454. https://doi.org/10.3390/ijms23010454
Fan S, Li B, Yu F, Han F, Yan S, Wang L, Sun J (2015) Analysis of additive and epistatic quantitative trait loci underlying fatty acid concentrations in soybean seeds across multiple environments. Euphytica 206:689–700. https://doi.org/10.1007/s10681-015-1491-3
FAO (2019) FAOSTAT. Food and Agriculture Organization of the United Nations. http://www.fao.org/faostat/en/. Accessed 20 Aug 2024
FDA (2018) FDA Completes Review of Qualified Health Claim Petition for Oleic Acid and the Risk of Coronary Heart Disease. https://www.fda.gov/food/cfsan-constituent-updates/fda-completes-review-qualified-health-claim-petition-oleic-acid-and-risk-coronary-heart-disease. Accessed 19 Sept 2024
Fehr WR, Hammond EG (2000) Reduced linolenic acid production in soybeans. Google Patents
Fehr WR, Welke GA, Hammond EG, Duvick DN, Cianzio SR (1992) Inheritance of reduced linolenic acid content in soybean genotypes A16 and A17. Crop Sci 32:903–906. https://doi.org/10.2135/cropsci1992.0011183X003200040013x
Focks N, Benning C (1998) wrinkled1: a novel, low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate metabolism. Plant Physiol 118:91–101. https://doi.org/10.1104/pp.118.1.91
Heller-Uszynska K, Schnippenkoetter W, Kilian A (2002) Cloning and characterization of rice (*Oryza sativa* L.) telomerase reverse transcriptase, which reveals complex splicing patterns. Plant J 31:75–86. https://doi.org/10.1046/j.1365-313X.2001.01337.x
Hong MJ, Jang YE, Kim DG, Kim JM, Lee MK, Kim JB, Eom SH, Ha BK, Lyu JI, Kwon SJ (2019) Selection of mutants with high linolenic acid contents and characterization of fatty acid desaturase 2 and 3 genes during seed development in soybean (*Glycine max*). J Sci Food Agric 99:5384–5391. https://doi.org/10.1002/jsfa.9798
Hu C, Li J, Liu J, Zhang D, Jin L, Yang N, Bai B, Wang Z, Feng S, Ru Z (2023) Genome-Wide Association Study on Seedling Phenotypic Traits of Wheat under Different Nitrogen Conditions. Plants 12:4050. https://doi.org/10.3390/plants12234050
Hyten DL, Pantalone VR, Saxton AM, Schmidt ME, Sams CE (2004) Molecular mapping and identification of soybean fatty acid modifier quantitative trait loci. J Am Oil Chem Soc 81:1115–1118. https://doi.org/10.1007/s11746-004-1027-z
Ivanov DS, Lević JD, Sredanović SA (2010) Fatty acid composition of various soybean products. Food Feed Res 37:65–70-65-70
Jurečková JF, Sýkorová E, Hafidh S, Honys D, Fajkus J, Fojtová M (2017) Tissue-specific expression of telomerase reverse transcriptase gene variants in Nicotiana tabacum. Planta 245:549–561. https://doi.org/10.1007/s00425-016-2624-1
Kim D-G, Lyu JI, Lee M-K, Kim JM, Hung NN, Hong MJ, Kim J-B, Bae C-H, Kwon S-J (2020) Construction of soybean mutant diversity pool (MDP) lines and an analysis of their genetic relationships and associations using TRAP markers. Agronomy 10:253. https://doi.org/10.3390/agronomy10020253
Kim D-G, Lyu J-I, Lim Y-J, Kim J-M, Hung N-N, Eom S-H, Kim S-H, Kim J-B, Bae C-H, Kwon S-J (2021) Differential gene expression associated with altered isoflavone and fatty acid contents in soybean mutant diversity pool. Plants 10:1037. https://doi.org/10.3390/plants10061037
Kim D-G, Lyu JI, Kim JM, Seo JS, Choi H-I, Jo YD, Kim SH, Eom SH, Ahn J-W, Bae C-H (2022) Identification of loci governing agronomic traits and mutation hotspots via a gbs-based genome-wide association study in a soybean mutant diversity pool. Int J Mol Sci 23:10441. https://doi.org/10.3390/ijms231810441
Kinney A, Clemente T (2005) Modifying soybean oil for enhanced performance in biodiesel blends. Fuel Process Technol 86:1137–1147. https://doi.org/10.1016/j.fuproc.2004.11.008
Kong Q, Yang Y, Guo L, Yuan L, Ma W (2020) Molecular basis of plant oil biosynthesis: insights gained from studying the WRINKLED1 transcription factor. Front Plant Sci 11:24. https://doi.org/10.3389/fpls.2020.00024
Lakhssassi N, Zhou Z, Liu S, Colantonio V, AbuGhazaleh A, Meksem K (2017) Characterization of the FAD2 gene family in soybean reveals the limitations of gel-based TILLING in genes with high copy number. Front Plant Sci 8:324. https://doi.org/10.3389/fpls.2017.00324
Lee J-D, Bilyeu KD, Shannon JG (2007) Genetics and breeding for modified fatty acid profile in soybean seed oil. J Crop Sci Biotechnol 10:201–210
Liu H-R, White PJ (1992) Oxidative stability of soybean oils with altered fatty acid compositions. J Am Oil Chem Soc 69:528–532. https://doi.org/10.1007/BF02636103
Liu W, Inbaraj BS, Chen B (2007) Analysis and formation of trans fatty acids in hydrogenated soybean oil during heating. Food Chem 104:1740–1749. https://doi.org/10.1016/j.foodchem.2006.10.069
Liu H, Zhai Z, Kuczynski K, Keereetaweep J, Schwender J, Shanklin J (2019) WRINKLED1 regulates BIOTIN ATTACHMENT DOMAIN-CONTAINING proteins that inhibit fatty acid synthesis. Plant Physiol 181:55–62. https://doi.org/10.1104/pp.19.00587
Messina M, Shearer G, Petersen K (2021) Soybean oil lowers circulating cholesterol levels and coronary heart disease risk, and has no effect on markers of inflammation and oxidation. Nutrition 89:111343. https://doi.org/10.1016/j.nut.2021.111343
Meydani SN, Lichtenstein AH, White PJ, Goodnight SH, Elson CE, Woods M, Gorbach SL, Schaefer EJ (1991) Food use and health effects of soybean and sunflower oils. J Am Coll Nutr 10:406–428. https://doi.org/10.1080/07315724.1991.10718168
Napolitano GE, Ye Y, Cruz-Hernandez C (2018) Chemical characterization of a high-oleic soybean oil. J Am Oil Chem Soc 95:583–589. https://doi.org/10.1002/aocs.12049
Narkhede N, Patel A (2013) Biodiesel production by esterification of oleic acid and transesterification of soybean oil using a new solid acid catalyst comprising 12-tungstosilicic acid and zeolite Hβ. Ind Eng Chem Res 52:13637–13644. https://doi.org/10.1021/ie402230v
Oguchi K, Liu H, Tamura K, Takahashi H (1999) Molecular cloning and characterization of AtTERT, a telomerase reverse transcriptase homolog in Arabidopsis thaliana. FEBS Lett 457:465–469. https://doi.org/10.1016/S0014-5793(99)01083-2
Okuley J, Lightner J, Feldmann K, Yadav N, Lark E, Browse J (1994) Arabidopsis FAD2 gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. Plant Cell 6:147–158. https://doi.org/10.1105/tpc.6.1.147
Panthee D, Pantalone V, Saxton A (2006) Modifier QTL for fatty acid composition in soybean oil. Euphytica 152:67–73. https://doi.org/10.1007/s10681-006-9179-3
Pham A-T, Lee J-D, Shannon JG, Bilyeu KD (2010) Mutant alleles of FAD2-1A and FAD2-1B combine to produce soybeans with the high oleic acid seed oil trait. BMC Plant Biol 10:1–13. https://doi.org/10.1186/1471-2229-10-195
Phuah E-T, Yap JW-L, Lau C-W, Lee Y-Y, Tang T-K (2022) Vegetable oils and animal fats: sources, properties and recovery. Recent advances in edible fats and oils technology: Processing, health implications, economic and environmental impact, Singapore, pp 1–26. https://doi.org/10.1007/978-981-16-5113-7_1
Poth U (2000) Drying oils and related products. Ullmann's Encyclopedia Industrial Chem. https://doi.org/10.1002/14356007.a09_055
Priolli R, Carvalho C, Bajay M, Pinheiro JB, Vello NA (2019) Genome analysis to identify SNPs associated with oil content and fatty acid components in soybean. Euphytica 215:54. https://doi.org/10.1007/s10681-019-2378-5
Romero A, Cuesta C, Sánchez-Muniz FJ (2000) Trans fatty acid production in deep fat frying of frozen foods with different oils and

frying modalities. Nutr Res 20:599–608. https://doi.org/10.1016/S0271-5317(00)00150-0

Ryu J, Kwon S-J, Ahn J-W, Jo YD, Kim SH, Jeong SW, Lee MK, Kim J-B, Kang S-Y (2017) Phytochemicals and antioxidant activity in the kenaf plant (*Hibiscus cannabinus* L.). J Plant Biotechnol 44:191–202. https://doi.org/10.5010/JPB.2017.44.2.191

Schlueter JA, Vasylenko-Sanders IF, Deshpande S, Yi J, Siegfried M, Roe BA, Schlueter SD, Scheffler BE, Shoemaker RC (2007) The FAD2 gene family of soybean: insights into the structural and functional divergence of a paleopolyploid genome. Crop Sci 47. https://doi.org/10.2135/cropsci2006.06.0382tpg

Séveno M, Séveno-Carpentier E, Voxeur A, Menu-Bouaouiche L, Rihouey C, Delmas F, Chevalier C, Driouich A, Lerouge P (2010) Characterization of a putative 3-deoxy-D-manno-2-octulosonic acid (Kdo) transferase gene from Arabidopsis thaliana. Glycobiology 20:617–628. https://doi.org/10.1093/glycob/cwq011

Shahbandeh M (2024) Global Vegetable Oil Consumption, 2019/20. https://www.statista.com/statistics/263937/vegetable-oils-global-consumption/. Accessed 20 Aug 2024

Shen B, Allen WB, Zheng P, Li C, Glassman K, Ranch J, Nubel D, Tarczynski MC (2010) Expression of ZmLEC1 and ZmWRI1 increases seed oil production in maize. Plant Physiol 153:980–987. https://doi.org/10.1104/pp.110.157537

Singer SD, Jayawardhane KN, Jiao C, Weselake RJ, Chen G (2021) The effect of AINTEGUMENTA-LIKE 7 over-expression on seed fatty acid biosynthesis, storage oil accumulation and the transcriptome in Arabidopsis thaliana. Plant Cell Rep 40:1647–1663. https://doi.org/10.1007/s00299-021-02715-3

Sung M, Van K, Lee S, Nelson R, LaMantia J, Taliercio E, McHale LK, Mian MR (2021) Identification of SNP markers associated with soybean fatty acids contents by genome-wide association analyses. Mol Breed 41:1–16. https://doi.org/10.1007/s11032-021-01216-1

Thapa R, Carrero-Colón M, Addo-Quaye C, Held J, Dilkes B, Hudson KA (2018) New alleles of FAD3A lower the linolenic acid content of soybean seeds. Crop Sci 58:713–718. https://doi.org/10.2135/cropsci2017.08.0490

Vega LP, Bautista KT, Campos H, Daza S, Vargas G (2024) Biofuel production in Latin America: A review for Argentina, Brazil, Mexico, Chile, Costa Rica and Colombia. Energy Rep 11:28–38. https://doi.org/10.1016/j.egyr.2023.10.060

Wang P-w, Di Q, Liu X-Y (2020) Genome-Wide association Study Identifies Candidate Genes Related to Oleic acid content of Soybean Seed. https://doi.org/10.21203/rs.3.rs-17853/v1

Wardlaw GM, Snook JT (1990) Effect of diets high in butter, corn oil, or high-oleic acid sunflower oil on serum lipids and apolipoproteins in men. Am J Clin Nutr 51:815–821. https://doi.org/10.1093/ajcn/51.5.815

Wilcox J, Cavins J (1985) Inheritance of low linolenic acid content of the seed oil of a mutant in *Glycine max*. Theor Appl Genet 71:74–78. https://doi.org/10.1007/BF00278257

Wilson RF (2004) Seed composition. Soybeans: improvement, production. uses 16:621–677. https://doi.org/10.2134/agronmonogr16.3ed.c13

Zhang L, Yang X-d, Zhang Y-y, Yang J, Qi G-x, Guo D-q, Xing G-j, Yao Y, Xu W-j, Li H-y (2014) Changes in oleic acid content of transgenic soybeans by antisense RNA mediated posttranscriptional gene silencing. Int J Genomics 2014:921950. https://doi.org/10.1155/2014/921950

Zhang X, Hong M, Wan H, Luo L, Yu Z, Guo R (2019) Identification of key genes involved in embryo development and differential oil accumulation in two contrasting maize genotypes. Genes 10:993. https://doi.org/10.3390/genes10120993

Zhao X, Chang H, Feng L, Jing Y, Teng W, Qiu L, Zheng H, Han Y, Li W (2019a) Genome-wide association mapping and candidate gene analysis for saturated fatty acid content in soybean seed. Plant Breed 138:588–598. https://doi.org/10.1111/pbr.12706

Zhao X, Jiang H, Feng L, Qu Y, Teng W, Qiu L, Zheng H, Han Y, Li W (2019b) Genome-wide association and transcriptional studies reveal novel genes for unsaturated fatty acid synthesis in a panel of soybean accessions. BMC Genomics 20:1–16. https://doi.org/10.1186/s12864-019-5449-z