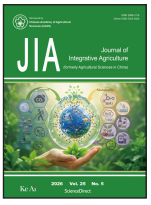
Journal of Integrative Agriculture 2026, 25(5): 1846–1856

Available online at ScienceDirect

Journal of Integrative Agriculture



RESEARCH ARTICLE

QTL mapping and allele stacking for enhanced lignan content in sesame (*Sesamum indicum* L.) using genotyping-by-sequencing

Yeon Ju An[1, 3], Min Young Kim[1], Sungup Kim[1], Jeongeun Lee[1], Sang Woo Kim[1], Jung In Kim[1], Eunyoung Oh[1], Heungsu Lee[1], Kwang-Soo Cho[2], Seung-Hyun Kim[3], Myoung Hee Lee[1], Eunsoo Lee[1#]

[1] *Upland Crop Breeding Research Division, Department of Southern Area Crop Science, National Institute of Crop Sciences, Rural Development Administration (RDA), Miryang 50424, Republic of Korea*

[2] *Central Crop Breeding Research Division, Department of Central Area Crop Science, National Institute of Crop Sciences, Rural Development Administration (RDA), Suwon 16429, Republic of Korea*

[3] *Department of Crop Science, College of Sanghuh Life Science, Konkuk University, Seoul 05029, Republic of Korea*

Highlights

- Key QTLs regulating lignan content (sesamin, sesamolin, and sesaminol) in sesame seeds were identified.
- The QTLs *qLIGNAN6* and *qLIGNAN11* showed allele stacking effects, confirming gene pyramiding potential for enhanced lignan content.

Abstract

Lignans are critical bioactive compounds in sesame (*Sesamum indicum* L.), valued for their health benefits and industrial potential. However, the genetic mechanisms underlying the lignan biosynthesis pathway remain poorly understood. This study aimed to identify key quantitative trait loci (QTLs) associated with lignan content using 140 F_9 recombinant inbred lines (RILs) derived from a cross between the sesame cultivars Areum and Gomazou, and to develop molecular markers for breeding programs to enhance lignan content. Genotyping-by-sequencing (GBS) technology was employed, combining single (*Ape*KI) and double-enzyme (*Pst*I/*Msp*I) digestions. This approach provided comprehensive genome coverage and substantially improved the quality of the linkage map. A total of 19 QTLs associated with lignan components, including oil-soluble and water-soluble lignans, were identified. Among these, loci on chromosomes 6 and 11 exhibited high LOD scores of 17.8 (PVE 13.9%) and 51.5 (PVE 68.7%), respectively, for total lignan contents in seeds. Two major QTLs, *qLIGNAN6* and *qLIGNAN11*, were identified as key contributors to lignan content. Significant single nucleotide polymorphisms (SNPs) linked to these QTLs were used to develop CAPS markers for potential application in marker-assisted selection (MAS). These markers were validated in RILs and germplasm to evaluate beneficial allele stacking and gene pyramiding. This study provides valuable genetic resources and molecular markers to facilitate the effective implementation of sesame breeding strategies to improve nutritional value. The findings contribute to the development of sesame varieties with higher lignan content, supporting both health and industrial applications.

Keywords: sesame, lignan, quantitative trait locus (QTL), genotyping-by-sequencing (GBS), CAPS marker

1. Introduction

Sesame (*Sesamum indicum* L.), a member of the Pedaliaceae family, is an oilseed crop grown mainly in tropical and subtropical regions (Ruperao *et al.* 2023). This oilseed is often regarded as the “queen of oilseeds” owing to its nutritional and health benefits for humans and animals (Yadav *et al.* 2022; Dossou *et al.* 2023). The seeds are rich in bioactive compounds, including essential minerals, vitamins, phytosterols, polyunsaturated fatty acids, tocopherols, and a specific group of lignans such as sesamin and sesamolin (Pathak *et al.* 2014). In particular, the oil-soluble lignans, sesamin, and sesamolin, lower blood glucose and cholesterol levels, increase gamma-tocopherol levels, and inhibit rancidity in sesame oil, thereby improving its storage stability

Received 20 September 2024; Received in revised form 11 December 2024; Accepted 3 January 2025; Available online 18 February 2025
[#]Correspondence Eunsoo Lee, E-mail: awesomelee@korea.kr

© 2026 CAAS. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/). Peer review under responsibility of Editorial Board of *Journal of Integrative Agriculture*.
doi: 10.1016/j.jia.2025.02.025

(Yadav *et al.* 2022). In addition, sesame seeds also contain large amounts of water-soluble lignan glucosides, such as pinoresinol and sesaminol glucosides (Moazzami *et al.* 2006). Among them, sesaminol and its glucosides exhibit antioxidant, anti-inflammatory, and estrogen-like activities (Kim *et al.* 2008; Wu *et al.* 2019; Andargie *et al.* 2021). Despite extensive research on the functional roles of lignans, studies investigating the genetic mechanisms underlying their biosynthesis and regulation remain limited.

Next-generation sequencing (NGS) technologies have revolutionized genome-wide analyses by enabling the development of ultra-high-density genetic linkage maps across various crops, thereby allowing the identification of high-resolution candidate loci within a few kilobases (Jaganathan *et al.* 2020). Recently, single nucleotide polymorphisms (SNPs) derived from NGS have been increasingly used in genetic diversity studies (Cui *et al.* 2017). SNPs are advantageous over conventional markers owing to their abundance and the ease with which numerous unbiased SNPs can now be obtained for non-model species (Fischer *et al.* 2017). Among the various NGS methods for SNP marker development, genotyping-by-sequencing (GBS) is a prominent technique owing to its use of restriction enzymes (REs) to reduce genome complexity. GBS also uses a DNA barcode system to process numerous samples simultaneously, generating a representation library that identifies numerous variants while reducing costs (Davey *et al.* 2011; Elshire *et al.* 2011). Genome reduction in GBS is achieved using REs, and selecting appropriate REs for species-specific GBS optimization is crucial for library development (Hamblin and Rabbi 2014). Commonly used REs in plants include *Ape*KI, *Pst*I, and *Eco*RI, often combined with other cutters such as *Msp*I, *Mse*I, and *Hpa*II based on genome specificity (Wang *et al.* 2017; Nguyen *et al.* 2018; Ruperao *et al.* 2023). However, single-enzyme digestion methods can pose challenges, such as a high proportion of short fragments and inconsistent read counts across individuals. Consequently, double-enzyme digestion is used to achieve consistent results and improve the overall data quality (Wang *et al.* 2017). This study evaluated GBS library construction using both a single (*Ape*KI) and a combination of two REs (*Pst*I/*Msp*I) in sesame, assessing their influence on genome coverage and efficiency to determine the optimal method for obtaining comprehensive genetic information.

Genetic research on sesame has made notable progress; however, significant gaps remain in understanding the genetic basis of functional components such as sesamin, sesamolin, and sesaminol, which are crucial for enhancing the nutritional value of sesame. Previous studies have investigated genetic research related to oil-soluble lignans. Two candidate genes (*SIN_1005755* and *SIN_1005756*) that regulate the biosynthesis of sesamin and sesamolin have been identified at the same locus on chromosome 11 (Xu *et al.* 2021). A GWAS study identified 17 and 72 SNP loci associated with sesamin and sesaminol, respectively, with *SiNST1* on chromosome 11 as the major candidate gene (Dossou *et al.* 2023). Additionally, a QTL-seq study identified three loci, *qLignan1-1*, *qLignan6-1*, and *qLignan11-1*, regulating sesamin and sesamolin content (Kim *et al.* 2023). Although studies on sesamin and sesamolin have been conducted, further studies are still needed to understand the water-soluble lignans that contribute to the overall lignan content in sesame seeds. This study addresses gaps in sesame's genetics using the F_9 generation of an RIL derived from a cross between sesame varieties Areum and Gomazou. The contents of key lignans — sesamin, sesamolin, sesaminol, sesaminol-diglucoside (sesaminol_DG), and sesaminol-triglucoside (sesaminol_TG) — were analyzed in seeds from both parents and RIL populations. This study aimed to provide valuable genetic resources and insights to guide breeding strategies that enhance sesame's nutritional value and advance molecular breeding efforts.

2. Materials and methods

2.1. Plant materials and growth conditions

This study used a RIL population developed from a cross between the sesame varieties Areum and Gomazou. The maternal parent, Areum, is a black seed variety bred in Korea, whereas the paternal parent, Gomazou, is a high-lignan variety bred in Japan (Yasumoto and Katsuta 2006). A total of 140 lines were advanced to the F_9 generation *via* single-seed descent in 2016, using field and greenhouse conditions. RILs and their parental lines were grown at the Upland Crop Breeding Research Division, Department of Southern Area Crop Science, National Institute of Crop Sciences (NICS), Rural Development Administration (RDA), Miryang, Republic of Korea, with a spacing of 55 cm×15 cm, from June to August 2020. After harvesting at least five individuals per line, seeds were pooled following seed quality sorting and used for the analysis of sesamin, sesamolin, sesaminol, and its glucosides.

2.2. Lignan content phenotyping and analysis

Harvested seeds of Areum, Gomazou, and RIL populations were dried and ground for analysis. The contents of lignans — sesamin, sesamolin, sesaminol, sesaminol_DG, and sesaminol_TG — were analyzed using high-performance liquid chromatography (HPLC; Dionex 3000, Thermo Fisher Scientific Co., Waltham, MA, USA). Sesamin and sesamolin were extracted by placing 1.0 g of ground sesame sample in a rotary shaker (Multi Shaker, MMS-110, Eyela Co., Tokyo, Japan) at room temperature for 24 h, and 20 mL of methanol was subsequently added. Thereafter, the supernatant was filtered using a syringe filter (Whatman 0.2 μm, NYL; Whatman, Maidstone, UK) and analyzed. HPLC was performed under isocratic conditions, using a mobile phase composed of 60% (v/v) methanol and 40% (v/v) distilled water. The column used was a YMC Triart C-18 (1.9 μm, 2.0 mm×50 mm, 30°C; YMC Co., Kyoto, Japan), with a flow rate of 0.3 mL min^{-1}. Detection was conducted using a diode-array detector (DAD 3000, Thermo Fisher Scientific Co., Waltham, MA, USA) set to UV 290 nm for 15 min. In addition, sesamin and sesamolin contents were measured in the Gomazou-derived RILs from the previous study (Kim *et al.* 2023) and the 100 germplasm harvested in 2024 using the same methods.

Sesaminol, sesaminol_DG, and sesaminol_TG were analyzed following defatting. A total of 1 g of ground sesame sample was first combined with 10 mL of hexane, followed by the addition of 20 mL of 70% ethanol. Extraction was performed using a rotary shaker (Multi Shaker, MMS-110, Eyela Co., Tokyo, Japan) at room temperature for 24 h. Thereafter, the supernatant was filtered using a syringe filter (Whatman 0.2 μm, NYL; Whatman, Maidstone, UK) and analyzed. The column used was a Supersil ODS2 (5 μm, 4.6 mm×25 mm, 30°C; Dalian Analytical Instruments Co., Dalian, China) with a flow rate of 1.0 mL min^{-1}. HPLC analysis was performed using a mobile phase gradient of distilled water (A) and methanol (B). Gradient was as follows: 0 min, 90%A/10%B; 2 min, 30%A/70%B; 25 min, 20%A/80%B; 30 min, 90%A/10%B, for re-equilibration. Standards were obtained from Sigma (Sigma-Aldrich Co., Burlington, MA, USA). Raw data for the lignan components, including technical replicates, are deposited in the RDA-NaRDA data repositories (https://doi.org/10.23138/epops.20241223-00017.r1) for public access. Correlation among the eight traits was evaluated by Pearson's correlation coefficient (Pearson 1896), and analyzed with the 'corrplot' library in R studio version 4.2.0.

2.3. DNA extraction, whole-genome resequencing, and genotyping-by-sequencing (GBS)

Genomic DNA was extracted from the young leaves of Areum, Gomazou, and their F_9 RILs using the GenoAll® Exgene™ Plant SV Kit (GenoAll Biotechnology, Seoul, Korea). DNA quality was assessed using a Nanodrop 3000 spectrometer (Thermo Fisher Scientific Co., Waltham, MA, USA) at 260/280 nm, and the concentration was adjusted to 20 ng μL^{-1} for sequencing. Parental lines (Goenbaek and Gomazou) were sequenced to a depth of approximately 20× using paired-end sequencing with the TruSeq Nano DNA library on the Illumina Novaseq 6000 platform. The resulting raw reads were deposited at the National Agricultural Biotechnology Information Center (NABIC) under the biological project number NV-0849. Genotype information for the RILs was obtained using the GBS method. Two different enzyme setups were used: single-enzyme digestion using *Ape*KI and double-enzyme digestion using *Pst*I/*Msp*I, following previously established barcodes and library construction protocols (Elshire *et al.* 2011; Poland *et al.* 2012; Lee *et al.* 2021). Barcoded adapters were ligated with digested DNA fragments (Appendix A) and the ligated DNA was amplified. Amplification conditions were 95°C for 1 min, followed by 18 cycles of 95°C for 30 s, 62°C for 20 s, and 68°C for 15 s. Separate libraries were constructed for each restriction enzyme setup, and paired-end sequencing was performed across four lanes using the Illumina HiSeq X Ten platform.

2.4. Genetic mapping and quantitative trait loci (QTLs) analysis

The re-sequenced reads for the parental lines were trimmed using Trimmomatic v0.39 to remove residual adapters and low-quality sequences. Trimmed reads were aligned to the sesame reference genome (cv. Zhongzhi 13 version 2) using BWA-MEM v0.7.17 and SAMtools v1.15.1 for variant calling (Wang *et al.* 2016). For GBS-derived RILs, barcode sequences were removed using the Fastx toolkit, and the alignment pipeline followed the same procedure as for parental lines. SNPs with polymorphisms between parents were selected for genetic mapping and further filtered to include those with a mapping quality≥30, mapping depth≥5, and a genotypic missing rate≤25% in RILs, using VCFtools v0.1.16. Missing genotypes in RILs were refined via imputation and binning using Genotype-Corrector (Miao *et al.* 2018), applying representative markers for genetic map construction, and forming 13 linkage groups based on the regression mapping method. Marker order was calculated as recombination distance (cM) using the Kosambi function of the JoinMap v4.1 program. Quantitative trait loci (QTLs) analysis was conducted using IciMapping v4.1 with the ICIM-ADD mapping method and a step size of 0.5 cM. Logarithm of odds (LOD) thresholds were established through a 1,000-permutation test at a significance level of 0.05.

2.5. Variant analysis, marker development, and phenotypic impact assessment of allele combinations

Variant analysis identified candidate genes associated with sesame lignan content regulation using resequencing data from the mapping parents. To further confirm the variation specific to the high-lignan variety, we used the low-lignan sesame 'Goenbaek' sequence from NABIC (NN-8889). The significance of selected variants was assessed and categorized as HIGH, MODERATE, LOW, and MODIFIER based on their gene positions using snpEff software (Cingolani *et al.* 2012). Putative functional information for sesame genes was obtained through BLASTp similarity analysis with the *Arabidopsis* gene protein sequences (TAIR version 10). Orthologous genes with the highest similarity were identified, and their functions were inferred. Based on the polymorphic SNPs identified in two QTLs, *qLIGNAN6* (chr6:15158030) and *qLIGNAN11* (chr11:803989), CAPS markers were developed to assess the phenotypic impact of alleles on sesame lignan regulation (Appendix B). REs were selected using dCAPS Finder 2.0 (Neff *et al.* 2002), and primers for PCR were designed using Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/). Statistical analysis was performed using analysis of variance (ANOVA) with R software, and statistically significant differences are marked with asterisks as follows: *** (P<0.001), ** (P<0.01), and * (P<0.05).

3. Results

3.1. Lignan content variation and correlation analysis

Sesamin, sesamolin, sesaminol, sesaminol_DG, and sesaminol_TG contents in Areum, Gomazou, and their RILs were investigated (Table 1). Oil-soluble lignans, sesamin and sesamolin, were detected at concentrations of 1.18 and 0.83 mg g^{-1} in Areum, and 3.92 and 2.43 mg g^{-1} in Gomazou, respectively. Soluble lignan glycosides, namely sesaminol, sesaminol_DG, and sesaminol_TG, were found at concentrations of 0.12, 0.06, and 0.63 mg g^{-1} in Areum and 0.01, 0.66, and 8.05 mg g^{-1} in Gomazou, respectively.

Table 1 Summary statistics of functional compounds in Areum (AR) and Gomazou (GM), and their derived values from a recombinant inbred line (RIL) population

Traits (mg g^{-1})[1)]	Mapping parents		RIL population					
	AR	GM	Average	SD	Min	Max	Skewness	Kurtosis
Sesamin	1.18±0.01	3.92±0.02	4.15	2.73	0.57	10.80	–0.92	0.55
Sesamolin	0.83±0.01	2.43±0.01	1.80	0.72	0.60	3.64	–1.05	0.13
TOLC	2.00±0.01	6.35±0.03	5.95	3.37	1.36	13.63	–1.05	0.43
Sesaminol	0.12±0.09	0.01±0.00	0.10	0.07	0	0.31	0.30	0.98
Sesaminol_DG	0.06±0.02	0.66±0.00	0.37	0.22	0.05	1.38	2.15	1.11
Sesaminol_TG	0.63±0.03	8.05±0.06	4.04	2.40	0.21	11.89	0.26	0.76
TSLC	0.80±0.14	8.72±0.07	4.51	2.59	0.54	13.31	0.29	0.76
TLC	2.80±0.13	15.08±0.10	10.46	5.46	2.50	22.03	–1.24	0.27

[1)] TOLC, total oil-soluble lignan contents (sesamin and sesamolin); sesaminol_DG, sesaminol-diglucoside; sesaminol_TG, sesaminol-triglucoside; TSLC, total soluble lignan contents (sesaminol, sesaminol-diglucoside, and sesaminol-triglucoside); TLC, total lignan content. Values are mean±SD (*n*=3).

Lignan content in the RIL population displayed a wide range of values, indicating considerable variability across lines (Appendix C-a). In the RIL population, sesamin content ranged from 0.57 to 10.8 mg g^{-1}, with a mean of 4.15 mg g^{-1}. Sesamolin content ranged from 0.6 to 3.64 mg g^{-1}, with a mean of 1.8 mg g^{-1}. Sesaminol, sesaminol_DG, and sesaminol_TG contents ranged from 0.00 to 0.31 mg g^{-1}, 0.05 to 1.38 mg g^{-1}, and 0.21 to 11.89 mg g^{-1}, with respective means of 0.1, 0.37, and 4.04 mg g^{-1}. Transgressive segregation was observed for both Areum and Gomazou in total oil-soluble lignans (TOLC), total soluble lignans (TSLC), and total combined lignan content (TLC). Among the RILs, 61 lines showed higher TOLC than Gomazou, with a mean of 9.31 mg g^{-1}, while 11 lines exhibited lower TOLC than Areum, with a mean of 1.67 mg g^{-1}. For TSLC, eight lines had higher values than Gomazou, with a mean of 10.7 mg g^{-1}, while only one line had lower TSLC than Areum, with a value of 0.54 mg g^{-1}. Regarding TLC, 38 lines had higher levels than Gomazou, with a mean of 17.66 mg g^{-1}, while two lines had lower levels than Areum, with a mean of 2.61 mg g^{-1}. Although the data showed a skewed distribution, as indicated by skewness and kurtosis absolute values ranging from 0.13 to 2.15 (Table 1), the QTL analysis was performed without normalization. Most of the investigated lignan contents showed a wide range of values exceeding those of the parental lines, supporting the assumption that these traits are quantitative and likely regulated by multiple genes.

A correlation analysis examined the relationships between the contents of each lignan component (Appendix C-b). A positive correlation was evident among the lignan components, whereas a negative correlation was observed specifically between sesaminol content and other lignan components. Notably, TSLC showed a significant positive correlation with sesaminol-TG content (r=1.00, P≤0.001). This suggested that sesaminol-TG constitutes approximately 92% of the TSLC, highlighting its dominant role among the soluble lignans.

3.2. Sequencing and GBS analysis of parental lines and recombinant inbred lines

Parental lines Areum and Gomazou were re-sequenced, generating 32.3 and 32 Gb of data, respectively, covering 90× and 89× of the sesame reference genome (357 Mb) (Wang *et al.* 2016). After raw data quality trimming, 165,996,600 high-quality reads from Areum and 165,527,464 high-quality reads from Gomazou were used for mapping, yielding 96.1% and 96.2% coverage across 13 chromosomes, respectively. This extensive coverage provided a sufficient basis for comparing genotypes between populations. Genotyping for the RIL population derived from Areum and Gomazou was performed using GBS with two different strategies involving distinct REs. Initially, genotyping efficiency was tested using two different restriction enzyme approaches on a set of 96 pooled individuals. The single-enzyme method employing *Ape*KI, which recognizes and cleaves the 'GCWGC' sequence, generated 623,630,782 total reads, with an average of 451,627,998 aligned reads per individual and a coverage of 4.6% (15.7% chromosome-specific). In contrast, the double-enzyme method, using *Pst*I (which recognizes the 'CTGCAG' sequence) and *Msp*I (which recognizes the 'CCGG' sequence), generated 704,092,980 reads, providing an average of 530,771,158 aligned reads per individual, with a coverage of 0.7% (2.9% chromosome-specific). Combining results from both GBS approaches yielded an average of 982,404,595 reads per individual, with an overall coverage efficiency of 4.9% (17.2% chromosome-specific) (Appendix D). To compare the distribution of variants across the genome using the *Ape*KI and *Pst*I/*Msp*I, SNP variants with a mapping quality≥30 were extracted and plotted based on their physical position with a 0.1 Mb window size. The results indicated that these methods complement each other in their variant distribution patterns (Fig. 1).

3.3. Genetic mapping and QTLs identification for lignan traits

In total, 2,034,169 SNPs were detected from resequencing and population-GBS reads, which were filtered for polymorphisms between parental lines. Thereafter, selected SNPs were imputed and binned, resulting in 1,515 bin markers. Genetic map for the Areum×Gomazou RIL population, constructed using Kosambi calculations, included 1,334 markers (Appendix E). This map exhibited a range of 57 to 201 markers per chromosome, with an average of 102.6 markers. Total map length was 1,441.9 cM, with an average chromosome length of 110.9 cM and an average marker spacing of 1.1 cM (Appendix F).

QTL analysis was performed using the phenotypic data and the constructed genetic map. This analysis identified 19 QTLs associated with five lignan-related traits distributed across five chromosomes (Chr4, Chr5, Chr6, Chr9, and Chr11) (Table 2; Fig. 2). Specifically, QTLs for sesamin and sesamolin were located on chromosomes 6 and 11, and associated with the total oil-soluble lignans (TOLC). Additionally, QTLs for sesaminol, sesaminol_DG, and sesaminol_TG were identified on chromosomes 4, 5, and 9, respectively, and were associated with the TSLC. TLC, including both soluble and oil-soluble lignans, was mapped to chromosomes 6 and 11. LOD scores for TOLC were 17.64 for *qTOLC6-1* locus and 51.29 for *qTOLC11-1*, with phenotypic variations of 13.95 and 69.16%, respectively. Higher phenotypic variation associated with *qTOLC11-1* indicates its significant effect on TOLC. For the TSLC, the *qTSLC6-1* had a LOD score of 61.62 and explained 47.88% of the phenotypic variation, whereas *qTSLC11-1* had a LOD score of 22.66 with 8.50% variation explained. This suggested that *qTSLC6-1* has a greater impact on TSLC despite the higher LOD score of *qTSLC11-1*. For the TLC, *qTLC6-1* had a LOD score of 17.77 and contributed 13.88% to the phenotypic variation, whereas the *qTLC11-1* had a LOD score of 51.50 and explained 68.74% of the variation, indicating its greater impact on TLC.

The analysis revealed that loci were co-located on chromosome 11 for all traits and on chromosome 6 for all traits except sesaminol. Although *qTOLC6-1* and *qTSLC6-1* are located on chromosome 6 and are only 0.5 cM apart, both appear to be beneficial alleles inherited from Gomazou,

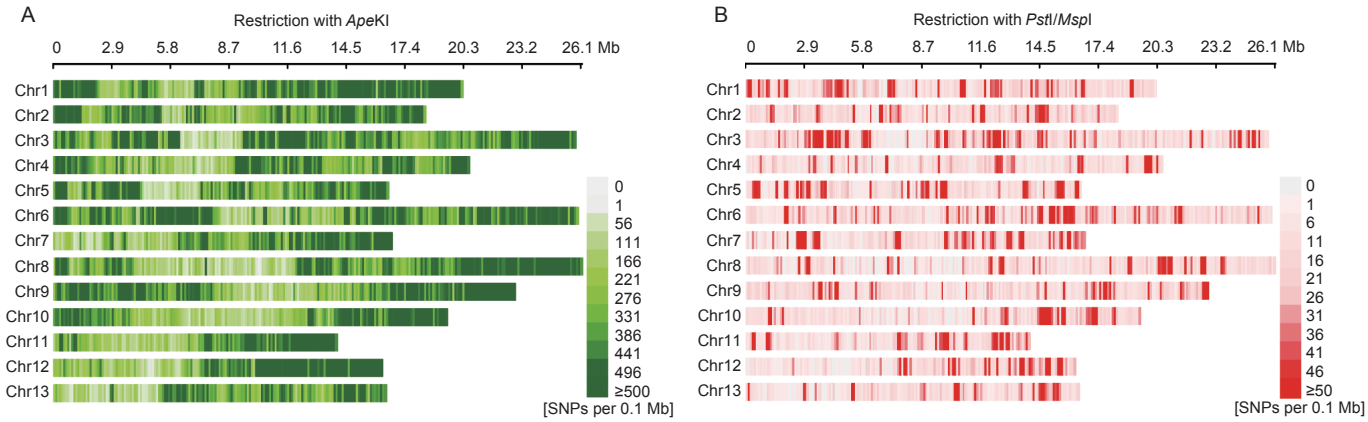


Fig. 1 Genotyping efficiency using restriction enzymes *Ape*KI and *Pst*I/*Msp*I on 96 individuals pooled for sequencing. *Ape*KI (A) and *Pst*I/*Msp*I (B) results of the genotyping-by-sequencing (GBS) approaches. SNP variants with a mapping quality≥30 were extracted and displayed in 0.1 Mb windows based on physical position.

Table 2 Quantitative trait loci result of genetic mapping for the lignan contents in the recombinant inbred line (RIL) population derived from Areum×Gomazou[1)]

Trait[2)]	Locus	Chr	Pos	Marker interval	LOD	PVE (%)	Add	No. of genes
Sesamin	*qSesamin6-1*	6	71.5	chr6_15065347–chr6_15174720	12.45	10.15	–0.84	11
	qSesamin11-1	11	1.0	chr11_134577–chr11_996151	46.98	70.80	–2.24	96
Sesamolin	*qSesamolin6-1*	6	71.5	chr6_15065347–chr6_15174720	25.10	31.67	–0.39	11
	qSesamolin11-1	11	1.5	chr11_134577–chr11_996151	35.59	45.59	–0.47	96
TOLC	*qTOLC6-1*	6	71.5	chr6_15065347–chr6_15174720	17.64	13.95	–1.22	11
	qTOLC11-1	11	1.0	chr11_134577–chr11_996151	51.29	69.16	–2.73	96
Sesaminol	*qSesaminol4-1*	4	69.0	chr4_10053158–chr4_10181297	3.74	7.65	0.02	14
	qSesaminol11-1	11	1.0	chr11_134577–chr11_996151	14.12	35.93	0.04	96
Sesaminol_DG	*qSDG6-1*	6	71.0	chr6_13694483–chr6_15065347	82.28	48.12	–0.49	120
	qSDG11-1	11	0.5	chr11_134577–chr11_996151	19.84	3.29	–0.13	96
Sesaminol_TG	*qSTG5-1*	5	50.5	chr5_6220044–chr5_6300555	3.18	0.93	0.47	2
	qSTG6-1	6	70.5	chr6_13694483–chr6_15065347	59.30	50.34	–3.45	120
	qSTG9-1	9	57.0	chr9_15100790–chr9_15552480	3.18	0.96	–0.48	46
	qSTG11-1	11	1.0	chr11_134577–chr11_996151	23.54	10.27	–1.56	96
TSLC	*qTSLC5-1*	5	50.5	chr5_6220044–chr5_6300555	3.17	0.81	0.51	2
	qTSLC6-1	6	71.0	chr6_13694483–chr6_15065347	61.62	47.88	–3.90	120
	qTSLC11-1	11	1.0	chr11_134577–chr11_996151	22.66	8.50	–1.65	96
TLC	*qTLC6-1*	6	71.0	chr6_13694483–chr6_15065347	17.77	13.88	–1.95	120
	qTLC11-1	11	1.5	chr11_134577–chr11_996151	51.50	68.74	–4.35	96

[1)] PVE, percentage of phenotype variance explained by quantitative trait loci; Add, allelic additive effect.
[2)] TOLC, total oil-soluble lignan contents (sesamin and sesamolin); sesaminol_DG, sesaminol-diglucoside; sesaminol_TG, sesaminol-triglucoside; TSLC, total soluble lignan contents (sesaminol, sesaminol-diglucoside, and sesaminol-triglucoside); TLC, total lignan content.

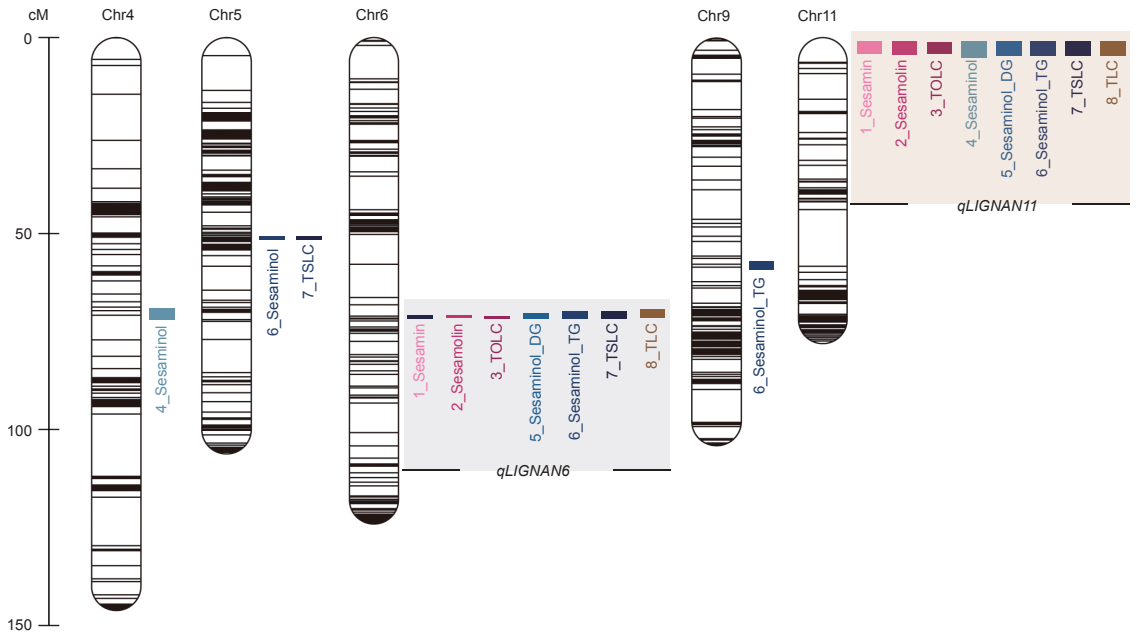


Fig. 2 QTLs mapped to linkage groups for lignan-related traits. Each trait was represented by a different color; chromosome without QTLs were excluded. Black bars on chromosomes indicate bin markers. Trait names are indicated on the right of each chromosome. Sesaminol_DG, sesaminol-diglucoside; sesaminol_TG, sesaminol-triglucoside; TOLC, total oil-soluble lignan contents; TSLC, total soluble lignan contents; TLC, total lignan contents.

sharing chr6_15065347 as a flanking marker. Despite their close physical proximity, the limited marker resolution in the QTL analysis resulted in the identification of separate loci. However, since these loci are likely to originate from the same inherited gene or genetic variant, we merged them to expand the candidate genetic regions to include potential common causative genes influencing lignan content in sesame seeds. Finally, two major QTL regions were defined: *qLIGNAN6* (chr6:13694483–chr6:15065347) and *qLIGNAN11* (chr11:134577–chr11:996151), identified as significant loci (Fig. 2). SNP flanking markers (chr6:13694483–chr6:15065347 and chr6:15065347–chr6:15174720) identified genetic regions encompassing 11 genes within the 0.1 Mb region and 120 genes within the 1.3 Mb region, respectively (Table 2). By merging these loci, the *qLIGNAN6* region was found to overlap a common gene at the chr6_15065347, resulting in a final total of 130 genes within the 1.4 Mb region. Similarly, 96 genes were identified within the 0.8 Mb region of the *qLIGNAN11* locus (Table 2).

3.4. Functional classification of variants within major QTLs

To identify candidate genes within the two major QTL regions controlling lignan content, variants specific to the high-lignan variety Gomazou were explored by comparing them with the low-lignan variety Areum used in this study and the additional low-lignan variety Goenbaek from a previous study (Kim *et al.* 2023). We compared SNPs/InDels in 130 genes within *qLIGNAN6* and 96 genes within *qLIGNAN11*, including 1 kb upstream and downstream of each gene. As a result, 201 variants unique to the high-lignan variety were identified, located within 37 gene regions (Appendix G). The functional potential of these candidate genes was assessed for similarity by inferring their protein sequences with homologous genes in the model plant, *Arabidopsis*. Since coding regions directly determine protein structure and function, and upstream regulatory regions control gene expression, variants located in these regions were prioritized for those with higher functional impact. Ultimately, 82 SNPs/InDels within 17 genes were selected as candidate variants and genes (Appendix H).

3.5. Impact of allele stacking on lignan content in recombinant inbred lines

To assess the impact of allele combinations from two major QTLs, *qLIGNAN6* and *qLIGNAN11*, we analyzed the RIL genotypes at these loci. From 201 high-lignan-specific variants previously selected, SNPs located at 15,158,030 bp on chromosome 6 and 803,989 bp on chromosome 11 that can be converted to CAPS markers were selected and assayed for segregation in the population (Fig. 3). *qLIGNAN6* and *qLIGNAN11* are beneficial alleles from Gomazou, whereas *qlignan6* and *qlignan11* are deleterious alleles from Areum. RILs were grouped according to their genotype combinations. RILs with the genotype *qLIGNAN6*+*qLIGNAN11* had the highest average TOLC at 9.92 mg g^{-1}, which was significantly higher than the 5.91 mg g^{-1} observed in RILs with *qLIGNAN6*+*qlignan11* or *qlignan6*+*qLIGNAN11*, and notably higher than the 2.35 mg g^{-1} found in those with *qlignan6*+*qlignan11* (Fig. 4). Similarly, the average TSLC was highest in RILs with the genotype *qLIGNAN6*+*qLIGNAN11*, measuring 6.93 mg g^{-1}, compared to 4.75 mg g^{-1} in RILs with *qLIGNAN6*+*qlignan11* or *qlignan6*+*qLIGNAN11*, and 2.15 mg g^{-1} in those with

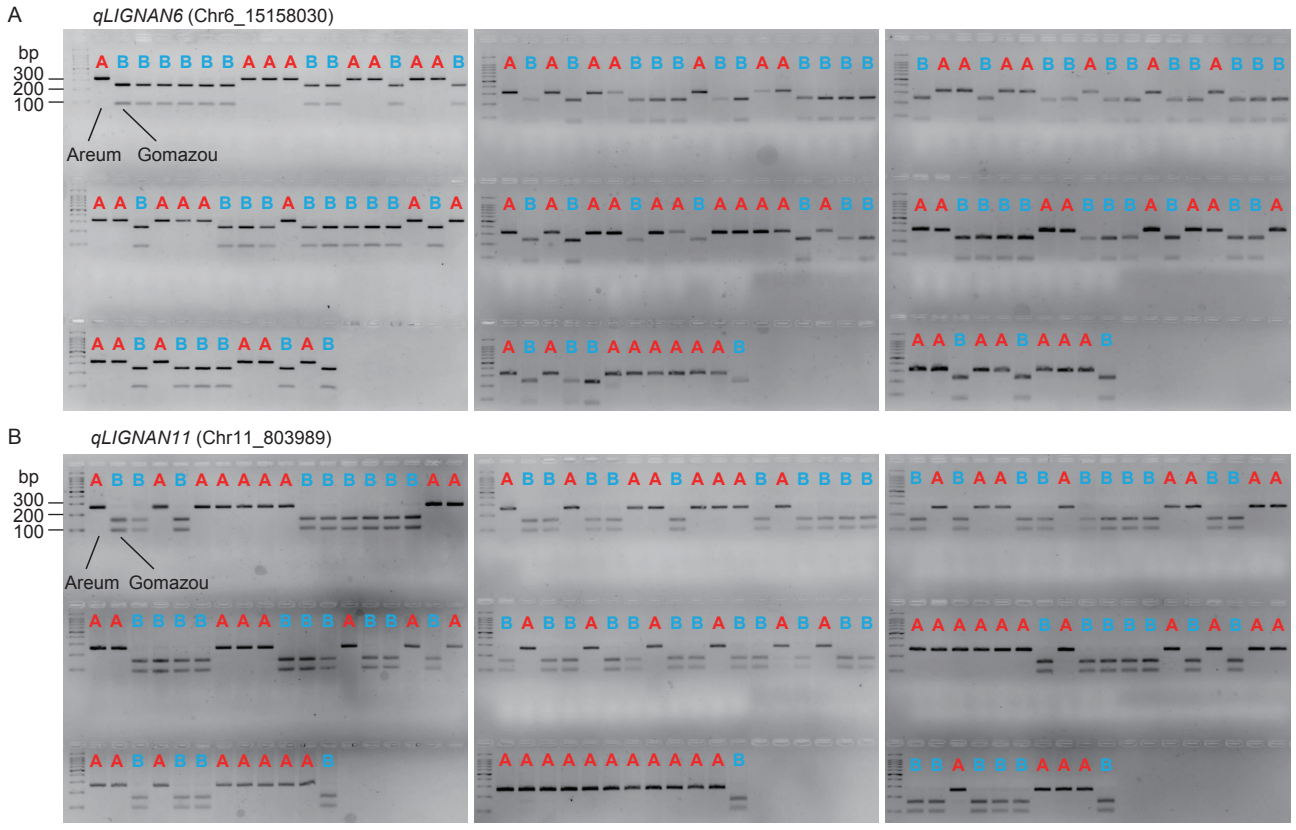


Fig. 3 Gel electrophoresis results of SNP-based CAPS marker from two major QTLs, *qLIGNAN6* and *qLIGNAN11*, in the recombinant inbred line (RIL) population. PCR products digested with restriction enzymes *Mse*I and *Hpy*188I are shown for parental lines (lanes 1 & 2) and an RIL mapping population. A, differential band for a polymorphic SNP at physical position 15,158,030 bp on chromosome 6 resulting from *Mse*I restriction enzyme digestion. Target band size for PCR amplification is 333 bp (A allele from Areum), and the cut region caused by restriction enzyme action is 83/250 bp (B allele from Gomazou). B, differential band for a polymorphic SNP at physical position 803,989 bp on chromosome 11 resulting from *Hpy*188I restriction enzyme digestion. Target band size for PCR amplification is 264 bp (A allele from Areum), and the cut region caused by restriction enzyme action is 163/101 bp (B allele from Gomazou).

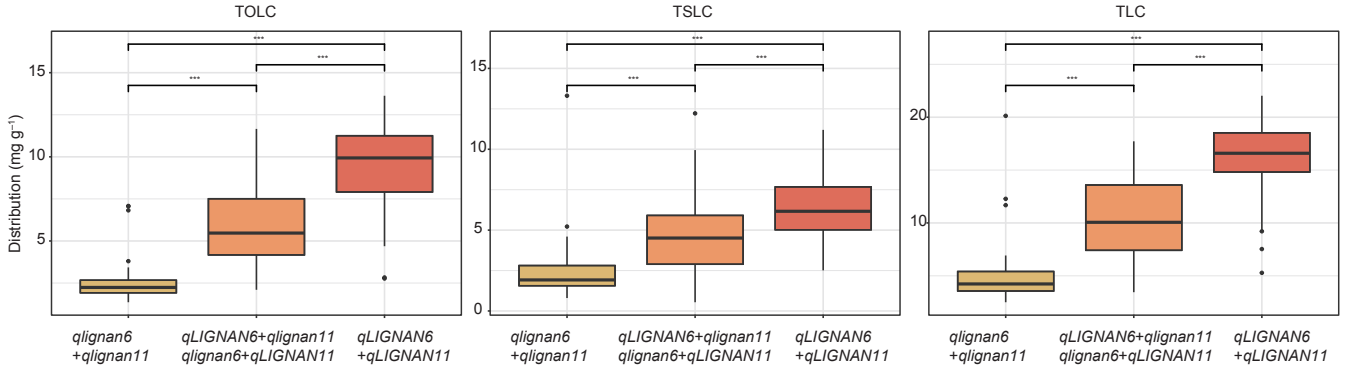


Fig. 4 Additive allele effects of *qLIGNAN6* and *qLIGNAN11* on total oil-soluble lignan contents (TOLC), total soluble lignan contents (TSLC), and total lignan contents (TLC) in recombinant inbred lines (RILs). Beneficial alleles, *qLIGNAN6* and *qLIGNAN11* originate from the paternal genotype Gomazou, whereas the deleterious alleles, *qlignan6* and *qlignan11*, originate from the maternal genotype Areum. Boxplots show the phenotypic distribution of RILs containing specific combinations of parental alleles. Statistical significance between allele combinations is indicated by ***, with $P<0.001$, as determined by a *t*-test.

qlignan6+qlignan11. Furthermore, RILs with the genotype *qLIGNAN6+qLIGNAN11* had the greatest average TLC of 16.86 mg g^{-1}, in contrast to 10.66 mg g in lines with *qLIGNAN6+qlignan11* or *qlignan6+qLIGNAN11*, and 4.5 mg g^{-1} in those with *qlignan6+qlignan11*.

For validation of these beneficial alleles from two major loci, an additional Gomazou-derived RIL population (Kim *et al.* 2023) and 100 germplasm were used to evaluate the practicality of the markers with TOLC (Appendix I-a–d). Statistical analysis was performed to investigate the

relationship between TOLC content and allele segregation in 255 RILs derived from Goenbaek×Gomazou (Appendix I-a and b)). Significant differences were found among the genotypes, with *qLIGNAN6*+*qLIGNAN11* exhibiting the highest TOLC content of 8.71 mg g^{-1}, followed by *qLIGNAN6*+*qlignan11* or *qlignan6*+*qLIGNAN11* with 7.58 mg g^{-1}, and *qlignan6*+*qlignan11* showing the lowest content at 5.61 mg g^{-1} (Appendix I-e). Similarly, the practical applicability of the CAPS markers was validated using 100 germplasm (Appendix I-c and d), revealing that TOLC content was highest with genotype *qLIGNAN6*+*qLIGNAN11* at 6.46 mg g^{-1}, followed by *qLIGNAN6*+*qlignan11* or *qlignan6*+*qLIGNAN11* at 4.76 mg g^{-1}, and *qlignan6*+*qlignan11* at 4.18 mg g^{-1}, demonstrating a significant increase with favorable allele stacking (Appendix I-e). Differences in TOLC among genotypes were statistically significant, with varying levels of significance: *** ($P \leq 0.001$), ** ($P \leq 0.01$), and * ($P \leq 0.05$). In conclusion, analysis of potential interactions among the QTLs revealed that the combination of *qLIGNAN6* and *qLIGNAN11* from the Gomazou allele substantially enhanced TOLC, TSLC, and TLC content in sesame seeds.

4. Discussion

4.1. Significance of sesame lignans in health and breeding research

Sesame contains oil-soluble lignan compounds such as sesamin and sesamolin, known for their antioxidative, anticancer, anti-inflammatory, anti-proliferative, antihypertensive, and anti-melanogenesis effects (Senouwa *et al.* 2021; Kim *et al.* 2022). Additionally, soluble lignans, of which sesaminol_TGis the most abundant, are suggested to be the active ingredients in defatted seeds (Moazzami *et al.* 2006; Ben Othman *et al.* 2015). During oil extraction, oil-soluble lignans remain in the oil, contributing to the oxidative stability of sesame oil, whereas soluble lignans are retained in the seed meal (Li *et al.* 2023). These oil-soluble and soluble lignans are stored in different forms within sesame seeds and are absorbed differently when consumed. This distinction highlights the need for research on these specific lignan types to comprehensively understand their health benefits and metabolic pathways. Moreover, enhancing the overall lignan content in sesame seeds is crucial for increasing the economic value of sesame, as higher lignan levels improve both the nutritional quality and marketability of sesame-based products. Despite the benefits of lignans and increasing global sesame consumption, genetic studies on their molecular traits for breeding are limited.

The genetic mechanisms involved in lignan synthesis in sesame have been studied usingvarious approaches. Functional characterization of the cytochrome P450 enzymes CYP81Q1 and CYP92B14, both critical for lignan biosynthesis, was performed to investigate the lignan synthesis in sesame (Satake *et al.* 2013). CYP81Q1 initiates the conversion of the precursor lignan pinoresinol to (+)-piperitol and then to (+)-sesamin, forming two methylenedioxy bridges that underscore its pivotal role in the lignan biosynthetic pathway (Ono *et al.* 2006). Genetic mapping identified *CYP92B14* on LG2, which complements lignan biosynthesis by converting (+)-sesamin to (+)-sesamolin and, unexpectedly, to (+)-sesaminol through an oxidative rearrangement of α-oxy-substituted aryl groups (ORA), implicating functional coordination with CYP81Q1 (Murata *et al.* 2017). However, no CYP genes were identified among the candidate genes in this study, suggesting that other genetic factors may influence lignan synthesis. *SiNST1* (*SIN_1005755*), which encodes a NAC domain-containing protein, was previously reported to play a potential role in the regulation of sesamin and sesamolin by association analysis and gene expression comparison (Xu *et al.* 2021; Dossou *et al.* 2023). In this study, the *qLIGNAN11* region, which contains 96 genes, was also found to contain *SIN_1005755*. However, the gene coding regions of the parental lines used in this study did not show any nucleotide sequence variants. The reason why this gene is located at the physical position 129,873–130,205 bp on chromosome 11, a region with relatively low recombination frequency. In addition to gene studies, QTL-seq research has provided valuable insights into lignan content and identified a locus on chromosome 11 that contributes to lignan levels in sesame (Kim *et al.* 2023). QTL mapping in this study identified major loci influencing TLC, with the most prominent located on chromosomes 6 and 11. These findings align with previously reported results, suggesting that the observed effects may be attributed to a beneficial allele from the Gomazou variety, which was specifically bred to enhance lignan content. Additionally, association analysis using RILs offered a detailed evaluation of the genetic contributions to lignan-related traits, including sesamin, sesamolin, sesaminol, sesaminol_DG, and sesaminol_TG. This approach identifies major and minor QTLs, improving understanding of genetic additive effects that enhance the role and lignan content in sesame breeding programs. Although the QTL analysis allowed the identification of nearby regions, further studies are needed to functionally validate candidate genes.

4.2. Enhancing linkage mapping *via* optimized GBS strategies

As NGS technology has advanced, GBS has emerged as a cost-effective, efficient, and representative method for gathering genetic information from large populations (Elshire *et al.* 2011; He *et al.* 2014). GBS is useful for analyzing genetic populations requiring genotyping data from many individuals, especially for orphan (or minor) crops, where large genome sizes make sequencing costly and the development of specific tools such as arrays is limited (Kim *et al.* 2016). GBS success can depend on the restriction enzyme selection used in the process. For crops with large genomes, using a combination of two REs increases the number of randomized sites, thus improving genome coverage compared to using a single enzyme (Poland and Rife *et al.* 2012; Fu *et al.* 2016). As the restriction enzyme efficiency varies by crop, several enzyme combinations are often tested to identify the most effective one for each specific application (Sonah *et al.* 2013; Hamblin and Rabbi 2014). For example, in soybean, *in silico* analysis of restriction sites recognized by *Mse*I, *Ape*KI, and *Pst*I as been conducted to determine their genome coverage

and optimal enzyme combinations for GBS (Sonah *et al*. 2013).

Previous GBS studies in sesame used single enzymes (*Ape*KI and *Taq*I) for linkage mapping (Uncu *et al.* 2016; Zhang *et al.* 2018) and genetic diversity studies (Seay *et al.* 2024), and double enzymes (*Pst*I/*Mse*I) for GWAS (He *et al.* 2019). In our study, libraries generated with both *Pst*I/*Msp*I and *Ape*KI identified 1,049,549 SNPs with *Ape*KI and added 52,921 SNPs with *Pst*I/*Msp*I, improving the integrity of the genetic map. However, 21,443 SNPs overlapped between the two enzyme strategies, indicating that *Pst*I/*Msp*I recognition sequences are not widely distributed and do not differ significantly from those of *Ape*KI. Although this additional approach represents only a modest 5% increase over the approximately 10% coverage achieved with the *Ape*KI library alone, our results provide valuable information for selecting restriction enzymes to optimize GBS strategies in sesame research.

4.3. Evaluation and application of molecular markers for enhancing lignan content in sesame

In breeding new varieties, a new paradigm for maximizing the industrial value of crops involves targeting useful ingredients that offer health benefits (Hansson *et al.* 2018; Loskutov and Khlestkina 2021). However, breeding for enhanced ingredient content is challenging because verifying ingredient presence requires long cultivation periods and breeding processes are vulnerable to losses from seed or plant damage. To address these limitations, marker-assisted selection (MAS) strategies using molecular markers associated with quantitative traits are applied in various crops (He *et al.* 2014). For example, in wheat, MAS is used to select backcrossed individuals for introgression of the *Glu-B1i* allele, improving bread quality by increasing protein content and gluten strength (Cha *et al.* 2024). Similarly, SNP markers in rice categorize *Waxy* gene haplotypes, regulating amylose content to breed varieties with desired cooked textures (Chen *et al.* 2010). For quantitative traits controlled by multiple genes, the phenotype is often maximized by accumulating key genes with high expression levels. Consequently, breeding strategies such as gene pyramiding are employed to combine alleles for each trait for additive effects (Ashikari and Matsuoka 2006). In soybeans, high-protein selection markers targeting three major additive QTLs generate transgressive progenies (Zhang *et al.* 2015).

In the present study, CAPS markers were developed for sesame breeding to produce high-quality cultivars with maximized lignan content. These markers are specific to the lignan-rich cultivar Gomazou and are located within the major QTL regions (*qLIGNAN6* and *qLIGNAN11*). The correlation between markers and lignan content in two RIL populations and 100 germplasm lines showed that pyramiding beneficial alleles from two major QTLs effectively maximized lignan content. These results demonstrate that the markers are effective for selecting individuals with high-lignan content, minimizing seed loss associated with functional analysis, and are promising for future breeding. However, to improve selection accuracy, further functional validation of candidate genes is required, as results obtained under limited conditions and genetic backgrounds may not fully represent the phenotypic variability of QTLs across different genetic resources. Future research should therefore include experiments across diverse environmental conditions and genetic backgrounds to better understand and validate the effectiveness of these markers in sesame.

5. Conclusion

This study identified two loci, *qLIGNAN6* and *qLIGNAN11*, associated with lignan content using a cross between the sesame cultivars Areum and Gomazou and the F_9 RIL population. GBS libraries were constructed using both a single restriction enzyme (*Ape*KI) and a combination of two restriction enzymes (*Pst*I/*Msp*I). The results indicated that the enzyme combination successfully covered regions of the genome missed when using a single enzyme, thereby enhancing linkage map integrity. Marker-assisted selection developed molecular markers specific to these loci, demonstrating that stacking beneficial alleles from both loci significantly enhances lignan content. These markers can be effectively utilized in breeding programs, enabling efficient selection of sesame plants with higher lignan content. In conclusion, this study provides valuable genetic resources and practical tools for enhancing the nutritional value of sesame. It establishes a foundation for future molecular breeding efforts, facilitating the development of improved sesame varieties through targeted marker applications.

Acknowledgements

This research was supported by the Cooperative Research Program for Agriculture Science and Technology Development, Rural Development Administration, Republic of Korea (Project PJ017387072024). It was also supported by the 2024 Collaborative Research Program Between the University and the Rural Development Administration, Republic of Korea.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this manuscript, the authors used ChatGPT 4.0 to assist in improving the clarity and readability of the language. All content generated was carefully reviewed and revised by the authors as necessary, and the authors take full responsibility for the final content of the article.

Appendices associated with this paper are available at https://doi.org/10.1016/j.jia.2025.02.025

References

Andargie M, Vinas M, Rathgeb A, Möller E, Karlovsky P. 2021. Lignans of sesame (*Sesamum indicum* L.): A comprehensive review. *Molecules*, **26**, 883.

Ashikari M, Matsuoka M. 2006. Identification, isolation, and pyramiding of quantitative trait loci for rice breeding. *Trends in Plant Science*, **11**, 344–350.

Cha J K, Park H, Kwon Y, Lee S M, Jang S G, Kwon S W, Lee J H. 2024. Synergizing breeding strategies *via* combining speed breeding,

phenotypic selection, and marker-assisted backcrossing for the introgression of *Glu-B1i* in wheat. *Frontiers in Plant Science*, **15**, 1402709.

Chen M H, Fjellstrom R G, Christensen E F, Bergman C J. 2010. Development of three allele-specific codominant rice *Waxy* gene PCR markers suitable for marker-assisted selection of amylose content and paste viscosity. *Molecular Breeding*, **26**, 513–523.

Cingolani P, Platts A, Wang L L, Coon M, Nguyen T, Wang L, Land S J, Lu X, Ruden D M. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w^{1118}; *iso*-2; *iso*-3. *Fly*, **6**, 80–92.

Cui C, Mei H, Liu Y, Zhang H, Zheng Y. 2017. Genetic diversity, population structure, and linkage disequilibrium of an association-mapping panel revealed by genome-wide SNP markers in sesame. *Frontiers in Plant Science*, **8**, 1189.

Davey J W, Hohenlohe P A, Etter P D, Boone J Q, Catchen J M, Blaxter M L. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, **12**, 499–510.

Dossou S S K, Xu F T, Dossa K, Zhou R, Zhao Y Z, Wang L H. 2023. Antioxidant lignans sesamin and sesamolin in sesame (*Sesamum indicum* L.): A comprehensive review and future prospects. *Journal of Integrative Agriculture*, **22**, 14–30.

Elshire R J, Glaubitz J C, Sun Q, Poland J A, Kawamoto K, Buckler E S, Mitchell S E. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, **6**, e19379.

Fischer M C, Rellstab C, Leuzinger M, Roumet M, Gugerli F, Shimizu K K, Holderegger R, Widmer A. 2017. Estimating genomic diversity and population differentiation — An empirical comparison of microsatellite and SNP variation in *Arabidopsis halleri*. *BMC Genomics*, **18**, 69.

Fu Y B, Peterson G W, Dong Y. 2016. Increasing genome sampling and improving SNP genotyping for genotyping-by-sequencing with new combinations of restriction enzymes. *G3* (Bethesda), **6**, 845–856.

Hamblin M, Rabbi I. 2014. The effects of restriction-enzyme choice on properties of genotyping-by-sequencing libraries: A study in cassava (*Manihot esculenta*). *Crop Science*, **54**, 2603–2608.

Hansson S O, Åman P, Becker W, Koning D J, Lagerkvist C, Larsson I, Lehrman A, Risérus U, Stymne S. 2018. Breeding for public health: A strategy. *Trends in Food Science & Technology*, **80**, 131–140.

He J, Zhao X, Laroche A, Lu Z X, Liu H, Li Z. 2014. Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Frontiers in Plant Science*, **5**, 484.

He Q, Xu F, Min M H, Chu S H, Kim K W, Park Y J. 2019. Genome-wide association study of vitamin E using genotyping by sequencing in sesame (*Sesamum indicum*). *Genes Genomics*, **41**, 1085–1093.

Jaganathan D, Bohra A, Thudi M, Varshney R K. 2020. Fine mapping and gene cloning in the post-NGS era: Advances and prospects. *Theoretical and Applied Genetics*, **133**, 1791–1810.

Kim C, Guo H, Kong W, Chandnani R, Shuang L S, Paterson A H. 2016. Application of genotyping by sequencing technology to a variety of crop breeding programs. *Plant Science*, **242**, 14–22.

Kim H M, Lee J M, Park J Y, Lee S L, Han S, Kim H Y, Son D W, Choi S Y, Lee S H. 2008. Phytosterols and lignans from the sesame dregs of *Sesamum indicum*. *Korean Journal of Plant Resources*, **21**, 420–426.

Kim M, Kim S, Kim J, Oh E, Kim S, Lee J, Lee E, Lee M. 2022. Antioxidant and hepatoprotective effects of different varieties of sesame (*Sesamum indicum* L.) with variation in lignan content. *Journal of the Korean Society of Food Science and Nutrition*, **51**, 660–670.

Kim S, Lee E, Lee J, An Y J, Oh E, Kim J I, Kim S W, Kim M Y, Lee M H, Cho K S. 2023. Identification of QTLs and allelic effect controlling lignan content in sesame (*Sesamum indicum* L.) using QTL-seq approach. *Frontiers in Genetics*, **14**, 1289793.

Lee E, Yang X, Ha J, Kim M Y, Park K Y, Lee S H. 2021. Identification of a locus controlling compound raceme inflorescence in mungbean [*Vigna radiata* (L.) R. Wilczek]. *Frontiers in Genetics*, **12**, 642518.

Li Q Z, Zuo Z W, Liu Y. 2023. Recent status of sesaminol and its glucosides: Synthesis, metabolism, and biological activities. *Critical Reviews in Food Science and Nutrition*, **63**, 12043–12056.

Loskutov I G, Khlestkina E K. 2021. Wheat, barley, and oat breeding for health benefit components in grain. *Plants* (Basel), **10**, 86.

Miao C, Fang J, Li D, Liang P, Zhang X, Yang J, Schnable J C, Tang H. 2018. Genotype-corrector: Improved genotype calls for genetic mapping in F_2 and RIL populations. *Scientific Reports*, **8**, 10088.

Moazzami A, Andersson R, Kamal-Eldin A. 2006. HPLC analysis of sesaminol glucosides in sesame seeds. *Journal of Agricultural and Food Chemistry*, **54**, 633–638.

Murata J, Ono E, Yoroizuka S, Toyonaga H, Shiraishi A, Mori S, Tera M, Azuma T, Nagano A J, Nakayasu M. 2017. Oxidative rearrangement of (+)-sesamin by CYP92B14 co-generates twin dietary lignans in sesame. *Nature Communications*, **8**, 1–10.

Neff M M, Turk E, Kalishman M. 2002. Web-based primer design for single nucleotide polymorphism analysis. *Trends in Genetics*, **18**, 613–615.

Nguyen T K, Yu J, Choi W, In B C, Lim J H. 2018. Optimization of genotyping-by-sequencing (GBS) in *Chrysanthemums*: Selecting proper restriction enzymes for GBS library construction. *Horticultural Science and Technology*, **36**, 108–114.

Ono E, Nakai M, Fukui Y, Tomimori N, Fukuchi-Mizutani M, Saito M, Satake H, Tanaka T, Katsuta M, Umezawa T. 2006. Formation of two methylenedioxy bridges by a *Sesamum* CYP81Q protein yielding a furofuran lignan, (+)-sesamin. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 10116–10121.

Ben Othman S, Katsuno N, Kanamaru Y, Yabe T. 2015. Water-soluble extracts from defatted sesame seed flour show antioxidant activity *in vitro*. *Food Chemistry*, **175**, 306–314.

Pathak N, Rai A K, Kumari R, Bhat K V. 2014. Value addition in sesame: A perspective on bioactive components for enhancing utility and profitability. *Pharmacognosy Reviews*, **8**, 147–155.

Pearson K. 1896. Mathematical contributions to the theory of evolution. III. Regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society A*, **187**, 253–318.

Poland J A, Brown P J, Sorrells M E, Jannink J L. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE*, **7**, e32253.

Poland J A, Rife T. 2012. Genotyping-by-sequencing for plant breeding and genetics. *The Plant Genome Journal*, **5**, 92–102.

Pootakham W, Sonthirod C, Naktang C, Jomchai N, Sangsrakru D, Tangphatsornruang S. 2016. Effects of methylation-sensitive enzymes on the enrichment of genic SNPs and the degree of genome complexity reduction in a two-enzyme genotyping-by-sequencing (GBS) approach: A case study in oil palm (*Elaeis guineensis*). *Molecular Breeding*, **36**, 154.

Ruperao P, Bajaj P, Subramani R, Yadav R, Reddy Lachagari V B, Lekkala S P, Rathore A, Archak S, Angadi U B, Singh R, Singh K, Mayes S, Rangan P. 2023. A pilot-scale comparison between single and double-digest RAD markers generated using GBS strategy in sesame (*Sesamum indicum* L.). *PLoS ONE*, **18**, e0286599.

Satake H, Ono E, Murata J. 2013. Recent advances in the metabolic engineering of lignan biosynthesis pathways for the production of transgenic plant-based foods and supplements. *Journal of Agricultural and Food Chemistry*, **61**, 11721–11729.

Schröder S, Mamidi S, Lee R, Osorno J, Pe M. 2016. Optimization of genotyping by sequencing (GBS) data in common bean (*Phaseolus vulgaris* L.). *Molecular Breeding*, **36**, 1–9.

Seay D, Szczepanek A, De La Fuente G N, Votava E, Abdel-Haleem H. 2024. Genetic diversity and population structure of a large USDA sesame collection. *Plants* (Basel), **13**, 1765.

Sonah H, Bastien M, Iquira E, Tardivel A, Légaré G, Boyle B, Normandeau É, Laroche J, Larose S, Jean M, Belzile F. 2013. An improved genotyping by sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. *PLoS ONE*, **8**, e54603.

Uncu A, Frary A, Karlovsky P, Doganlar S, Stamatov S. 2016. High-throughput single nucleotide polymorphism (SNP) identification and mapping in the sesame (*Sesamum indicum* L.) genome with genotyping by sequencing (GBS) analysis. *Molecular Breeding*, **36**, 173.

Wang L, Xia Q, Zhang Y, Zhu X, Zhu X, Li D, Ni X, Gao Y, Xiang H, Wei X, Yu J, Quan Z, Zhang X. 2016. Updated sesame genome assembly and fine mapping of plant height and seed coat color QTLs using a new high-density genetic map. *BMC Genomics*, **17**, 31.

Wang Y, Cao X, Zhao Y, Fei J, Hu X, Li N. 2017. Optimized double-digest genotyping by sequencing (ddGBS) method with high-density SNP markers and high genotyping accuracy for chickens. *PLoS ONE*, **12**, e0179073.

Wu D, Wang X P, Zhang W. 2019. Sesamolin exerts anti-proliferative and apoptotic effect on human colorectal cancer cells *via* inhibition of JAK2/STAT3 signaling pathway. *Cellular and Molecular Biology*, **65**, 96–100.

Xu F, Zhou R, Dossou S S K, Song S, Wang L. 2021. Fine mapping of a major pleiotropic QTL associated with sesamin and sesamolin variation in sesame (*Sesamum indicum* L.). *Plants* (Basel), **10**, 1343.

Yadav R, Kalia S, Rangan P, Pradheep K, Rao G P, Kaur V, Pandey R, Rai V, Vasimalla C C, Langyan S, Sharma S, Thangavel B, Rana V S, Vishwakarma H, Shah A, Saxena A, Kumar A, Singh K, Siddique K H M. 2022. Current research trends and prospects for yield and quality improvement in sesame, an important oilseed crop. *Frontiers in Plant Science*, **13**, 863521.

Yasumoto S, Katsuta M. 2006. Breeding a high-lignan-content sesame cultivar in the prospect of promoting metabolic functionality. *Japan Agricultural Research Quarterly*, **40**, 123–129.

Zhang Y, Wang L, Gao Y, Li D, Yu J, Zhou R, Zhang X. 2018. Genetic dissection and fine mapping of a novel *dt* gene associated with determinate growth habit in sesame. *BMC Genetics*, **19**, 38.

Zhang Y H, Liu M F, He J B, Wang Y F, Xing G N, Li Y, Yang S P, Zhao T J, Gai J Y. 2015. Marker-assisted breeding for transgressive seed protein content in soybean [*Glycine max* (L.) Merr]. *Theoretical and Applied Genetics*, **128**, 1061–1072.

Executive Editor-in-Chief Xueyong Zhang
Managing Editor Ning Wang