Molecular variation of lipoxygenase-associated genes in grain of commercial Mexican soybean cultivars

Variación molecular de genes asociados a lipoxigenasas en grano de variedades de soya comercial mexicana

Variação molecular de genes associados a lipoxigenases em grãos de variedades comerciais de soja mexicana

Mónica López Fernández[1]
Ana María Sifuentes Rincón[2]
Francisco Alejandro Paredes Sánchez[3]
Nicolás Maldonado Moreno[4]
Homar Rene Gill Langarica[1*]

[1]Laboratorio de Biotecnología Vegetal, Centro de Biotecnología Genómica, Instituto Politécnico Nacional, Reynosa, Tamaulipas, México.
[2]Laboratorio de Biotecnología Animal, Centro de Biotecnología Genómica, Instituto Politécnico Nacional, Reynosa, Tamaulipas, México.
[3]Universidad Autónoma de Tamaulipas. IA-UAM, Mante, Tamaulipas, México.
[4]INIFAP-Campo Experimental Las Huastecas. Carretera Tampico-Mante, km 55. Villa Cuauhtémoc, Altamira, Tamaulipas, México.

Rev. Fac. Agron. (LUZ). 2024, 41(2): e244111
ISSN 2477-9407
DOI: https://doi.org/10.47280/RevFacAgron(LUZ).v41.n2.01

Crop production
Associate editor: Dr. Jorge Vilchez-Perozo
University of Zulia, Faculty of Agronomy
Bolivarian Republic of Venezuela

Received: 19-01-2024
Accepted: 28-02-2024
Published: 18-03-2024

Keywords:
Soybean grain
Lipoxygenase
Lox genes
Single nucleotide polymorphisms

Abstract

Lipoxygenase enzymes encoded by the *Lox*1, *Lox*2 and *Lox*3 genes play a crucial role in soybean grain, particularly in the development of off-flavors. Understanding molecular variation within *Lox* genes is essential for the improvement of soybean organoleptic traits. This study investigated the genetic variation in the internal regions of the *Lox*1, *Lox*2, and *Lox*3 genes in mature grain of commercially grown soybean cultivars in Mexico. Genomic DNA from a diverse panel of Mexican soybean cultivars was analyzed using resequencing techniques and *in-silico* analysis. Single nucleotide polymorphisms (SNP) within the *Lox*1, *Lox*2, and *Lox*3 genes were identified and characterized. The findings indicated that *Lox*3 gene displayed lower genetic variability compared to *Lox*1 and *Lox*2 genes, specifically, was identified a total of 26 SNPs in the *Lox*1 gene, 11 SNPs in the *Lox*2 gene, and 5 SNPs in the *Lox*3 gene among the examined cultivars. A non-synonymous SNP variant of the C/C genotype located in exon 6 of the *Lox*2 gene was associated with a destabilizing effect on the lipoxygenase 2 enzyme in the Guayparime S-10 and Huasteca 300 cultivars. These findings provide insights into the molecular variation of lipoxygenase-associated genes in Mexican soybean cultivars.

*Corresponding author: hgill@ipn.mx

Resumen

Las enzimas lipoxigenasas codificadas por los genes *Lox*1, *Lox*2 y *Lox*3 desempeñan un papel crucial en el grano de soya, particularmente en el desarrollo de sabores desagradables. Comprender la variación molecular dentro de los genes *Lox* es esencial para el mejoramiento de las características organolepticas de la soya. Este estudio investigó la variación genética en las regiones internas de los genes *Lox*1, *Lox*2 y *Lox*3 en grano maduro de variedades de soya cultivados comercialmente en México. Se analizó el ADN genómico de un panel diverso de variedades de soya mexicanas mediante técnicas de resecuenciación y análisis *in silico*. Se identificaron y caracterizaron polimorfismos de un solo nucleótido (SNP) dentro de los genes *Lox*1, *Lox*2 y *Lox*3. Los hallazgos indicaron que el gen *Lox*3 mostró una menor variabilidad genética en comparación con los genes *Lox*1 y *Lox*2; específicamente, fueron identificados un total de 26 SNPs en el gen *Lox*1, 11 SNPs en el gen *Lox*2 y 5 SNPs en el gen *Lox*3 entre las variedades examinadas. Un SNP no sinónimo variante del genotipo C/C ubicado en el exón 6 del gen *Lox*2, fue asociado con un efecto desestabilizador en la enzima lipoxigenasa 2 en las variedades Guayparime S-10 y Huasteca 300. Estos hallazgos proporcionan información sobre la variación molecular de los genes asociados a las lipoxigenasas en variedades de soya mexicanas.

Palabras clave: grano de soya, lipoxigenasa, genes *Lox*, polimorfismos de un solo nucleotido.

Resumo

As enzimas lipoxigenases codificadas pelos genes *Lox*1, *Lox*2 e *Lox*3 desempenham um papel crucial na grão de soja, particularmente no desenvolvimento de sabores estranhos. Compreender a variação molecular dentro dos genes *Lox* é essencial para a melhoria das características organolépticas da soja. Este trabalho visou pesquisar a variação genética nas regiões internas dos genes *Lox*1, *Lox*2 e *Lox*3 em grãos maduros de variedades de soja cultivadas comercialmente no México. O DNA genômico de um painel diversificado de variedades mexicanas de soja foi analisado usando técnicas de ressequenciamento e análise in silico. Foram identificados y caracterizados polimorfismos de nucleotídeo único (SNP) nos genes *Lox*1, *Lox*2 e *Lox*3. Os resultados indicaram que o gene *Lox*3 apresentou menor variabilidade genética em comparação com os genes *Lox*1 e *Lox*2, especificamente, um total de 26 SNPs no gene *Lox*1, 11 SNPs no gene *Lox*2 e 5 SNPs no gene *Lox*3 foram identificados entre as variedades examinadas. Uma variante SNP não-sinônima do genótipo C/C localizada no exon 6 do gene *Lox*2 foi associada a um efeito desestabilizador na enzima lipoxigenase 2 nas variedades Guayparime S-10 e Huasteca 300. Estes interesantes ressultados proporcionam informações sobre a variação molecular dos genes associadas a lipoxigenases nas variedades mexicanas de soja.

Palabras-chave: grão de soja, lipoxigenase, genes *Lox*, polimorfismos de nucleotídeo único.

Introduction

Soybean grain [*Glycine max* (L.) Merr.] is globally recognized for its significance as a protein and oil source in human diets and as a balanced animal feed (Singh *et al*., 2020). However, the consumption of soybean products is limited due to the presence of anti-nutritional compounds like lipoxygenase enzymes that trigger volatile compound reactions, resulting in unpleasant flavors in commercial soybean-based products (Carpentieri-Pipolo *et al*., 2021). The lipoxygenases facilitate the oxidation of polyunsaturated fatty acids such as linoleic acid (18:2) and α-linolenic acid (18:3), resulting in the production of hydroperoxides of unsaturated fatty acids (Lenis *et al*., 2010). The enzymatic activity of soybean lipoxygenases has significance, particularly in relation to the generation of undesirable flavors during grain processing (Wang *et al*., 1994; Lee *et al*., 2014). The Lox (lipoxygenase) *loci* play a crucial role in the expression of the three known lipoxygenase genes in mature soybean grains, these genes are located in three specific genomic regions. *Lox*1 (Glyma13g347600; GmLox1) and *Lox*2 (Glyma13g347500; GmLox2) genes are found on chromosome 13, while the *Lox*3 gene (Glyma13g026300; GmLox3) is located on chromosome 15 (Shin *et al*., 2012). The structural composition of all three *Lox* genes consists of nine exons and eight introns (Reinprecht *et al*., 2011). The *Lox*2 gene is primarily responsible for the undesirable beany flavor in mature soybean grains (Davies *et al*., 1987). In past years, research on lipoxygenases has focused on identifying natural soybean mutants and generating single, double and triple null mutants through mutagenesis that lack lipoxygenase activity (Suda *et al*., 1995; Lenis *et al*., 2010). Conventional genetic improvement techniques have also allowed breeders to develop lines and cultivars that exhibit a null effect of lipoxygenases (Lee *et al*., 2014). In Mexico, soybean production takes place at latitudes below 25° North with the soybean plant adapted to short-day photoperiods of approximately 13.3 light hours (Hinson and Hartwig, 1977). In Mexico, commercial soybeans are grown in states such as Sonora, Sinaloa, and Tamaulipas and are produced under a seasonal production system. The productive zones of Sonora and Sinaloa are classified as the Mexican dry tropics, while the southern zone of Tamaulipas and some coastal states are known as the Mexican humid tropics. Both dry and humid tropics, have their own genetic plant breeding program (Ascencio and Maldonado, 1998) aimed at developing cultivars with high grain yield and tolerance to pathogens (Maldonado and Ascencio, 2012; Rodríguez-Cota *et al*., 2017). It is worth noting that Mexican soybean genetic programs employ radiation with cobalt-60 to generate genetic variation during the improvement process (Rodríguez-Cota *et al*., 2017). The objective of this study was to investigate the genetic variability of *Lox* genes through resequencing and *in silico* analysis in a Mexican soybean population, alongside soybean materials devoid of the influence of lipoxygenases.

Materials and methods

In this study, a total of 13 soybean materials were analyzed. Among them, 11 cultivars are commercially available cultivars and they were developed in Mexico by the National Institute of Forestry, Agriculture, and Livestock Research (INIFAP) to suit the conditions of the Mexican humid and dry tropics. The Huasteca-named cultivars, along with Tamesi and Vernal cultivars, were generously donated by MSc. Nicolás Maldonado Moreno, a soybean breeder from INIFAP, for research and academic purposes. Additionally, the Nainari, Guayparime S-10, and Suaqui 86 soybean cultivars were obtained from a commercial supplier in the state of Sonora, Mexico. The materials JP30790 and JP28955 were generously provided for research and academic purposes by Dr. Shoshi Kikuchi from the Genetic Resources Center of the National Agriculture and Food Research Organization at the National Institute of Agrobiological Sciences (NIAS) in Japan.

The soybean germplasm bank at NIAS (https://www.gene.affrc.go.jp/databases_plant_search_en.php) confirms that JP30790 and JP28955 materials are free from the effect of lipoxygenase or the triple null effect of *Lox* genes in mature grain. The Nainari variety is a mutant variety developed by cobalt 60 from seeds of Suaqui 86 variety and the Guayparime S-10 variety was developed from the Nainari variety through crossing and selection (Rodríguez-Cota *et al*., 2017) (table 1).

Table 1. Soybean germplasm analyzed.

Number	Germplasm	Adaptation	Progenitors
1	JP30790*	USA	Unknown
2	JP28955*	USA	Unknown
3	Vernal^{+}	Mexico	D77-12244 X Bedford
4	Huasteca 100°	Mexico	Santa Rosa X Jupiter
5	Huasteca 200°	Mexico	F815344 X Santa Rosa
6	Huasteca 300°	Mexico	H82-1930 X H80-2535
7	Huasteca 400°	Mexico	DM301
8	Tamesi°	Mexico	Santa Rosa X H80-2535
9	Huasteca 600°	Mexico	H88-1880 X H88-3868
10	Huasteca 700°	Mexico	Santa Rosa X F81-5517
11	Nainari**	Mexico	Derived from Suaqui 86
12	Suaqui 86**	Mexico	(Rad X Cajeme) X (Tetabiate X Cajeme)
13	Guayparime S-10**	Mexico	Nainari X PI-171443

*= Soybean accessions deposited in the Japan NIAS gene bank; += North of the Tamaulipas state; °= South of the Tamaulipas state (humid tropics); **= South of the Sonora state (dry tropics).

Sequencing of mature grain *Lox* genes

To amplify the specific regions of the *Lox*1, *Lox*2, and *Lox*3 genes, specific primers were designed using the DNAStar LaserGene software. The reference sequences of these genes from the cultivar Williams 82 V1.1 (GFC_000004515.3), deposited in the NCBI database, were used for primer design (table 2). For PCR amplification, 25 ng.µL of template DNA, isolated from young leaves with Wizard® genomic kit protocol and the PCR Master Mix enzyme from Promega were used. The thermocycling conditions for each gene were as follows: an initial denaturation step at 94 °C for 2 min, followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 50 °C for 30 s, and elongation at 72 °C for 2.5 min. A final elongation step was performed at 72 °C for 5 min. The quality and specificity of the PCR products were assessed by analyzing them on a 1.2 % agarose gel electrophoresis, which allowed verification of the expected amplicon sizes and confirmation of specific amplification (Kumar and Rani, 2019).

Library construction and sequencing

For sequencing the amplified *Lox* gene fragments, the samples were sent to the LANMDA-CBG-IPN National Laboratory. To prepare the libraries for sequencing, the Nextera Flex Library Kit was used. Individual indices were incorporated into the libraries for barcoding, allowing for multiplexing of samples. The library preparation process followed the Illumina reference guide #1000000025416 v07. Finally, the prepared libraries were sequenced using the MiniSeq™ Sequencing System, which is an Illumina sequencing platform. This system performs high-throughput sequencing of the libraries, generating sequence data for each of the *Lox* genes in the samples.

Bioinformatic analysis of sequencing data

The sequence reads obtained from the MiniSeq™ Sequencing System were aligned to the reference genome Williams 82 using the Burrows-Wheeler Aligner (BWA-MEM) v0.7. This alignment process helps to map the reads to their corresponding genomic positions. The aligned reads were then processed using the Picard v1.135 program, the processed reads were saved in BAM files. To identify variations in the sequenced data, the Genomic Variant Call Format (GVCF) workflow was employed, specifically using the HaplotypeCaller program. The identified SNPs and single nucleotide variants (SNVs) were saved in Variant Call Format (VCF) files and underwent filtering based on specific criteria. These criteria include depth-normalized variant confidence (QD) <2.0, mapping quality (MQ) <40.0, strand bias (FS) >60.0, HaplotypeScore >13.0, MQRankSum < -12.5, and Read-PosRank-Sum < -8.0. Filtering helps to ensure the quality and reliability of the detected variations. Finally, the Infogen software (Balzarini and Di Rienzo, 2004) was used to determine allelic frequencies.

Predictive analysis of the effect of SNPs on the protein

The potential effect of each SNP on the structure of the *Lox*1, *Lox*2, and *Lox*3 genes was analyzed using the Sanger Institute's

Table 2. Specific primers for the amplification of grain *Lox* genes of 13 soybean cultivars.

Number	Direction	Primers sequence	Amplicon size*	Position of amplicon
		***Lox*1 gene**		
1	Forward	5' TTCTTCTTCTTTATTTTCTCATTT 3'	2,567bp	32-2598
	Reverse	5' GCCAAATTGTGCTCTCA 3'		
2	Forward	5' ACAATTATCCCCTTACCAGTG 3'	2,495bp	1648-4142
	Reverse	5' TGATGACAGGAGCTAAACACAAAC 3'		
		***Lox*2 gene**		
1	Forward	5' AGCTTTTCTTTTTCTTGTT 3'	1,530bp	11-1540
	Reverse	5' AAAAATAAATCAGAATCATAGCAC 3'		
2	Forward	5' CTTTGGGAGCAGGGGAGTC 3'	3,461bp	741-4201
	Reverse	5' AATAGTGCTCGGTGCTCTTA 3'		
		***Lox*3 gene**		
1	Forward	5' TGATGCGCAAGAATGTG 3'	4,346bp	105-4340
	Reverse	5' CACAAAGCAAAGCAGTA 3'		

*= nucleotide base pairs

PFAM 34.0 database (http://pfam.xfam.org/) to identify traits of the L-1, L-2 and L-3 proteins and obtain information about the function of the protein. Subsequently, through predictions in the PROVEAN program (http://provean.jcvi.org/seq_submit.php) the amino acid changes produced by each SNP were analyzed and their potential effect on protein stability was determined. The algorithm used 173 supporting sequences and the score thresholds for prediction were as follows: variants with a score equal to or less than -2.5 are considered "destabilizing" and variants with a score greater than -2.5 are considered "neutral". This program pools BLAST results using CD-HIT with a parameter of 75 % overall sequence identity. The 30-best cluster of closely related sequences from the set of supporting sequences that is used to generate the prediction, a delta alignment score is calculated for each supporting sequence. Scores are averaged within and across clusters to generate the final PROVEAN score. The last step consisted in analyzing the mutated sequence in the Eukariotic Linear Motif (ELM) platform (http://elm.eu.org/search.html) to identify putative motifs and finally modeling the protein in the SWISS-MODEL program (https://swissmodel.expasy.org/interactive) to determine its functionality.

Results and discussion

Size and traits of *Lox* genes

The sequences of exons and introns obtained from the *Lox*1, *Lox*2, and *Lox*3 genes in Mexican soybean grains were compared to the reference genome Williams 82. The coverage for *Lox*1, *Lox*2, and *Lox*3 genes was 96 % (4,265 bp), 97.5 % (4,297 bp), and 100 % (4,346 bp), respectively. Interestingly, the average sequence size of exons and introns observed in the Mexican soybean population was found to be greater than that reported by (Reinprecht *et al*., 2011; Lee *et al*., 2014), who studied a mutant genotype and wild-type genotype. The differences in sequence size are particularly noticeable in the *Lox*1 gene, as both studies reported a deletion of 74 bp in exon 8 of the mutant genotype.

***Lox*1 gene**

The *Lox*1 gene sequences in the Mexican soybean population exhibited 26 SNPs compared to the reference sequence of the Williams 82 cultivar (table 3). Other studies have reported a lower number of SNPs associated with normal and null activity of the *Lox*1 when compared to the reference sequence of the Williams 82 cultivar (Lenis *et al*., 2010; Reinprecht *et al*., 2011; Lee *et al*., 2014). Of the 26 SNPs found in the *Lox*1 gene of the Mexican soybean population, 14 were found in introns, while 12 were identified in exons. Out of the 12 SNPs identified in exons, eight were non-synonymous changes, meaning they resulted in amino acid substitutions. Among these eight SNPs, five were transversions (C/A, A/T, G/C) and three were transitions (A/G, G/A, G/A). The transversion SNPs were located in exon 2 (C/A), resulting in a His-Asn amino acid change; in exon 4 (A/T), leading to a Glu-Asp amino acid change; and in exon 6 (G/C), causing a Ser-Thr amino acid change. The most variable exon was exon 9, which exhibited five non-synonymous changes. It included two transversions (G/C, C/A) resulting in Val-Leu and Leu-Ile amino acid changes, as well as three transitions (A/G, G/A, G/A) leading to amino acid changes of Ile-Val, Ala-Thr, and Gly-Asp, respectively (table 3).

***Lox*2 gene**

In the *Lox*2 gene of the Mexican soybean population, a total of 17 SNPs were identified (table 3). Among these SNPs, 11 were located in non-coding regions, while six were found in coding regions. Out of the SNPs detected in exons, three SNPs were transversion with non-synonymous changes. One transversion SNP (G/C) was found in exon 6, resulting in a Glu-Asp amino acid change. Additionally, two transversion SNPs (C/A) in exon 9 led to Pro-Thr and Pro-His amino acid changes, respectively. The predictive analysis of non-synonymous SNPs in the Mexican soybean population revealed that the G/C SNP present in exon 6 destabilized the L-2 protein in the homozygous C/C genotype in Guayparime S-10 and Huasteca 300 cultivars. Other non-synonymous SNPs in the *Lox*2 gene have been reported in previous studies. For example, Wang *et al*. (1994), Reinprecht *et al*. (2011) and Lee *et al*. (2014) detected a non-synonymous exchange of T/A in exon 8, resulting in an amino acid change from histidine to glutamine. They observed that this mutation affected an iron ligand essential for the activity of L-2, causing enzyme disfunction. Additionally, Reinprecht *et al*. (2011) detected an A/A SNP at position 678, leading to the substitution of threonine with lysine in the *Lox*2 gene. The change to observed in the nsSNP G/C from exon 6 in the Mexican soybean population led to the substitution of glutamic acid (GAG) to aspartic acid (GAC) (tables 3, 4) and consequently, an error occurred in the conformation of the LH2 globular domain (Lipoxygenase homology 2) within the protein sequence due to the low conservation of the DOC_PP4_FxxP_1 motifs (position 2-5) and DOC_USP7_MATH_1 (position 5-9). These modifications altered the conserved ligand-binding site of the protein in the mutated protein sequence compared to the corresponding protein sequence of the Williams 82 reference material (accession SM00308, position 17-176). By examining the direct ancestors of the Guayparime S-10 cultivar, it was observed that both the Nainari mutant and the normal Suaqui 86 cultivars did not contribute the C allelic variant to the Guayparime S-10 (tables 3 and 4). This is evident from their normal G/G genotype. Instead, the C allelic variant in the Guayparime S-10 variety originated from the PI-171443 line, which served as a direct parent in the genetic cross with the Nainari mutant variety during the development of Guayparime S-10. The PI-171443 line carries the *Rym*1 and *Rym*2 genes, which confer tolerance to Mung Bean Yellow Mosaic Begomovirus (Rodríguez-Cota *et al*., 2017; Rani and Kumar, 2020). The allelic contribution of the PI-171443 line to the soybean population in the Mexican dry tropics, is likely associated with an improvement in the reduced activity of the L-2 enzyme in mature grains where Guayparime S-10 exhibited a desirable attribute of low beany or rancid flavor (López-Fernández *et al*., 2022).

The Huasteca 300 cultivar in the Mexican soybean population shares genetic information with the Huasteca 100, Tamesi, and Huasteca 600 cultivars, which were developed using the Iowa and Jupiter cultivars parents in single (Huasteca 100) and double genetic crosses (Huasteca 300, Tamesi, and Huasteca 600). However, the Huasteca 300 cultivar also possesses unique genetic information contributed by the parent F76-9835, which is not found in any other Mexican soybean cultivar (table 1). Upon genotyping the nsSNP G/C, it was observed that the Huasteca 100, Tamesi, and Huasteca 600 cultivars do not exhibit the allelic change from G to C (tables 3 and 4). Therefore, the C allelic variant causing destabilization of the L-2 protein in the Huasteca 300 cultivar is likely attributed to the parent F76-9835. The F76-9835 line was introduced to Mexico for use as a parent in genetic improvement, primarily aimed at enhancing the long juvenile trait in Mexican soybean populations. This trait contributes to a delayed flowering time and improved plant size under short photoperiods.

Table 3. Location and frequency of SNPs.

Number	Reference location	Gene location	SNPs	Allele frequency				Amino acid change
				A	G	C	T	
				***Lox*1 gene**				
1	42329191	Exon 1	T/A	0.2000			0.8000	Arg-Arg
2	42329248	Intron 1	A/T	0.5000			0.5000	
3	42329367	Intron 1	G/T		0.8000		0.2000	
4	42329383	Intron 1	C/T			0.8000	0.2000	
5	42329522	Intron 1	T/A	0.2000			0.8000	
6	42329549	Exon 2	C/A	0.2000		0.8000		His-Asn
7	42330262	Intron 3	C/G		0.2917	0.7083		
8	42330372	Intron 3	T/G		0.2917		0.7083	
9	42330423	Exon 4	A/T	0.7083			0.2917	Glu-Asp
10	42330876	Intron 4	C/T			0.4615	0.5385	
11	42330884	Intron 4	T/C			0.2692	0.7308	
12	42330888	Intron 4	T/G		0.2692		0.7308	
13	42331050	Intron 5	A/C	0.5000		0.5000		
14	42331083	Intron 5	G/T		0.4583		0.5417	
15	42331134	Intron 5	A/T	0.7083			0.2917	
16	42331199	Exon 6	G/C		0.7083	0.2917		Ser-Thr
17	42331273	Intron 6	A/C	0.4167		0.5833		
18	42331544	Exon 7	C/T			0.7083	0.2917	Val-Val
19	42332038	Intron 8	C/A	0.2917		0.7083		
20	42332188	Exon 9	G/C		0.7083	0.2917		Val-Leu
21	42332220	Exon 9	C/T			0.7083	0.2917	Ala-Ala
22	42332242	Exon 9	G/A	0.2917	0.7083			Ala-Thr
23	42332262	Exon 9	T/C			0.2692	0.7308	Val-Val
24	42332424	Exon 9	C/A	0.2692		0.7308		Leu-Ile
25	42332433	Exon 9	A/G	0.7308	0.2692			Ile-Val
26	42332806	Exon 9	G/A	0.0385	0.9615			Gly-Asp
				***Lox*2 gene**				
1	42322045	Intron 1	T/C			0.3571	0.6071	
2	42322161	Intron 1	G/A	0.3846	0.6154			
3	42323112	Intron 3	G/A	0.3846	0.6154			
4	42323153	Intron 3	T/C			0.3846	0.6154	
5	42323180	Intron 3	T/G		0.3846		0.6154	
6	42323648	Intron 4	C/T			0.6364	0.3636	
7	42323835	Intron 5	G/A	0.5000	0.5000			
8	42323904	Intron 5	A/G	0.8077	0.1923			
9	42324015	Exon 6	G/C		0.8462	0.1538		Glu-Asp
10	42324197	Intron 6	C/A	0.1667		0.8333		
11	42324352	Exon 7	C/T			0.9583	0.0417	Thr-Thr
12	42324587	Intron 7	T/A	0.1667			0.8333	
13	42324588	Intron 7	A/T	0.8333			0.1667	
14	42325120	Exon 9	C/A	0.0385		0.9615		Pro-Thr
15	42325121	Exon 9	C/A	0.0385		0.9615		Pro-His
16	42325167	Exon 9	T/C			0.0385	0.9615	Gly-Gly
17	42325177	Exon 9	C/T			0.9615	0.0385	Leu-Leu
				***Lox*3 gene**				
1	2125400	Intron 2	A/G	0.3846	0.6154			
2	2126115	Exon 5	T/G		0.6154		0.3846	Asp-Glu
3	2126489	Exon 6	G/A	0.6154	0.3846			Arg-Arg
4	2126715	Intron 6	T/G		0.6154		0.3846	
5	2127691	Exon 8	A/G	0.3846	0.6154			Ala-Ala

This scientific publication in digital format is a continuation of the Printed Review: Legal Deposit pp 196802ZU42, ISSN 0378-7818.

Table 4. Predictive analysis of the effect of nsSNPs on the protein in the mexican soybean population.

*Position	**nsSNP	Exon	Amino acid variant	Effect	Destabilized genotype
			***Lox*1 gene**		
42329549	C/A	2	H53N	Neutral	
42331199	G/C	6	S379T	Neutral	
42332188	G/C	9	V600L	Neutral	
42332242	G/A	9	A618T	Neutral	
42332424	C/A	9	L679I	Neutral	
42332433	A/G	9	I682V	Neutral	
42332806	G/A	9	G806D	Neutral	
			***Lox*2 gene**		
42324015	G/C	6	E393D	Destabilizing	C/C (Guayparime S-10 and Huasteca 300)
42325120	C/A	9	P669T	Neutral	
42325121	C/A	9	P669H	Neutral	
			***Lox*2 gene**		
2126115	T/G	9	D382E	Neutral	

*= Location of the SNP conforms to the reference sequence of the Williams 82 cultivar; **= non synonymous SNPs

In studies by Reinprecht *et al.* (2011) and Lee *et al.* (2014), only three non-synonymous substitutions in exons 2 and 3 were detected, resulting in amino acid changes. The predictive analysis of the eight non-synonymous SNPs detected in the *Lox*1 gene of the Mexican soybean population showed no destabilization of the L-1 protein. The PROVEAN analysis supported these findings by indicating a neutral effect of the nsSNPs on the L-1 protein (table 4). This neutral effect was observed not only in the mutant cultivars Nainari and Guayparime S-10 but also in the JP28955 and JP30790 cultivars, as well as in the overall Mexican soybean population. These results contrast with previous reports by Lenis *et al*. (2010) and Lee *et al*. (2014), who identified a 74 bp deletion in exon 8 that influenced the transcription of the L-1 protein in mutant lines. The limited genetic variation observed in the *Lox*1 gene of the Mexican soybean population can be attributed to the genome homogenization resulting from the selection and fixation of specific alleles associated with soybean breeding programs in Mexico (Ascencio and Maldonado, 1998; Rodríguez-Cota *et al*., 2017).

***Lox*3 gene**

The *Lox*3 gene in the Mexican soybean population exhibits lower genetic variation compared to the *Lox*1 and *Lox*2 genes, as observed in the reference sequence of the Williams 82 material. The limited genetic variability of the *Lox*3 gene in the Mexican soybean population is consistent with previous reports at both the intron and exon levels (Reinprecht *et al*., 2011; Lee *et al*., 2014), suggesting that the *Lox*3 gene is less variable than *Lox*1 and *Lox*2. In the *Lox*3 gene of the Mexican soybean population, a total of five SNPs were identified (table 3). Two SNPs were located in introns, while three transversion SNPs were found in exons 5, 6, and 8. Among these SNPs, the nsSNP T/G in exon 5 resulted in an amino acid change from aspartic acid to glutamic acid. However, the predictive analysis of the three non-synonymous SNPs detected in the *Lox*3 gene of the Mexican soybean population showed no destabilization of the L-3 protein. Similar changes and polymorphisms have been reported in the *Lox*3 gene, both with and without effects on the L-3 protein. For example, Lenis *et al*. (2010) detected a nucleotide change in exon 1 compared to the reference sequence of the Williams 82 cultivar. This nucleotide change altered the reading frame of the *Lox*3 gene, leading to the truncation of the functional L-3 protein in mutant soybean genotypes. Reinprecht *et al*. (2011) identified three SNPs in exons 6, 7, and 9 between a mutant and wild-type sequence, but these SNPs did not have an effect on the functioning of the L-3 protein. Lee *et al*. (2014) also reported changes in the *Lox*3 gene. They identified two non-synonymous G/A nucleotide changes in exons 6 and 9, resulting in histidine to arginine and isoleucine to valine amino acid changes, respectively, in a wild-type and mutant genotype.

In silico analysis of SNPs effect on the protein

The predictive analysis of this nsSNP G/C detected in exon 6 revealed the presence of 79 and 65 putative motifs before and after data filtering, respectively, which is one more than the reference sequence of cultivar Williams 82. However, it is worth noting that the L-2 protein affected by the nsSNP G/C at position 42324015 (E393D) did not exhibit changes in structural conformation compared to the normal protein as shown in figure 1.

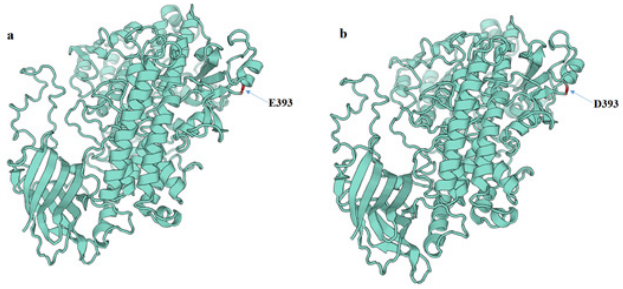

Figure 1. Model of normal L-2 protein (a) and model of the L-2 protein affected by SNP G/C at position 42324015 (b).

Conclusions

In conclusion, our study reveals significant genetic diversity within the introns and exons of *Lox* genes in the Mexican soybean population, with *Lox*3 exhibiting less variability compared to *Lox*1 and *Lox*2. Notably, nsSNPs found in the Mexican population have a neutral impact on L-1 and L-3 proteins. Moreover, our analysis suggests no destabilizing effects on lipoxygenase proteins in lipoxygenase-free JPs materials, prompting further exploration of alternative genetic regions and biochemical analyses, including enzymatic activity assessments and volatile compound quantification. The G/C E392D nsSNP in exon 6, showcasing destabilizing effects on L-2 protein in specific cultivars, holds potential as a valuable polymorphism for enhancing grain quality in Mexican soybean populations through breeding strategies.

Acknowledgments

Mónica López-Fernández thanks the IPN (BEIFI scholarship) and CONAHCYT (CONAHCYT scholarship) for providing the student scholarships that allowed me to carry out doctoral studies and publish this manuscript.

Literature cited

Ascencio, L. G., & Maldonado, M. N. (1998). Crecimiento y desarrollo de la soya en el sur de Tamaulipas, México. *Agricultura Técnica en México*, *24*(2), 99-110. https://www.acuedi.org/ddata/1274.pdf.

Balzarini, M., & Di Rienzo, J. (2004). Info-Gen: Software para análisis estadístico de datos genéticos. Universidad Nacional de Córdoba. Córdoba. Argentina. https://www.info-gen.com.ar/.

Carpentieri-Pipolo, V., Barreto, T. P., da Silva, D. A., Abdelnoor, R. V., Marin, S. R., & Degrassi, G. (2021). Soybean Improvement for Lipoxygenase-free by Simple Sequence Repeat (SSR) Markers Selection. *Journal of Botanical Research, 3*(1), 23-31. https://doi.org/10.30564/jrb.v3i1.2818.

Davies, C. S., Nielsen, S. S., & Nielsen, N. C. (1987). Flavor improvement of soybean preparations by genetic removal of lipoxygenase-2. *Journal of the American Oil Chemists' Society, 64*(10), 1428-1433. https://doi.org/10.1007/BF02636994.

Hinson, K., & Hartwig, E. E. (1977). Soybean production in the tropics. FAO. *4*, p222. https://www.cabdirect.org/cabdirect/abstract/19786724286.

Kumar, V., & Rani, A. (2019). Development of lipoxygenase-3 free Indian soybean using gene-mutation based molecular markers. *Indian Journal of Experimental Biology, 57*, 540-544. https://nopr.niscpr.res.in/handle/123456789/48474.

Lee, K. J., H. Wang, J. E., Velusamy, V., Ha, B. K., Kim, J. B., Kim, S. H., & Kim, D. S. (2014). Selection and molecular characterization of a lipoxygenase-free soybean mutant line induced by gamma irradiation. *Theoretical and applied genetics, 127*, 2405-2413. https://doi.org/10.1007/s00122-014-2385-9.

Lenis, J. M., Gillman, J. D., Lee, J. D., Shannon, J. G., & Bilyeu, K. D. (2010). Soybean seed lipoxygenase genes: molecular characterization and development of molecular marker assays. *Theoretical and Applied Genetics, 120*, 1139-1149. https://doi.org/10.1007/s00122-009-1241-9.

López-Fernández M., Castillo-Ruíz O., & Gill-Langarica H. R. (2022). First sensory analysis of soybean drinks made from commercial grain grown in Mexico. *Czech Journal of Food Sciences, 40*(5), 400-406. https://doi.org/10.17221/153/2021-CJFS.

Maldonado-Moreno, N., & Ascencio-Luciano, G. (2012). Tamesí, nueva variedad de soya para el trópico húmedo de México. *Revista mexicana de ciencias agrícolas*, *3*(8), 1671-1677). https://www.scielo.org.mx/scielo.php?pid=S2007-09342012000800017&script=sci_arttext.

Rani, A., & Kumar, V. (2020). Translational genomics and breeding in soybean. *Accelerated Plant Breeding, Volume 3: Food Legumes*, pp. 343-367. https://doi.org/10.1007/978-3-030-47306-8_11.

Reinprecht, Y., Luk-Labey, S. Y., Yu, K., Poysa, V. W., Rajcan, I., Ablett, G. R., & Peter Pauls, K. (2011). Molecular basis of seed lipoxygenase null traits in soybean line OX948. *Theoretical and applied genetics, 122,* 1247-1264. https://doi.org/10.1007/s00122-011-1528-5.

Rodríguez-Cota, F. G., Manjarrez-Sandoval, P., Cortez-Mondaca, E., Sauceda-Acosta, R. H., Valenzuela-Herrera, V., González-González, D., & Velarde-Félix, S. (2017). Guayparime S-10, nueva variedad de soya resistente a mosca blanca y geminivirus para Sinaloa. *Revista mexicana de ciencias agrícolas, 8*(1), 241-245. https://doi.org/10.29312/remexca.v8i1.88.

Shin, J. H., Van, K., Kim, K. D., Lee, Y. H., Jun, T. H., & Lee, S. H. (2012). Molecular sequence variations of the lipoxygenase-2 gene in soybean. *Theoretical and Applied Genetics, 124*, 613-622. https://doi.org/10.1007/s00122-011-1733-2.

Singh, G., Dukariya, G., & Kumar, A. (2020). Distribution, importance and diseases of soybean and common bean: A review. *Biotechnology Journal International, 24*(6), 86-98. https://doi.org/10.9734/bji/2020/v24i630125.

Suda, I., Hajika, M., Nishiba, Y., Furuta, S., & Igita, K. (1995). Simple and rapid method for the selective detection of individual lipoxygenase isoenzymes in soybean seeds. *Journal of Agricultural and Food Chemistry, 43*(3), 742-747. https://doi.org/10.1021/jf00051a034.

Wang, W. H., Takano, T., Shibata, D., Kitamura, K., & Takeda, G. (1994). Molecular basis of a null mutation in soybean lipoxygenase 2: substitution of glutamine for an iron-ligand histidine. Proceedings of the National Academy of Sciences, *91*(13), 5828-5832. https://doi.org/10.1073/pnas.91.13.5828.

This scientific publication in digital format is a continuation of the Printed Review: Legal Deposit pp 196802ZU42, ISSN 0378-7818.