International Journal of Bio-resource and Stress Management https://ojs.pphouse.org/index.php/IJBSM

Crossref

IJBSM October 2025, 16(10): 01-13 **Article** AR6444

Research Article *Natural Resource Management*

DOI: HTTPS://DOI.ORG/10.23910/1.2025.6444

In-silico Sequence Alignment, Phylogeny, Docking, and Domain Analysis of Fatty Acid Desaturase (FAD) Proteins in Pearl Millet (*Pennisetum glaucum* L.)

Thalari Vasanthrao[1,2], Mazahar Moin[1], Prabhavathi Kona[2], Wricha Tyagi[1], S. K. Gupta[1], Sudhakar Reddy P.[1] and Shashi Bhushan Danam[2]

[1]International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, Patancheru, Telangana (502 324), India
[2]Dept. of Genetics and Plant Breeding, Professor Jayshankar Telangana Agricultural University (PJTAU), Rajendranagar, Hyderabad, Telangana (500 030), India

Open Access

Corresponding Wricha.Tyagi@icrisat.org

0000-0001-9625-4870

ABSTRACT

This *in-silico* analysis was conducted during May and June, 2025 at ICRISAT, Patancheru, Telangana, India to explore the properties of pearl millet FAD (fatty acid desaturase) proteins. The protein sequences were retrieved from the NovoGene Millet database, and analyzed using ExPASy ProtParamto investigate the FAD protein family in pearl millet. Phyre2.1 was used to predict the secondary structures, and protein phylogeny was analyzed using ClustalW to identify evolutionary relationships. Conserved motifs were identified using the MEME online tool. Multiple sequence alignment of PgFAD protein sequences across six genotypes in the pangenome was performed through ClustalW to identify the polymorphism. A total of 22 PgFAD proteins were identified, and their physiochemical analysis revealed diverse molecular weights, isoelectric points (pI) and hydrophobicity profiles, suggested functional variability. Three-dimensional secondary structure analysis revealed that all PgFADs are rich in alpha helical content (45–66%), suggesting a compact, stable and membrane spanning structure. Low to moderate disordered regions (6 to 22%) in these proteins indicated their enzymatic activities. Protein polymorphism analysis showed that PgFAD-5e and PgFAD-6c are conserved across all six genotypes in the pangenome, whereas PgFAD-5c, PgFAD-5f and PgFAD-1ω6 exhibited multiple nonsynonymous mutations particularly in PI526529 genotype. These findings provided insights into the structural and functional diversity of fatty acid desaturases and established a foundation for targeting these genes to enhance lipid stability and improve quality-related traits of pearl millet.

KEYWORDS: FADs, *Pennisetum glaucum*, lipid metabolism, phylogeny

***Citation* (VANCOUVER):** Vasanthrao et al., *In-silico* Sequence Alignment, Phylogeny, Docking, and Domain Analysis of Fatty Acid Desaturase (FAD) Proteins in Pearl Millet (*Pennisetum glaucum* L.). *International Journal of Bio-resource and Stress Management,* 2025; 16(10), 01-13. HTTPS://DOI.ORG/10.23910/1.2025.6444.

Data Availability Statement: Legal restrictions are imposed on the public sharing of raw data. However, authors have full right to transfer or share the data in raw form upon request subject to either meeting the conditions of the original consents and the original research study. Further, access of data needs to meet whether the user complies with the ethical and legal obligations as data controllers to allow for secondary use of the data outside of the original study.

Conflict of interests: The authors have declared that no conflict of interest exists.

RECEIVED on 12th June 2025 RECEIVED in revised form on 25th September 2025 ACCEPTED in final form on 12th October 2025 PUBLISHED on 31st October 2025

1. INTRODUCTION

Pearl millet (*Pennisetum glaucum* L.) is a nutrient-dense and climate-resilient cereal grown on marginal soils in Asia and Africa, highlighting its importance in global food security. It is rightly designated as a "nutricereal" due to its substantial source of energy, rich in carbohydrates, proteins, fats, dietary fibre, iron, and zinc (Hassan et al., 2021, Kumar et al., 2021 Slama et al., 2020). It is an important cereal globally, along with rice, wheat, maize, barley, and sorghum (Satyavathi et al., 2021, Varshney et al., 2021). Up to 74% of its total fat content is made up of polyunsaturated fatty acids (PUFA), which are important and vital for overall health, specially for the functioning of the brain. It is also rich in omega-3 fatty acids, oleic acid (25%), linoleic acid (45%), and linolenic acid (4%). It is a gluten–free grain that is beneficial to those with gluten allergies because it maintains its alkaline qualities even after cooking. Pearl millet also has a larger proportion of slowly digested starch (SDS) and RS, which contribute to a lower glycemic index (GI) (Satyavathi et al., 2020). Despite its nutritional significance, the variety of lipids and fatty acids and their genetic determinants in pearl millet, is poorly understood. However, high lipid content in pearl millet flour is susceptible to rancidity, and it is a significant challenge that affects the shelf life and consumer acceptability (Goswami et al., 2020, Ali et al., 2022). Rancidity is driven by the hydrolysis of unsaturated fatty acids. This process is closely linked with the activity of fatty acid desaturase enzymes, which introduce double bonds into fatty acid chains and convert them into unsaturated fatty acids (Cao et al., 2024, Zhang et al., 2021).

FADs are also one of the key enzymes in the Kennedy pathway of fatty acid metabolism, where they regulate the degree of unsaturation in membrane and storage lipids (Yogendra et al., 2024). These changes not only influence the fluidity of cell membranes, structural integrity but also determine the nutritional and oxidative stability of grain lipids (Jiao and Zhang 2013, Diaz et al., 2018, Wang et al., 2019). The desaturation of fatty acids catalyzed by the FADs results in the formation of mono and polyunsaturated fatty acids (Dong and Shang, 2013, Naughton et al., 2023, Xue et al., 2018, Vasanthrao et al., 2025, Zhu et al., 2018). After the pearl millet grains are milled, the endogenous lipases and lipoxygenases are released and hence, activated which rapidly hydrolyze and oxidize these UFAs and produce free FAs, hydrogen peroxides and volatile aldehydes responsible for inducing off-flavors to the fluor, thereby reduces the shelf life (Selvan et al, 2022, Aher et al., 2022, (Sharma and Chug 2017). Thus, activity of FADs which increases PUFA contents in the grain leads to both hydrolytic and oxidative rancidity (Yogendra et al., 2024). Therefore, regulation of endogenous levels of FAD expression through genetic or gene editing techniques can effectively improve pearl millet flour stability and quality.

Despite their functional significance, the FAD gene family has not yet been comprehensively characterized in pearl millet. Hence, understanding their sequence diversity, domain architecture, phylogeny, and evolutionary relationships of FADs can provide valuable insights into their roles in lipid metabolism and storage as well as their potential link to rancidity. Recent advances in bioinformatics tools and the availability of the genome and protein sequences make in silico analysis a powerful approach to investigate their structure, function, and phylogeny of FAD family.

The present study primarily focused on conducting a genome-wide computational characterization FAD protein in pearl millet, examining their structure, phylogenetic relationships, and conserved motifs to understand the functional landscape of FADs in pearl millet.

2. MATERIALS AND METHODS

The present analysis was conducted during May and June 2025 at ICRISAT, Patancheru, India, to study in-silico sequence alignment, phylogeny, docking, and domain analysis of FAD proteins in pearl millet (*Pennisetum glaucum* L.)

2.1. In-silico analysis

To conduct in silico analysis of pearl millet PgFAD proteins, complementary DNA (cDNA) sequences and the corresponding amino acid sequences were retrieved from the NovoGene Millet genome database (http://milletdb.novogene.com/). The analysis was done using ExPASy ProtParam to assess their physico-chemical properties. Phyre2.1 server was used to predict the 3-dimensional structural modeling of FAD proteins. Conserved motifs within FAD proteins were identified using the MEME online tool. Multiple sequence alignment and phylogenetic analysis were carried out through ClustalW to identify evolutionary relationships.

2.2. Protein properties, phylogeny, secondary structure prediction and regulatory network analyses

The protein sequences of 22 PgFADs were obtained from the millet database and analyzed using ExPASy and ProtParam, which predicts protein properties such as size, molecular weight, and isoelectric point (pI). The alignment of the amino acid sequences of these proteins was accomplished through ClustalW. The overall hydrophobicity of the PgFADs was assessed by computing the GRAVY (Grand average of hydropathicity) indices using ExPASy ProtParam. To predict the three-dimensional secondary structure of the 22 PgFADs proteins, we employed the Phyre2 program (Protein Homology/AnalogY Recognition Engine v2).

2.3. Conserved motifs of PgFAD proteins

To identify the conserved motifs of PgFAD proteins, the amino acid sequences of FADs were submitted to the MEME online web server. Subsequent analysis to understand the function and characterization of these identified motifs, the sequences were subjected to the Pfam (http://pfam.xfam.org) online tool (Goud et al., 2023; Bateman et al., 2004).

2.4. Alignment of PgFAD amino acid sequences

All the amino acid sequences of 22 PgFADs were retrieved from the Novogene Millet database. In addition, the protein sequences of 22 PgFADs from six pearl millet genotypes (two landraces, two wild, and two cultivated genotypes) in the pangenome were also retrieved to identify the non-synonymous variations across the pangenome leading to changes in the amino acids or corresponding domains. All sequences were aligned through ClustalW software.

3. RESULTS AND DISCUSSION

3.1. Identification of pearl millet fatty acid desaturases

A keyword search using 'Fatty Acid Desaturase' yielded 22 FAD genes and their protein sequences were retrieved from the Novogene MilletDB. Among the 22 FADs, 13 were ACP desaturases, one was Omega 3 FAD, two were Omega 6 FADs and one Sphingolipid desaturase. These are all classified and named according to their location on the chromosomes of the respective fatty acid desaturase proteins.

3.2. Cellular localization of PgFAD proteins

All PgFAD proteins were present in the chloroplast of green tissues and plastids of non-photosynthetic tissues, integral membranes, and some proteins were also present in the cytoplasm (Almagro et al., 2017, Yu et al., 2006) (Table 3).

3.3. Protein domain identification, phylogeny, secondary structure, and regulatory network analyses of FADs

Among the 22 FADs, PgFAD-1ω6 was found as the largest, possessing a molecular mass of 51394.1 Da, followed closely by PgFAD-5c with 48164.58 Da. The other FADs exhibited molecular masses below 48000 Da. All these proteins displayed isoelectric points that ranged from 6.44 to 11.63; this variation might serve to minimize columbic repulsions, as these proteins have interactive properties. The largest protein, PgFAD-1ω6 was comprised 444 amino acids,

Table 1: Properties of PgFAD protein

Gene ID	Protein name	Protein length	Mol Wt. (Da)	pI	GRAVY
PMD1G02263.1	PgFAD-1ω6	444	51394.72	9.31	-0.121
PMD2G02512.1	PgFAD-2ω6	406	45462.14	8.85	0.006
PMD1G00698.1	PgFAD4-1a	256	26958.54	8.27	0.04
PMD1G00701.1	PgFAD4- 1b	292	31030.02	8.3	-0.091
PMD1G00700.1	PgFAD4 -1c	282	29825.88	9.21	-0.002
PMD3G05617.1	PgFAD-3a	366	41788.29	6.69	-0.507
PMD4G02572.1	PgFAD4-4a	304	31971.79	6.44	-0.028
PMD4G02574.1	PgFAD4-4b	236	24563.97	8.68	0.223
PMD4G02703.1	PgFAD-4	365	40314.98	6.74	-0.281
PMD3G01595.1	PgFAD-3b	321	37149.91	7.93	-0.07
PMD3G03626.1	PgFAD-3c	392	44576.77	6.53	-0.492
PMD5G04087.1	PgFAD-5a	410	45725	6.98	-0.356
PMD6G06386.1	PgFAD-6a	384	42804.81	6.58	-0.34
PMD6G06388.1	PgFAD-6b	199	22855.34	8.93	-0.32
PMD6G06710.1	PgFAD-6c	399	45364.79	6.44	-0.45
PMD5G03309.1	PgFAD-5b	348	39436.73	8.79	-0.428
PMD5G03528.1	PgFAD-5c	433	48164.58	8.41	-0.408
PMD5G03553.1	PgFAD-5d	426	47143.63	7.72	-0.294
PMD5G03554.1	PgFAD-5e	426	47143.63	7.72	-0.294
PMD5G01612.1	PgFAD-5ω3	427	47820.28	9.23	-0.105
PMD5G05355.1	PgFAD-5f	171	18907.29	11.63	-0.591
PMD7G02412.1	PgFAD-7	414	45570.01	7.22	-0.214

Table 1: Continue...

Gene ID	Rich amino acid (%)	Disordered region	Alpha helix	Beta strand	TM helix	Aliphatic index
PMD1G02263.1	Ala (A) 8.6%, Leu (L) 8.6%	20%	45%	0%	28%	84.59
PMD2G02512.1	Ala (A) 13.5%	13%	52%	2%	35%	89.7
PMD1G00698.1	Ala (A) 8%	9%	52%	0%	25%	82.58
PMD1G00701.1	Ala (A)18.2%	16%	49%	1%	22%	78.12
PMD1G00700.1	Ala (A) 16.7%	19%	47%	0%	30%	76.7
PMD3G05617.1	Ala (A) 11.5%	14%	66%	1%		72.87
PMD4G02572.1	Ala (A) 20.4%	16%	46%	1%	29%	77.34
PMD4G02574.1	Ala (A) 20.3%	10%	56%	1%	26%	91.99
PMD4G02703.1	Ala (A) 16.4%	12%	53%	1%		82.25
PMD3G01595.1	Leu (L) 10.6%	7%	55%	0%	36%	85.98
PMD3G03626.1	Ala (A) 9.2%	22%	54%	1%		77.17
PMD5G04087.1	Ala (A) 12.4%	20%	50%	1%		71.73
PMD6G06386.1	Ala (A) 13.3%	20%	52%	2%		74.53
PMD6G06388.1	Arg (R) 11.6%	22%	55%	3%		84.72
PMD6G06710.1	Ala (A) 8.5%	21%	49%	3%		80.2
PMD5G03309.1	Arg (R) 13.2%	12%	57%	2%		78.88
PMD5G03528.1	Ala (A) 12.5%	17%	46%	3%		76.91
PMD5G03553.1	Ala (A) 12.7%	18%	47%	2%		79.34
PMD5G03554.1	Ala (A) 12.7%	18%	47%	2%		79.34
PMD5G01612.1	Leu (L) 12.2%	17%	44%	2%	31%	90.73
PMD5G05355.1	Arg (R) 15.8%	6%	58%	0%		56.14
PMD7G02412.1	Ala (A) 14.0%	14%	53%	1%		75.53

while PgFAD-5c followed with 433 amino acids, and the shortest, PgFAD-5f contained only 171 amino acids. Nearly all FAD proteins were characterized by a high content of alanine (Ala), albeit in different proportions. The proteins PgFAD-6b, PgFAD-5b, and PgFAD-5f.1 were noted for their enrichment in arginine (Arg), whereas the remaining proteins, such as PgFAD-5ω3 and PgFAD-3b, were rich in leucine. Interestingly, all the FAD proteins were rich in amino acids, alanine and leucine. The aliphatic index of FADs lay between 50 and 90. The GRAVY indices, which served as a measure of a peptide's hydrophobicity, were less than 0. Peptides with hydropathy scores below 0 were characterized as highly hydrophilic and were likely to form flexible structures in conjunction with other molecules. The secondary structure of FADs revealed varying proportions of α, β, and distorted regions. For instance, PgFAD-3c exhibited the highest percentage of α-helical content at 66%, while PgFAD-5ω3 showed the lowest at 44%. The most significant distorted structural regions, at 22% each, were observed in PgFAD-6b and PgFAD-3c, whereas PgFAD-5f displayed the least at 6%. The predicted 3D structure of all 22 PgFADs was illustrated in Figure 1. These disordered regions played a crucial role in promoting the organization of signaling complexes through reversible protein-protein interactions, which aided in the formation of reversible cellular assemblies and contributed to the establishment of stable amyloid frameworks (Li et al., 2012, Weatheritt et al., 2014, Tripathi and Roy, 2023). In

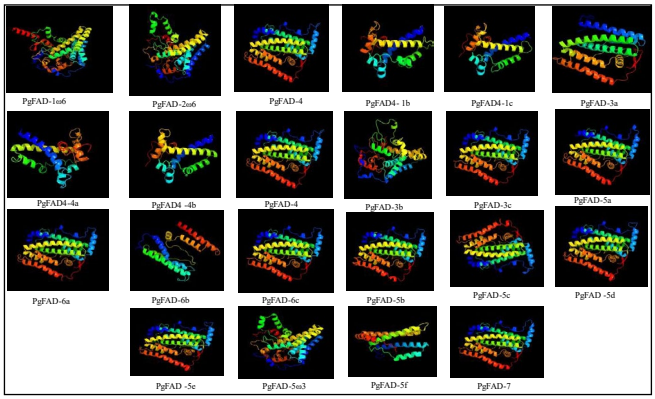

Figure 1: Predicted 3D structures of PgFAD proteins. The 3D structure model and the template of PgFADs were performed by the Phyre2 Software server (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgiid=index)

contrast, the β-sheet content varied from 0 to 3% across all proteins. The properties of FADs, such as size, pI, the percentage of α-helices and β-sheets, and GRAVY indices, were comprehensively outlined in Table 1.

3.4. Analysis of conserved motifs of PgFAD proteins

We identified 20 conserved motifs in the PgFAD protein sequences using the MEME web server. Subsequent analysis with Pfam (http://pfam.xfam.org) indicated that all conserved motifs of PgFAD were classified under the category of "FA_desaturase. Similar conserved motifs were also found in wheat and barley (Hajiahmadi et al., 2020, Bailey et al., 2006). PgFAD-3a, PgFAD-4, PgFAD-6c, and PgFAD-5c contained 10 motifs, while PgFAD-5a, PgFAD-6a, and PgFAD-5b had 9 motifs. In contrast, PgFAD4-1a, PgFAD4-1b, PgFAD4-1c, PgFAD-4a, and PgFAD-4b each contain 5 motifs, and PgFAD-1ω6, PgFAD2ω6, and PgFAD-6b presented 2, 3, and 4 motifs, respectively (Figure 2). Similar observations were also documented in *Argania spinosa* (Faqer et al., 2024).

Table 2: Analysis of non-synonymous variations/SNPs across six genotypes of PgFAD in the pangenome

Gene ID PI583800	Change in AA	Original amino acid (PI583800)	Position in change	Gene ID of the variations observed
PMD1G02263	Thiamine	Alanine	4	PMB1G02056
	Cysteine	Serine	5	PMB1G02056
	Lysin	Serine	7	PMB1G02056
	Cysteine	Leucine	9	PMB1G02056
	Alanine	Glutamine	16	PMB1G02056
	Valine	Leucine	17	PMB1G02056
	Alanine	Proline	18	PMB1G02056
	Alanine	Serine	19	PMB1G02056
	Valine	Leucine	20	PMB1G02056
	Glutamine	Arginine	21	PMB1G02056
PMD1G00698	Alanine	Valine	3	PME1G00709
	Serine	Thiamine	40	PME1G00709
	Glycine	Alanine	169	PME1G00709
	Alanine	Serine	172	PME1G00709
	Alanine	Glycine	183	PME1G00709
	Leucine	Valine	213	PME1G00709
PMD1G00701	Alanine	Leucine	21	PME1G00716
	Thiamine	Alanine	53	PME1G00716
	Arginine	Alanine	144	PME1G00716
	Glycine	Alanine	145	PME1G00716
	Arginine	Alanine	146	PME1G00716
	Alanine	Histidine	147	PME1G00716
	Arginine	Alanine	148	PME1G00716
PMD1G00700	Valine	Alanine	80	PME1G00715
	leucine	Valine	109	PMH1G00560
	Glycine	Alanine	156	PMH1G00560
	Arginine	Cysteine	178	PME1G00715
	Arginine	Cysteine	189	PMB1G00695
	Arginine	Cysteine	280	PMC1G00216
PMD2G02512	Arginine	Leucine	19	PMB3G07300

Table 2: Continue...

Gene ID PI583800	Change in AA	Original amino acid (PI583800)	Position in change	Gene ID of the variations observed
PMD3G05617	No variations observed			
PMD3G01595	No variations observed			
PMD3G03626	No variations observed			
PMD4G02703	Lysine	Aspartic acid	10	PMH4G01948
	Glutamic acid	Aspartic acid	11	PMH4G01948
	Arginine	Proline	45	PMA4G02756
	Lysine	Arginine	224	PMD4G02703
	Lysine	Arginine	224	PMC4G02514
PMD4G02572	No variations observed			
PMD4G02574	No variations observed			
PMD5G04087	Threonine	Serine	27	PMH5G03240
	Threonine	Serine	27	PMA5G03986
	Arginine	Histidine	55	PMB5G04247
	Arginine	Histidine	55	PME5G04083
	Arginine	Histidine	55	PMC5G04048
	Threonine	alanine	63	PMB5G04247
	Threonine	alanine	63	PMC5G04048
	Threonine	Histidine	386	PMC5G04048
	Asparagine	Aspartic acid	404	PME5G04083
PMD5G03309	Glycine	Alanine	292	PMA5G03255
	Arginine	lysine	323	PMA5G03255
	Alanine	Proline	364	PMA5G03255
PMD5G03528	Arginine	Glutamine	70	PMB5G03663
	Arginine	Glutamine	70	PMH5G02816
	Arginine	Glutamine	70	PMC5G03507
	Arginine	Histidine	237	PMB5G03663
	Arginine	Histidine	237	PMH5G02816
	Arginine	Histidine	237	PMC5G03507
	Glycine	Asparagine	421	PMC5G03507
	Leucine	Tryptophan	422	PMC5G03507
	Arginine	Isoleucine	423	PMC5G03507
	Aspartic acid	Arginine	426	PMC5G03507
	Isoleucine	Threonine	427	PMC5G03507
	Alanine	Isoleucine	428	PMC5G03507
	Arginine	Asparagine	429	PMC5G03507
	Cysteine	Valine	430	PMC5G03507
	Histidine	Arginine	431	PMC5G03507
	Serine	Valine	432	PMC5G03507
	Arginine	Proline	433	PMC5G03507

Table 2: Continue...

Gene ID PI583800	Change in AA	Original amino acid (PI583800)	Position in change	Gene ID of the variations observed
PMD5G03553	methionine	Valine	293	PMB5G03705
	methionine	Valine	293	PME5G03563
	methionine	Valine	293	PMA5G03499
PMD5G03554	No variations observed			
PMD5G05355	Serine	Tryptophan	12	PME5G05307
	Glycine	aspartic acid	38	PME5G05307
	Glycine	aspartic acid	38	PMA5G05237
	phenylalanine	Arginine	142	PME5G05307
	Threonine	Leucine	143	PME5G05307
	Threonine	Proline	144	PME5G05307
	Valine	Tryptophan	145	PME5G05307
	Alanine	Arginine	146	PME5G05307
	Serine	Proline	147	PME5G05307
	phenylalanine	Arginine	142	PMA5G05237
	Threonine	Leucine	143	PMA5G05237
	Threonine	Proline	144	PMA5G05237
	Valine	Tryptophan	145	PMA5G05237
	Alanine	Arginine	146	PMA5G05237
	Serine	Proline	147	PMA5G05237
PMD5G01612	Alanine	threonine	427	PMA5G01548
PMD6G06386	Phenylalanine	Serine	175	PMB6G07037
	glutamine	Lysine	371	PME6G06572
PMD6G06388	Alanine	threonine	50	PMA6G06351
	Alanine	Glycine	194	PMA6G06351
	Leucine	Arginine	195	PMA6G06351
	Serine	Glutamic acid	196	PMA6G06351
	Serine	Valine	197	PMA6G06351
	Arginine	Valine	198	PMA6G06351
	Tryptophan	Valine	199	PMA6G06351
PMD6G06710	No variations observed			
PMD7G02412	Alanine	Threonine	66	PMA7G02371
	Alanine	Threonine	66	PME7G02393
	Alanine	Threonine	66	PMC7G01966
	Aspartic acid	Glycine	236	PMB7G02602
	Serine	Alanine	289	PMB7G02602
	Glycine	Alanine	389	PMB7G02602

3.5. Multiple sequence alignment of PgFAD amino acid sequences across pangenome

The alignment of amino acid sequences from six pearl millet genotypes across the pangenome was conducted using ClustalW software, indicated the presence of mutations in the sequences for the majority of the PgFAD

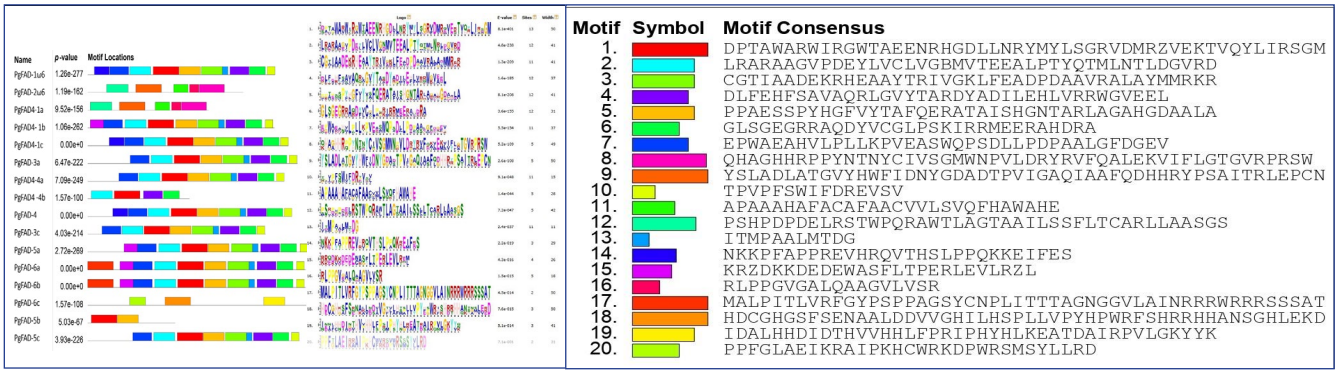

Figure 2: Identification of conserved motifs in PgFAD proteins. Twenty conserved motifs are presented in different colors. The identified motifs were detected by the MEME online tool

Table 3: Cellular localization of PgFAD proteins

Protein	Localization
PMD1G00698	Integral membrane protein
PMD1G02263	Chloroplast, Nuclear and Vacuolar
PMD2G02512	Integral membrane protein
PMD1G00701	Chloroplast thylakoid. Mitochondrial inner membrane
PMD1G00700	Mitochondrial inner membrane
PMD3G05617	Chloroplast of green tissue and plastids of no photosynthetic tissues
PMD4G02572	Chloroplast
PMD4G02574	Integral membrane protein; plasma membrane
PMD4G02703	Cytoplasmic
PMD3G01595	Cytoplasmic
PMD3G03626	Chloroplast of green tissue and plastids of no photosynthetic tissues
PMD5G04087	Mitochondrial
PMD6G06386	Chloroplast
PMD6G06388	Cytoplasmic
PMD6G06710	Chloroplast
PMD5G03309	Chloroplast
PMD5G03528	Chloroplast
PMD5G03553	Chloroplast
PMD5G03554	Chloroplast
PMD5G01612	Integral membrane protein; plasma membrane
PMD5G05355	Glycosomal
PMD7G02412	Chloroplast

genes. For instance, the amino acid sequence of gene ID PMD1G00698 from the PI583800 genotype, which served as a reference, was aligned with the amino acid sequences of five other genotypes (PI587025, PI527388, PI537069, PI521612, and PI526529). These results revealed that non-synonymous variations occur at various positions within the aligned sequences. For example, in the PME1G00709 gene of the PI526529 genotype, alanine, serine, glycine, alanine, and leucine were found in place of valine, thiamine, alanine, serine, glycine, and valine, respectively (Fig. 3). The changes in amino acids or SNPs for the other genotypes were detailed in Table 2. No mutations were observed in some FADs (Figures 3 and 4 and Table 2). Similar findings regarding the FAD gene family had also been reported in Brassica juncea (Dar et al., 2017).

3.6. Phylogenetic analysis of 22 PgFAD proteins revealed six distinct groups/clusters, reflecting their evolutionary divergence and potential functional differentiation

The phylogenetic analysis identified six distinct clusters among the pearl millet fatty acid desaturase protein family, highlighted their evolutionary divergence and possible functional differentiation. Cluster I consisted of PgFAD4-1a, PgFAD4-1c, PgFAD4- 1b and PgFAD4-4a with a bootstrap value of 1000 replicates, indicated a close evolutionary relation probably due to their recent gene duplication events. Similarly, Cluster II contained PgFAD4-4b with conserved function. Cluster III consisted of a broad array of proteins, such as PgFAD-3b, PgFAD-1ω6, PgFAD-2ω6, PgFAD-5ω3, PgFAD-5a, and PgFAD-7. This cluster likely represented a functionally diverse subfamily within the larger pearl millet FAD protein family, and it was characterized by an internal branching pattern at high 1000 bootstrap replicates indicated multiple rounds of duplication events and divergence within this protein family. Additionally, two more clusters, Cluster IV and V, encompassed PgFAD-5b, PgFAD-6a, PgFAD-6b and PgFAD-3a, PgFAD-3c, PgFAD-6c, respectively. Both clusters signified functionally distinct lineages that might have evolved independently. The cluster VI included

Figure 3: Alignment of amino acid sequences of PgFADs across six pearl millet genotypes. Non-synonymous SNPs are highlighted with red colour at 3, 40,169, 172, 183 and 213 amino acid positions. The Gene IDs PMD1G00698, PME1G00709, PMH1G00564, PMA1G00949, PMB1G00561 and PMC1G00213 of PI583800, PI587025, PI527388, PI537069, PI521612 and PI526529 genotypes respectively

```
CLUSTAL 2.1 multiple sequence alignment
PMB3G01505      MGAGMEADEGVMATDFFWSYTDEPHASRRREILAKYPQIKELFGPDPWAFLKIAVVVSLQ
PMD3G01595      MAAGMEADEGVMATDFFWSYTDEPHASRRREILAKYPQIKELFGPDPWAFLKIAVVVSLQ
PME3G01608      MGAGMEADEGVMATDFFWSYTDEPHASRRREILAKYPQIKELFGPDPWAFLKIAVVVSLQ
PMH3G01214      MGAGMEADEGVMATDFFWSYTDEPHASRRREILAKYPQIKELFGPDPWAFLKIAVVVSLQ
PMA3G01561      MGAGMEADEGVMATDFFWSYTDEPHASRRREILAKYPQIKELFGPDPWAFLKIAVVVSLQ
PMC3G01575      MGAGMEADEGVMATDFFWSYTDEPHASRRREILAKYPQIKELFGPDPWAFLKIAVVVSLQ
                ************************************************************

PMB3G01505      LWTSTFLRDASWMKLLIVAYFFGSFLNHNLFLAIHELSHNLAFATPSLNRWLGIFANLPI
PMD3G01595      LWTSTFLRDASWMKLLIVAYFFGSFLNHNLFLAIHELSHNLAFATPSLNRWLGIFANLPI
PME3G01608      LWTSTFLRDASWMKLLIVAYFFGSFLNHNLFLAIHELSHNLAFATPSLNRWLGIFANLPI
PMH3G01214      LWTSTFLRDASWMKLLIVAYFFGSFLNHNLFLAIHELSHNLAFATPSLNRWLGIFANLPI
PMA3G01561      LWTSTFLRDASWMKLLIVAYFFGSFLNHNLFLAIHELSHNLAFATPSLNRWLGIFANLPI
PMC3G01575      LWTSTFLRDASWMKLLIVAYFFGSFLNHNLFLAIHELSHNLAFATPSLNRWLGIFANLPI
                ************************************************************

PMB3G01505      GVPMSITFQKYHLEHHRFQGVDGIDMDVPSQAEAHAVKNAVSKSVWVVLQLFFYALRPLF
PMD3G01595      GVPMSITFQKYHLEHHRFQGVDGIDMDVPSQAEAHAVKNAVSKSVWVVLQLFFYALRPLF
PME3G01608      GVPMSITFQKYHLEHHRFQGVDGIDMDVPSQAEAHAVKNAVSKSVWVVLQLFFYALRPLF
PMH3G01214      GVPMSITFQKYHLEHHRFQGVDGIDMDVPSQAEAHAVKNAVSKSVWVVLQLFFYALRPLF
PMA3G01561      GVPMSITFQKYHLEHHRFQGVDGIDMDVPSQAEAHAVKNAVSKSVWVVLQLFFYALRPLF
PMC3G01575      GVPMSITFQKYHLEHHRFQGVDGIDMDVPSQAEAHAVKNAVSKSVWVVLQLFFYALRPLF
                ************************************************************

PMB3G01505      LKPKPPGLWEFTNLAIQVALDAGLVYLYGWKSLAYLILSTFVGGGMHPMAGHFISEHYVF
PMD3G01595      LKPKPPGLWEFTNLAIQVALDAGLVYLYGWKSLAYLILSTFVGGGMHPMAGHFISEHYVF
PME3G01608      LKPKPPGLWEFTNLAIQVALDAGLVYLYGWKSLAYLILSTFVGGGMHPMAGHFISEHYVF
PMH3G01214      LKPKPPGLWEFTNLAIQVALDAGLVYLYGWKSLAYLILSTFVGGGMHPMAGHFISEHYVF
PMA3G01561      LKPKPPGLWEFTNLAIQVALDAGLVYLYGWKSLAYLILSTFVGGGMHPMAGHFISEHYVF
PMC3G01575      LKPKPPGLWEFTNLAIQVALDAGLVYLYGWKSLAYLILSTFVGGGMHPMAGHFISEHYVF
                ************************************************************

PMB3G01505      SPDQETYSYYGPLNLMTWHVGYHNEHHDFPRIPGTRLHKVKEMAPEYYDSLRSYRSWSQV
PMD3G01595      SPDQETYSYYGPLNLMTWHVGYHNEHHDFPRIPGTRLHKVKEMAPEYYDSLRSYRSWSQV
PME3G01608      SPDQETYSYYGPLNLMTWHVGYHNEHHDFPRIPGTRLHKVKEMAPEYYDSLRSYRSWSQV
PMH3G01214      SPDQETYSYYGPLNLMTWHVGYHNEHHDFPRIPGTRLHKVKEMAPEYYDSLRSYRSWSQV
PMA3G01561      SPDQETYSYYGPLNLMTWHVGYHNEHHDFPRIPGTRLHKVKEMAPEYYDSLRSYRSWSQV
PMC3G01575      SPDQETYSYYGPLNLMTWHVGYHNEHHDFPRIPGTRLHKVKEMAPEYYDSLRSYRSWSQV
                ************************************************************

PMB3G01505      IYMYIMDQTVGPFSRMKRKAPKKDS
PMD3G01595      IYMYIMDQTVGPFSRMKRKAPKKDS
PME3G01608      IYMYIMDQTVGPFSRMKRKAPKKDS
PMH3G01214      IYMYIMDQTVGPFSRMKRKAPKKDS
PMA3G01561      IYMYIMDQTVGPFSRMKRKAPKKDS
PMC3G01575      IYMYIMDQTVGPFSRMKRKAPKKDS
                *************************
```

Figure 4: Representation showing no variations/SNPs across six genotypes of PgFAD amino acid sequences in (I583800, PI587025, PI527388, PI537069, PI521612 and PI526529)

proteins such as PgFAD-4, PgFAD-5c, PgFAD-5f, PgFAD-5d, and PgFAD-5e (Figure 5). Similar results aligned with (Hajiahmadi et al., 2020). Although these proteins diverged earlier in the evolutionary timeline, their close grouping suggested that they might preserve ancestral functions (Li et al., 2025, Faqer et al., 2024). In summary, the tree topology and bootstrap values indicated that the pearl millet fatty acid desaturase protein family had experienced significant expansion through extensive duplication events and divergence, resulting in the formation of various well-supported subfamilies. This evolutionary trajectory pointed to functional specialization and warranted further research through gene expansion studies and functional characterization.

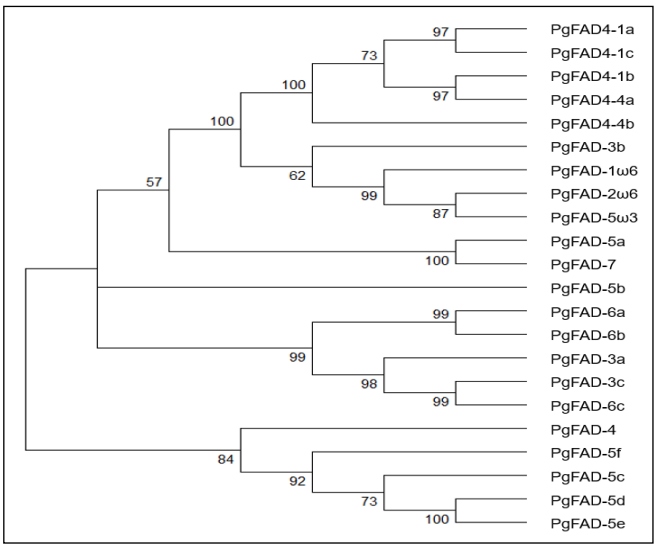

Figure 5: Phylogeny and analysis of fatty acid desaturase (PgFAD) proteins

4. CONCLUSION

The physicochemical properties of FAD proteins in pearl millet suggested diverse roles in lipid metabolism. Conserved motif analysis indicated strong functional conservation, while phylogenetic analysis revealed divergence among PgFAD families, suggested specialized functions. This divergence likely contributed to differences in lipid unsaturation.. Secondary structure analysis showed that PgFADs had a compact, membrane-bound confirmation, whereas protein polymorphism indicated that some PgFADs are conserved while others exhibited nonsynonymous mutations in the pangenome, which might contribute to their functional divergence.

5. DISCLAIMER ARTIFICIAL INTELLIGENCE

Author(s) hereby declares that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

6. ACKNOWLEDGEMENT

Thalari Vasanthrao acknowledges the UGC-CSIR, Govt. of India for providing fellowship. The authors also express their gratitude to the PJTAU, Hyderabad, ICRISAT, Patancheru, Hyderabad. All the authors have read and approved the manuscript.

7. REFERENCES

Aher, R.R., Reddy, P.S., Bhunia, R.K., Flyckt, K.S., Shankhapal, A.R., Ojha, R., Bhatnagar-Mathur, P., 2022. Loss-of-function of triacylglycerol lipases are associated with low flour rancidity in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. Frontiers in Plant Science 13, 962667.

Ali, A., Kumar, R.R., Vinutha, T., Singh, T., Singh, S.P., Satyavathi, C.T., Praveen, S., Goswami, S., 2022. Grain phenolics: Critical role in quality, storage stability and effects of processing in major grain crops-A concise review. European Food Research and Technology 248(8), 2197–2213. https://doi.org/10.1007/s00217-022-04026-7.

Almagro, Armenteros, J.J., Sønderby, C.K., Sonderby, S.K., Nielsen, H., Ole, W., 2017. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics 33(21), 3387–3395. DOI: 10.1093/bioinformatics/btx431.

Bailey, T.L., Williams, N., Misleh, C.L.I., 2006. MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Resources 34(suppl 2), W369–W73.

Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E.L., Studholme, D.J., Yeats, C., Eddy, S.R., 2004. The Pfam protein families database. Nucleic Acids Research 32, D138 D41.

Cao, T., Du, Q., Ge, R., Li, R., 2024. Genome-wide identification and characterization of FAD family genes in barley. Peer Journal 212, e16812.

Dar, A.A., Choudhury, A.R., Kancharla, P.K., Arumugam, N., 2017. The FAD2 gene in plants: occurrence, regulation, and role. Frontiers in plant science 8, 1789.

Diaz, M.L., Cuppari S., Soresi D., Carrera, A., 2018. In silico analysis of fatty acid desaturase genes and proteins in grasses. Applied Biochemistry Biotechnology 184(2), 484–499.

Dong, C.J., Shang, Q.M., 2013. Genome-wide characterization of phenylalanine ammonia-lyase

gene family in watermelon (*Citrullus lanatus*). Planta 238(1), 35–49.

El Faqer, A., Rabeh, K., Alami, M., Filali-Maltouf, A., Belkadi, B., 2024. Silico identification and characterization of fatty acid desaturase (FAD) genes in *Argania spinosa* L. skeels: Implications for oil quality and abiotic stress. Bioinformatics and Biology Insights 18, 11779322241248908.

Goswami, S., Asrani, P., Ali, A., Kumar, R.D., Vinutha, T., Veda, K., Kumari, S., Sachdev, A., Singh, S.P., Satyavathi, C.T, Kumar, R.R., Praveen, S., 2020. Rancidity matrix: Development of biochemical indicators for analysing the keeping quality of pearl millet flour. Food Analytical Methods 13(11), 2147–2164. https://doi. org/10.1007/s12161-020-01831-2.

Goud, C.A., Bhargava, K., Sadhana, P., Dattatray, B.N., Shivani, D., Nagamani, G., 2023. Structural and functional characterization of rice yield enhancing genes by in silico approach. International Journal of Bio-resource and Stress Management 14(3), 498–505.

Hajiahmadi, Z., Abedi, A., Wei, H., Sun, W., Ruan, H., Zhuge, Q., Movahedi, A., 2020. Identification, evolution, expression, and docking studies of fatty acid desaturase genes in wheat (*Triticum aestivum* L.). BMC genomics 21, 1–20.

Hassan, Z.M., Sebola, N.A., Mabelebele, M., 2021. The nutritional use of millet grain for food and feed: a review. Agriculture and Food Security 10, 1–14.

Jiao, J., Zhang, Y., 2013. Transgenic biosynthesis of polyunsaturated fatty acids: a sustainable biochemical engineering approach for making essential fatty acids in plants and animals. Chemical Reviews 113(5), 3799–3814. doi:10.1021/cr300007p.

Kumar, M., Singh, U., Raiger, P.R., Singh, L.N., Parewa, H.P., Chouhan, R., 2021. Micronutrient fortification of pearl millet (*Pennisetum glaucum* (L.) R. Br.) hybrids using customized fertilizer formulation. International Journal of Bio-resource and Stress Management 12(5), 416–425.

Li, J., McQuade, T., Siemer, A.B., Napetschnig, J., Moriwaki, K., Hsiao, Y.S., Damko, E., Moquin D., Walz, T., McDermott, A., Chan, F.K., 2012. The RIP1/RIP3 necrosome forms a functional amyloid signaling complex required for programmed necrosis. Cell 150(2), 339–50.

Li, X., Munir, M., Zeng, W., Sun, Z., Chang, X., Yang., W., 2025. Characterization of fatty acid desaturase gene family in *Glycine max* and their expression patterns in seeds after Fusarium fujikuroi infection. Frontiers in Plant Science 16, 1540003.

McNaughton, A.D., Shanklin, J., Raugei, S., Kumar, N., 2023. Data-Driven Modeling and Analysis of Fatty Acid Desaturase in Plants. BioRxiv, 2023-08.

Satyavathi, C.T., Khandelwal, V., Supriya, A., Beniwal, B.R., Sushila, B., Mahesh, C.K., 2020. Pearl millet-hybrids and varieties-2020. Jodhpur: ICAR-All India Coordinated Research Project on Pearl Millet.

Satyavathi, C.T., Ambawat, S., Khandelwal, V., Srivastava, R.K., 2021. Pearl millet: a climate-resilient nutricereal for mitigating hidden hunger and provide nutritional security. Frontiers in Plant Sciences 12, 659938.

Selvan, S.S., Mohapatra, D., Anakkallan, S., Kate, A., Tripathi, M.K., Singh, K., Kar, A., 2022. Oxidation kinetics and ANN model for shelf-life estimation of pearl millet (*Pennisetum glaucum* L.) grains during storage. Journal of Food Processing and Preservation 46(12), e17218.

Slama, A., Cherif, A., Sakouhi, F., Boukhchina, S., Radhouane, L., 2020. Fatty acids, phytochemical composition, and antioxidant potential of pearl millet oil. Journal of Consumer Protection and Food Safety 15, 145–151.

Sharma, B., Chugh, L.K., 2017. Two isoforms of lipoxygenase from mature grains of pearl millet [*Pennisetum glaucum* (L.) R. Br.]: Purification and physico-chemico-kinetic characterization. Journal of Food Science and Technology 54(6), 1577–1584.

Tripathi, A., Roy, N., 2023. Molecular Modelling and Structure analysis of lectins from *Vigna linearis*. International Journal of Bio-resource and Stress Management 14(9), 1295–1304.

Varshney, R.K, Shi, C., Thudi, M., Mariac, C., Wallace, J., Qi, P., Xu, X., 2017. Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. Nature Biotechnology 35(10), 969–976.

Vasanthrao, T., Moin, M., Kona, P., Tyagi, W., Gupta, S.K., Reddy, P.S., Danam, S.B., 2025. Genome-wide identification and in silico analysis of fatty acid desaturase (FAD) genes of pearl millet (*Pennisetum glaucum* L.). Journal of Advances in Biology and Biotechnology 28(6), 1646–1659.

Wang, X., Yu, C., Liu Y., Yang, L., Li, Y., Yao, W., Peng, X., 2019. GmFAD3A, aω-3fatty acid desaturase gene, enhances cold tolerance and seed germination rate under low temperature in rice. International Journal of Molecular Sciences 20(15), 3796.

Weatheritt, R.J., Gibson, T.J., Babu, M.M., 2014. Asymmetric mRNA localization contributes to fidelity and sensitivity of spatially localized systems. Nature Structural Molecular Biology 21(9), 833–839.

Xue, Y., Chen, B.J., Wang, R., Win, A.N., Li, J.N., Chai, Y.R., 2018. Genome-wide survey and characterization of fatty acid desaturase gene family in Brassica napus and its parental species. Applied Biochemistry

Biotechnology 184, 582–598. Doi: 10.1007/s12010-017-2563–8.

Yogendra, K., Sanivarapu, H., Avuthu, T., Gupta, S.K., Durgalla, P., Banerjee, R., Tyagi, W., 2024. Comparative metabolomics to unravel the biochemical mechanism associated with rancidity in pearl millet (*Pennisetum glaucum* L.). International Journal of Molecular Sciences 25(21), 11583.

Yu, C.S., Chen, Y.C., Lu, C.H., Hwang, J.K.J.P.S., 2006. Function, bioinformatics. Prediction of protein subcellular localization. Proteins 64(3), 643–51.

Zhang, B., Xia, P., Yu, H., Li, W., Chai, W., Liang, Z., 2021. Based on the whole genome clarified the evolution and expression process of fatty acid desaturase genes in three soybeans. International Journal of Biological Macromolecules 182, 1966–1980.

Zhu, S., Zhu, Z., Wang, H., Wang, L., Cheng, L., Yuan, Y., Zheng, Y., Li, D., 2018. Characterization and functional analysis of a plastidial FAD6 gene and its promoter in the mesocarp of oil palm (*Elaeis guineensis*). Scientia Horticulturae 239, 163–170.