Scientific Electronic Archives
Issue ID: Sci. Elec. Arch. Vol. 18 (3)
May/June 2025
DOI: http://dx.doi.org/10.36560/18220252068
Article link: https://sea.ufr.edu.br/SEA/article/view/2068

ISSN 2316-9281

Genome-Wide identification and characterization caleosin genes in lima-bean (*Phaseolus lunatus*)

Corresponding author
Maria Cidinaria Silva Alves
Universidade Federal de Pernambuco
cidinaria.alves@ufpe.br

Ramon da Silva de Souza
Universidade Federal de Pernambuco

Caio Gabriel Tavares Ferreira
Universidade Federal de Pernambuco

Luana de Souza Leandro
Universidade Federal de Pernambuco

Dalton Ferreira Matos
Universidade Federal do Rio Grande do Sul

Ruana Carolina Cabral da Silva
Faculdade Anhanguera Currais Novos

Abstract. *Phaseolus lunatus*, commonly known as lima bean or butter bean, is a leguminous crop with significant agricultural and nutritional value, particularly in tropical regions. Caleosin, a lipid-associated protein, plays a crucial role in seed germination, stress response, and lipid metabolism, presenting potential targets for genetic improvement. This study aimed to identify and characterize caleosin genes (*PlCLOs*) in *P. lunatus* using *in silico* methods. Six genes encoding caleosin proteins were identified, exhibiting molecular weights between 16.58 and 27.28 kDa and subcellular localization predominantly in chloroplasts. Conserved motifs, such as calcium-binding and phosphorylation sites, were identified, alongside structural elements crucial for lipid droplet anchoring. Phylogenetic analysis revealed three evolutionary groups, suggesting functional divergence. Structural modeling confirmed high-quality protein models dominated by α-helices and irregular loops. Functional annotations highlighted roles in stress tolerance, calcium signaling, and lipid metabolism. These findings deepen the understanding of caleosins role in plant biology, providing insights for sustainable agricultural practices and genetic improvement of *P. lunatus.*
Keywords: Caleosin; Genomic analysis, Lima-Bean; Protein modeling; Stress tolerance.

Introduction

Phaseolus lunatus, also known as lima bean or butter bean, is a leguminous plant from the Fabaceae family, originating in Central and South America. It ranks as the second most important legume species for human nutrition and agricultural production in tropical regions, surpassed only by the common bean (Bitocchi et al. 2017). Given its relevance, numerous omics-based investigations have recently been conducted to explore various aspects of its biology and cultivation (Garcia et al. 2021; Wisser et al. 2021; Heredia-Pech et al. 2022).

The lima bean, or butter bean, plays an important role in food security and sustainable agriculture, especially in tropical regions, where it stands out due to its adaptability to diverse environmental conditions. Its high nutritional value, including a significant protein content, makes it an essential component of the human diet, particularly in areas with limited access to animal-based protein sources. *P. lunatus*

contributes to soil enrichment through nitrogen fixation, promoting more sustainable agricultural practices and improving crop productivity (Palupi et al. 2021; Adebo, 2023).

In recent years, advancements in molecular biology and genomics have provided a deeper understanding of the genetic makeup of *P. lunatus*, enabling researchers to identify key genes involved in important traits such as drought resistance, pest tolerance, and seed quality. These discoveries have paved the way for the development of improved varieties through selective breeding and genetic engineering, with the goal of increasing both productivity and the crop's resilience to environmental stresses (Garcia et al. 2021; Silva Alves et al. 2024)

Caleosin is a protein associated with oleosomes, structures responsible for lipid storage in plants, playing a role in lipid metabolism, especially during seed germination, when stored lipids are mobilized to provide energy. Its function in reserve metabolism, caleosin, is also involved in stress response processes, such as drought and salinity conditions, contributing to the plant's adaptation to adverse environments (Hanano et al. 2023; Zhu et al. 2023; Feng et al. 2024).

Recent studies suggest that caleosin may act in cell signaling and seed development regulation, making it an interesting target for research related to genetic improvement and agricultural biotechnology. The identification of caleosin genes, in particular, represents a promising area of research due to their role in lipid metabolism, seed development, and stress response (Brunetti et al. 2021; Hanano et al. 2023).

Understanding the genetic basis of these traits not only benefits agricultural productivity but also provides crucial information for conserving genetic diversity within the species. By identifying and studying specific genes, such as caleosin genes, scientists can develop more sustainable cultivation methods and ensure the long-term viability of this crop. Thus, the objective of this research was to identify and characterize caleosin genes in *P. lunatus* through *in silico* approaches, aiming to contribute to the understanding of their role in plant metabolism and processes related to development and stress response.

Material and methods

Identification and physicochemical properties of the lima bean Caleosin gene

The genomic data for the lima bean was obtained from Phytozome (Genome ID: 563, Garcia et al. 2021). To identify caleosin genes in *P. lunatus*, a BlastP search was performed using published protein sequences of caleosin genes from various species as probes. Candidate domains were confirmed using CD-search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi, Wang et al. 2023).

To predict the physicochemical properties of the candidate proteins, the online prediction tool ProParam from the Expasy database was used (https://web.expasy.org/protparam/). Subcellular localization was predicted using the WoLF PSORT online tool (https://wolfpsort.hgc.jp/). Multiple sequence alignment was carried out using the Clustal Omega tool (Sievers; Higgins, 2014).

Conserved motifs, gene structure, chromosomal location and promoters of PlCLOs

The MEME Suite 5.5.3 tool (https://meme-suite.org/meme/doc/meme.html) was used to analyze the conserved motifs in the amino acid sequences of the candidate caleosin proteins, while gene structure was analyzed using the Gene Structure Display Server 2.0 (http://gsds.cbi.pku.edu.cn/). Promoter regions of *PlCLO* (2000 bp upstream of ATG) were uploaded to the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) for cis-regulatory element analysis. The results were visualized using TBtools (v1.120). Chromosomal localization was performed using the MG2C software (http://mg2c.iask.in/mg2c_v2.1/index.html).

Phylogenetic analysis

Caleosin protein sequences from various species were obtained from the UniProt database (https://www.uniprot.org/). The phylogenetic tree was constructed using the Neighbor-Joining (NJ) method in MEGA 7.0 software with 1000 bootstrap replicates (Kumar et al. 2016) . The tree was visualized using the online tool iTOL (https://itol.embl.de/; Letunic and Bork, 2016).

Structural analysis of lima-bean Caleosin proteins

AlphaFold (Varadi et al. 2022) was used to select the homolog with the highest identity (Culver et al. 2024). Subsequently, models were generated using the MODELLER tool (Webb; Sali, 2014) and selected based on the Discrete Optimized Protein Energy (DOPE) score metric (Shen, Min-yi; Sali, 2006).

To evaluate stereochemical conflicts in the models, PDBsum (Laskowski et al. 2018) was utilized through the PROCHECK tool. PROSA-Web was employed to calculate the model folding (Z-score) (Wiederstein; Sippl, 2007), and QMEANDisCo was used to assess the overall quality of the models (Studer et al. 2020).

The physicochemical characterization of biomolecular solvent-solute interactions was conducted using the APBS-PDB2PQR software suite (http://server.poisonboltzmann.org). The PDB2PQR algorithm facilitated model preparation, while pKa values were determined using the PROPKA algorithm to define protonation states at pH 7, employing the Charmm force field (Brooks et al. 2009). Solvation energy was estimated through the Poisson-Boltzmann equation, which provides a simplified representation of polar solvation energies. Structural analysis of CLO models, focusing on

conserved sites, was performed by comparing 3D models using PyMOL (The PyMOL Molecular Graphics System, Version 2.0, Schrödinger, LLC).

Results and discussion

After the genomic analysis, six genes belonging to the caleosin domain were identified. These genes exhibited significant variations in their sizes, ranging from 16.58 to 27.28 kDa. The isoelectric points of the identified genes varied between 5.98 and 8.65. Regarding subcellular localization, the analysis revealed that the majority of the identified caleosin proteins were predominantly located in the chloroplasts (3 genes), while two other genes were found in the cytoplasm, and one gene was located in the nucleus (Table 1).

Generally, caleosins are primarily found in the cytoplasm, where they associate with lipid droplets and play an essential role in lipid metabolism and energy storage (Liu et al. 2022). Caleosins can also be found in chloroplasts, contributing to photosynthetic processes and stress responses (Saadat, 2023). The study by Wasąg et al. (2024), observed caleosins localized in the nucleus as well, particularly calreticulin, similar to our findings, suggesting roles in gene regulation and calcium signaling (Saadat, 2023). Furthermore, caleosins interact with autophagy-related proteins to aid in the degradation of lipid droplets during seed germination (Miklaszewska et al. 2023).

Considering that caleosins are ubiquitously found across various plant species, such as *Arabidopsis thaliana* (Shen et al. 2014), *Phoenix dactylifera L.* (Rahman et al. 2018), *Carthamus tinctorius L.* (Lu et al. 2018), *Nicotiana benthamiana* (Brunetti et al. 2022), among others, according to the study by Cao et al. (2024), gene sequences of *Arabidopsis* caleosins were used to perform a local BLAST and characterize genes in three distinct species. Four caleosin genes were identified in *Carya cathayensis*, three in *Carya illinoinensis*, and three in *Juglans regia*, with molecular weights ranging from 16.91 to 27.08 kDa and isoelectric points (pI) varying between 5.70 and 9.40.

Although *caleosins* are present in all plant genomes characterized so far, their isoforms vary among different plant groups (Rahman et al. 2018). These isoforms are classified based on their molecular weight, being either the H-isoform, which has a high molecular weight, or the L-isoform, which has a low molecular weight (Saadat, 2023). In this context, all the caleosins in our study exhibited low molecular weight and were thus classified as L-type isoforms.

The alignment of the six gene sequences revealed several conserved regions, suggesting the presence of structural and functional elements that are crucial for the activity of caleosin proteins. The analysis highlighted variations in specific areas, with an emphasis on residues that exhibited 100% conservation, which may influence the physicochemical and functional properties of these proteins (Figure 1).

Table 1. Prediction of physicochemical properties of caleosin proteins of *P. lunatus*

Gene	Gene-ID	Length (aa)	MW (Kda)	pI	Gravy	Subcellular Localization
ICLO1	PI06G0000219800.1.v1	240	27.28	7.10	-0.278	Cytoplasm
PICLO2	PI07G0000221600.1.v1	240	27.03	5.98	-0.329	Cytoplasm
PICLO3	PI03G0000022900.1.v1	201	23.02	8.65	-0.654	Chloroplast
PICLO4	PI03G0000023200.1.v1	205	22.84	8,.65	-0.507	Chloroplast
PICLO5	PI03G0000023100.1.v1	199	22.41	8.64	-0.555	Chloroplast
PICLO6	PI07G0000221600.2.v1	148	16.58	6.58	-0.388	Nucleus

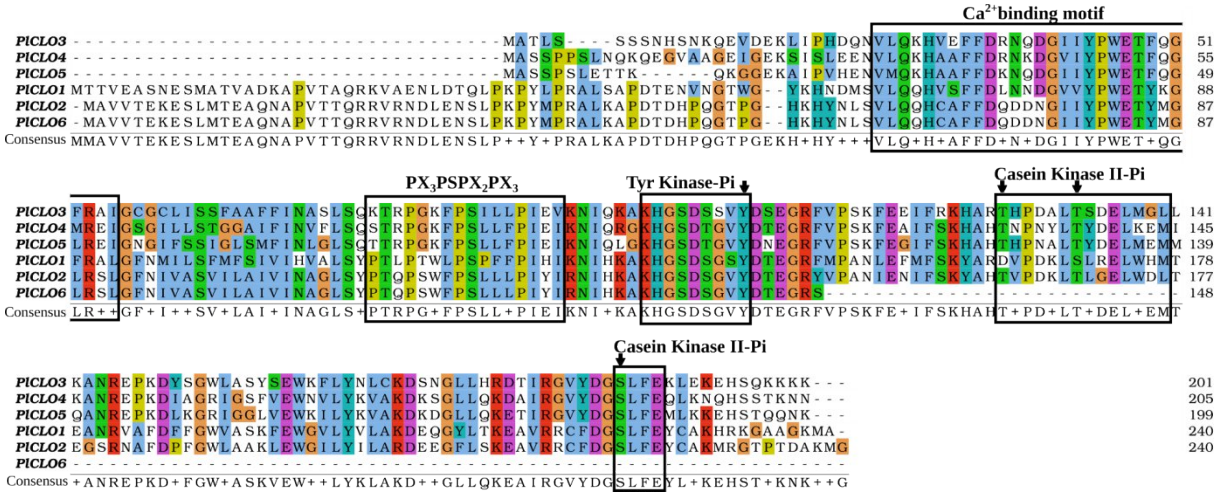

Figure. 1 Alignment of caleosin gene sequences, showing their conserved regions.

As in other studies, a calcium-binding motif and a proline knot, represented as PX3PSPX2PX3, were observed, along with a tyrosine kinase phosphorylation site and two casein kinase II phosphorylation sites. These motifs have already been reported in species such as *Carya illinoinensis*, *Juglans regia*, and *Arabidopsis thaliana* (Song et al. 2014; Cao et al. 2024). It should be noted that differences in amino acid residues among distinct species may contribute to the existing isoforms within this group, which are highly conserved and perform various biological functions. For instance, the PX3PSPX2PX3 motif is considered crucial for anchoring lipid droplets (LDs). Caleosins interact directly with LDs, generating lipid metabolites through peroxygenase activity, which subsequently influence other signaling compounds in plant cells (Kim et al. 2011; Chapman et al. 2012).

Exon and intron patterns in caleosins were analyzed, revealing a high degree of similarity among them and suggesting structural and functional differences between plant groups, potentially impacting gene structure and the patterns of conserved motifs (Figure 2) (Saadat, 2023; Cao et al. 2024). Previous studies have indicated that the number of introns and exons in caleosins has not varied significantly, remaining between 4 and 6 introns and 2 to 6 exons, consistent with our findings (Song et al. 2014; Shen et al. 2016).

Regarding the analysis of conserved motifs, results vary across studies. In Cao et al. (2024), the researchers identified only three conserved motifs: the first being a highly conserved proline, the second a calcium-binding site, and the third an NADPH-binding domain. Conversely, Hanano et al. (2018), reported that the predicted sequences exhibited high levels of identity or, in some cases, conservative substitutions.

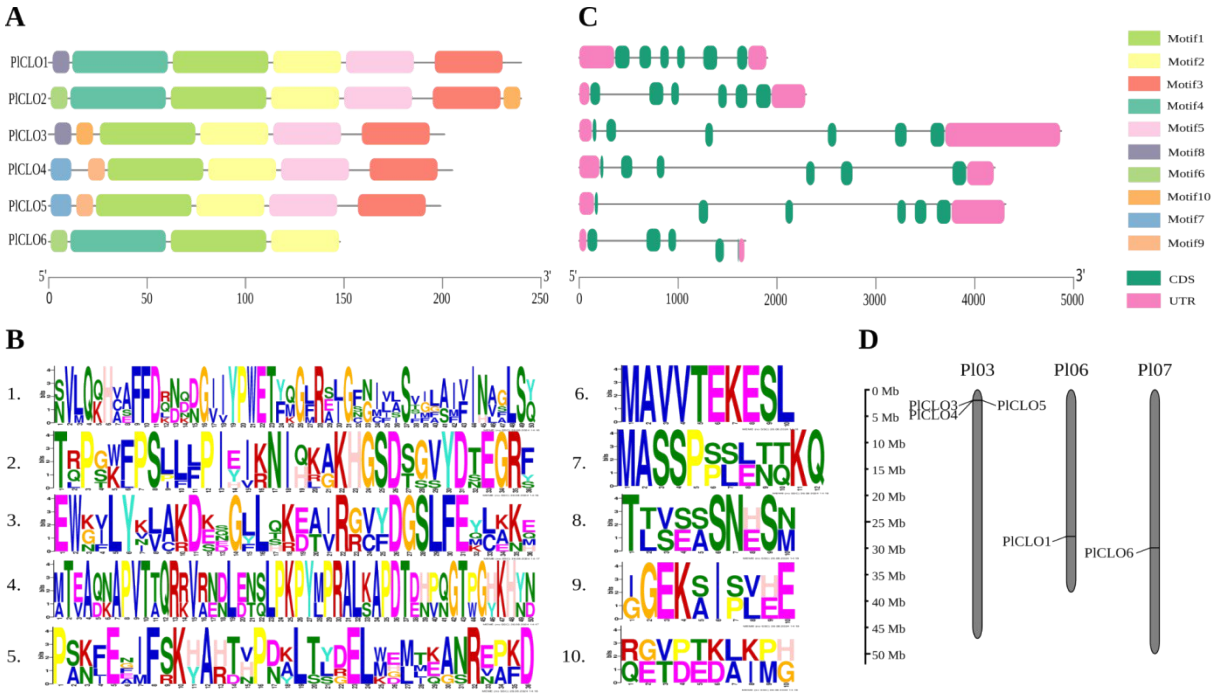

Figure. 2 A. Representation of Conserved Domains. **B**. Most frequent conserved motifs. **C**. Gene structure. **D**. Chromosomal location of CLOs of *P. lunatus*.

Several studies address cis-regulatory elements in caleosins, some of which are associated with stress responses. For example, the study by Shen et al. (2016), demonstrated that the promoter sequences of *BnCLO1-2* and *BnCLO1-3* support expression patterns that confer stress resistance in *Arabidopsis*.

In *Carya cathayensis*, a total of 13 cis-regulatory elements were identified, related to stress responses, hormonal signaling, and tissue-specific expression. Additionally, these elements are associated with responses to light, anaerobic induction, MYB-binding sites induced by drought, and components of defense and stress responses. In *Juglans regia*, these cis elements are involved in responses to plant hormones such as abscisic acid, jasmonate, gibberellin, salicylic acid, among others (Cao et al. 2024). Similarly, in *Phaseolus lunatus*, these cis-regulatory elements are also linked to hormonal responses, playing roles in defense, endosperm expression, and other adaptive responses (Figure 3).

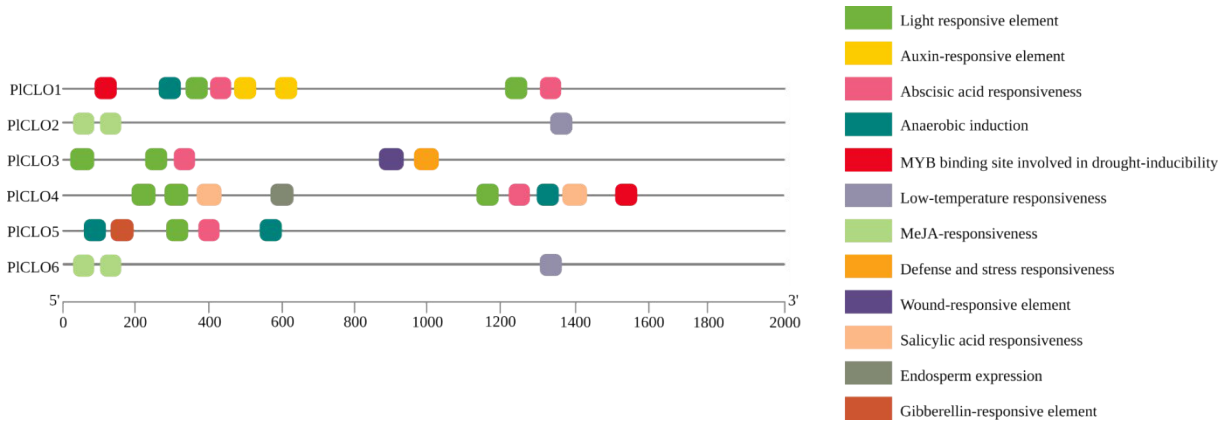

Figure. 3 Analysis of cis-regulatory elements in promoters of caleosin genes of *P. lunatus.*

Cis-regulatory element analysis in promoters of *caleosin* genes in *P. lunatus* can provide insights into the gene expression regulation, tissue-specific roles, and responses to biotic or abiotic stresses. However, within this gene family, a strong conservation of key motifs is observed, including the canonical EF-hand domains, which are associated with calcium binding (GO:0005509) as cofactors or regulators. Additionally, heme binding (GO:0020037), common in proteins involved in electron transport and oxidative reactions, and lipid binding (GO:0010888, GO:0012511, GO:0031408, GO:0034389) are prominent. Peroxidase activity (GO:0004601) and plant seed peroxygenase activity (GO:0009707) are linked to the role of eliminating reactive oxygen species, particularly in seeds (Figure 4).

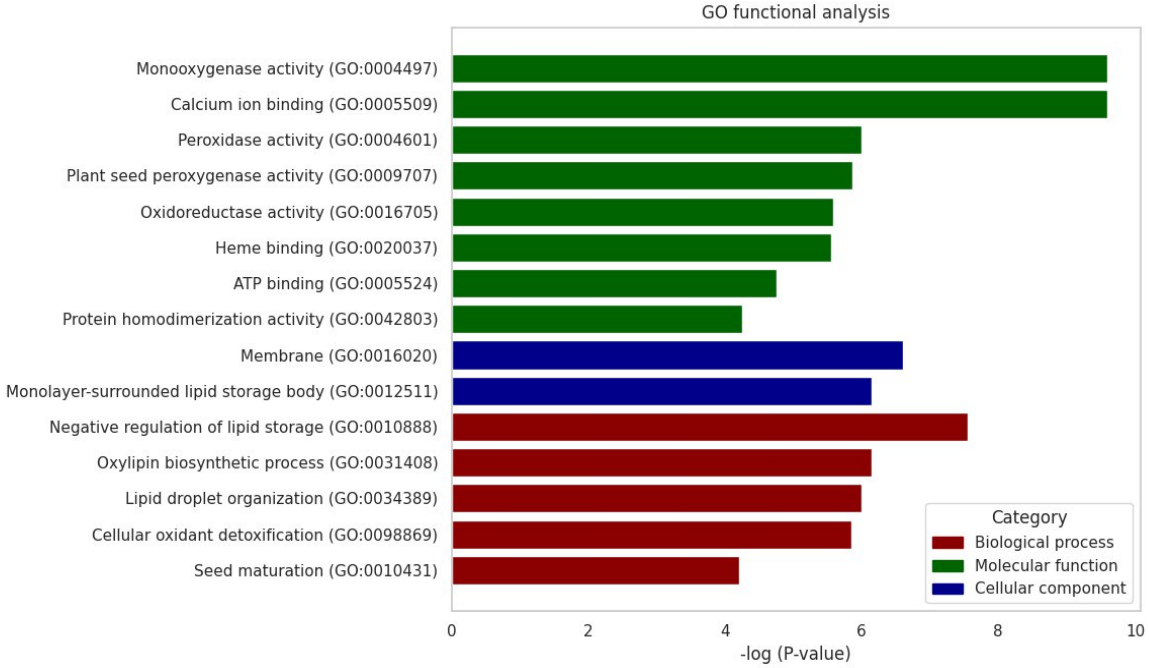

Figure. 4 Functional Annotation of GO Terms: Biological, Molecular and Cellular Categories of Genes of Interest in *PlCLOs.*

Regarding the phylogenetic analysis, three distinct groups were identified. The caleosin genes were associated with two of these groups, represented in green and pink. This separation suggests a divergence in the genetic and functional characteristics among the groups, indicating possible evolutionary adaptations and specializations in the functions of caleosin proteins (Figure 5).

Several studies have been conducted to analyze the phylogenetic relationships of caleosins, revealing significant variations in the classification of this protein family. According to Song et al. (2014), caleosins were divided into five conserved subfamilies. However, our analysis identified three main subfamilies. A more recent study by Saadat et al. (2023), expanded the classification to eight distinct families, with most sequences concentrated in the *Poaceae* and *Brassicaceae* families.

These evolutionary divergences among groups may be associated with the different botanical families analyzed. To date, it is known that these proteins have undergone specific adaptive diversification for functions in different plant species, such as response to environmental stresses or in the regulation of physiological processes (Cao *et al.*, 2024; Garcia *et al.*, 2021). Thus, further studies are needed to elucidate the specific evolutionary relationships among caleosins.

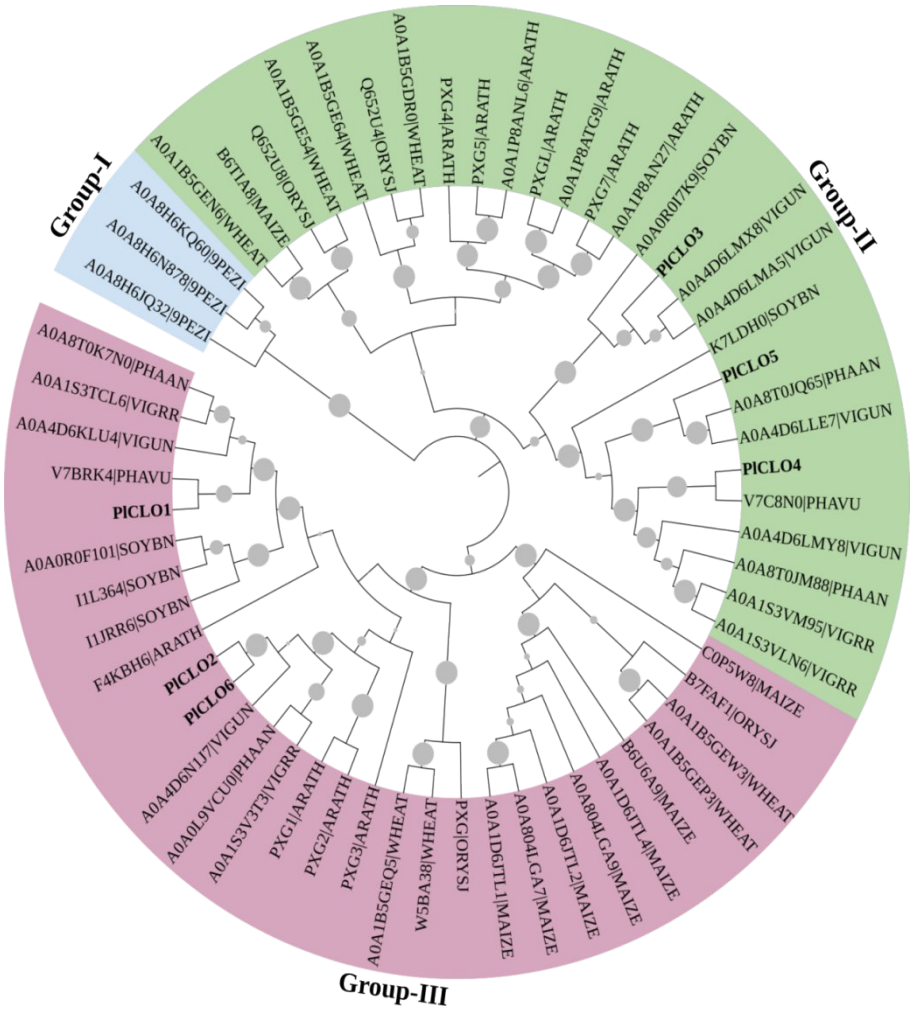

Figure. 5 Dendrogram generated from the protein sequences using the Neighbor-Joining method, demonstrating the distribution of caleosin in the lima bean genome and comparing it with different plant species. The values at the base of the nodes refer to bootstrap (1,000 replicates) and p-distance. In blue = group I; green = group II; pink = group III.

To analyze the differences among PICLO proteins, five proteins were modeled. The results of this analysis, considering the theoretical models of PICLO, modeling algorithms, and validation, are presented in Table 2. The percentage of residues in the Ramachandran plot, QMEANDisCo scores, and Z-score values for each model were calculated. All theoretical models showed at least 90% of amino acids in the most favorable regions, 0.60 ± 0.06 for QMEANDisCo, and -7.72 for the Z-score. These results highlight variations in the structural and statistical qualities of the theoretical models of PICLO proteins, which are important for interpreting their conformations and potential functional applications.

Table 2 – Results of the algorithms for modeling and validating the theoretical models of PICLOs.

TLP	Template	Organism	Identity	Ramachandran	QMEANDisCo	Z-score
PICLO1	V7BRK4	*Phaseolus vulgaris*	96%	92.12%	0.66 ± 0.06	-6.33
PICLO2	V7BE66	*Phaseolus vulgaris*	95%	93.12%	0.67 ± 0.06	-6.85
PICLO3	V7C517	*Phaseolus vulgaris*	90%	90.66%	0.71 ± 0.06	-7.65
PICLO4	V7C8N0	*Phaseolus vulgaris*	94%	91.02%	0.87 ± 0.06	-6.59
PICLO5	A0A0S3T1B6	*Vigna angularis*	89%	95.03%	0.60 ± 0.06	-7.72

The results of three-dimensional modeling indicate that lima bean caleosin proteins are composed of four structural elements: extended chains, irregular clusters, α-helices, and β-turns. The models show a predominance of α-helices, with some regions of disordered coiled loops and no β-sheets observed. The α-helices and irregular coils account for the highest percentages in PlCLOs, at 47.99% and 45.25%, respectively. Additionally, it was noted that PlCLO3 and PlCLO4 have the highest percentages of extended strands (Figure 6). Similar results were reported in the study by Cao et al. (2024).

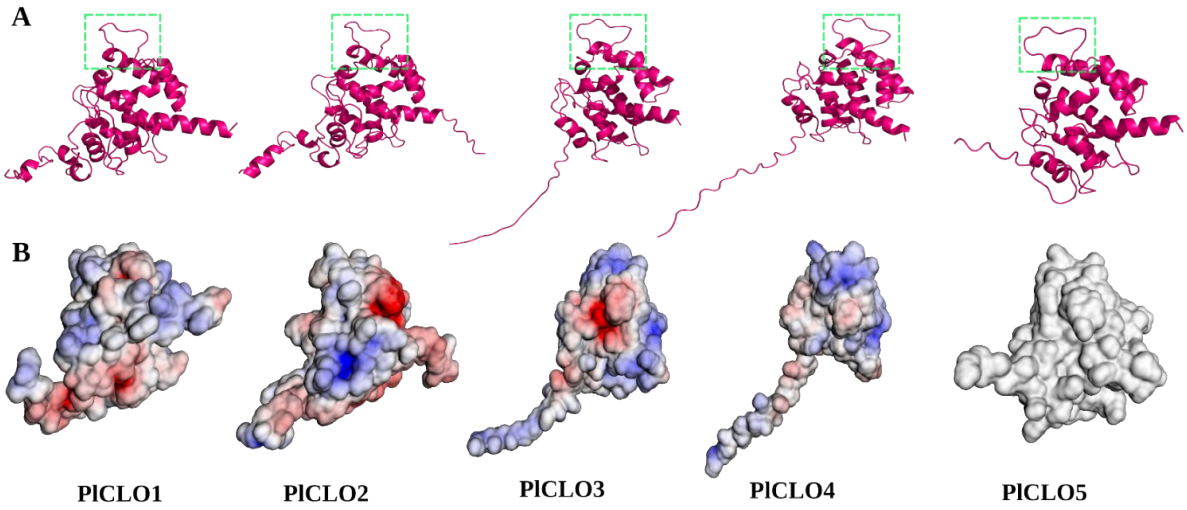

Figure. 6 Three-dimensional models of PlCLOs. A. The magenta structures were generated highlighting the proline node in green. B. The electrostatic surface potential for theoretical three-dimensional models of PlCLOs, where the red color indicates acidic charge, and the blue color indicates basic charge.

The results of three-dimensional modeling indicate that lima bean caleosin proteins are composed of four structural elements: extended chains, irregular clusters, α-helices, and β-turns. The models show a predominance of α-helices, with some regions of disordered coiled loops and no β-sheets observed. The α-helices and irregular coils account for the highest percentages in PlCLOs, at 47.99% and 45.25%, respectively. Additionally, it was noted that PlCLO3 and PlCLO4 have the highest percentages of extended strands (Figure 6). Similar results were reported in the study by Cao et al. (2024).

The three-dimensional structure and electrostatic potential of caleosins are significant for their role in lipid storage and degradation. Caleosin features a central hydrophobic domain and a unique proline knot motif, which are essential for its anchoring in oil bodies, similar to oleosin (Jiang and Tzen, 2010). Additionally, caleosins contain a conserved EF-hand calcium-binding domain, which is crucial for interaction with other proteins, such as the autophagy-related protein ATG8 (Song et al. 2014; Miklaszewska et al. 2023).

Our study highlights different electrostatic potentials, with PlCLO5 displaying a completely neutral surface, while the others showed acidic, basic, and neutral regions. The electrostatic potential of caleosin, determined by its amino acid composition, particularly in the calcium-binding region, is crucial for its functional interactions (Song et al. 2014). Changes in this potential can influence the protein's affinity for lipid droplets and its role in lipid mobilization during seed germination (Miklaszewska et al. 2023). Although essential for lipid storage and degradation, the evolutionary relationship between caleosin and oleosin suggests a divergence that may impact their functions in different plant species (Jiang; Tzen, 2010).

However, although caleosin is essential for lipid storage and degradation, its evolutionary relationship with oleosin suggests a divergence that may affect its functional roles in different plant species (Jiang and Tzen, 2010). Therefore, further studies across various plant groups are needed to elucidate the evolution and functional mechanisms of caleosins.

Conclusion

This study provided a comprehensive characterization of the caleosin genes in *Phaseolus lunatus,* identifying six genes with significant variations in size, isoelectric point, and subcellular localization. Sequence analysis revealed conserved motifs essential for caleosin function, including calcium-binding and phosphorylation sites, which are critical for stress response and lipid metabolism. Phylogeny indicated divergence among the caleosin genes of *P. lunatus*, suggesting specific evolutionary adaptations for different biological functions. Three-dimensional modeling of the proteins confirmed the presence of predominant alpha-helix and irregular loop structures, essential for their functionality. These results not

only expand the understanding of the role of caleosins in *P. lunatus*, but also provide valuable directions for future research in genetic improvement and agricultural biotechnology. The identification and characterization of these proteins may contribute to the development of lima bean varieties that are more resilient to environmental stresses, promoting more sustainable agricultural practices, and improving food security in tropical regions.

Acknowledgments

We thank CAPES (Coordination of Higher Education Personnel, Brazil) and CNPq (National Council for Scientific and Technological Development, Brazil) for fellowships and financial support. We also thank ChatGPT for assistance with the translation of this article.

References

ADEBO, J. A. A Review on the Potential Food Application of Lima Beans (Phaseolus lunatus L.), an Underutilized Crop. Applied Sciences, [*s. l.*], v. 13, n. 3, p. 1996, jan. 2023. https://doi.org/10.3390/app13031996.

BITOCCHI, E.; RAU, D.; BELLUCCI, E.; RODRIGUEZ, M.; MURGIA, M. L.; GIOIA, T.; SANTO, D.; NANNI, L.; ATTENE, G.; PAPA, R. Beans (Phaseolus ssp.) as a Model for Understanding Crop Evolution. Frontiers in Plant Science, [*s. l.*], v. 8, 8 maio 2017. DOI 10.3389/fpls.2017.00722. Disponível em: https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2017.00722/full. Acesso em: 22 ago. 2024.

BROOKS, B. R.; BROOKS III, C. L.; MACKERELL JR., A. D.; NILSSON, L.; PETRELLA, R. J.; ROUX, B.; WON, Y.; ARCHONTIS, G.; BARTELS, C.; BORESCH, S.; CAFLISCH, A.; CAVES, L.; CUI, Q.; DINNER, A. R.; FEIG, M.; FISCHER, S.; GAO, J.; HODOSCEK, M.; IM, W.; KUCZERA, K.; LAZARIDIS, T.; MA, J.; OVCHINNIKOV, V.; PACI, E.; PASTOR, R. W.; POST, C. B.; PU, J. Z.; SCHAEFER, M.; TIDOR, B.; VENABLE, R. M.; WOODCOCK, H. L.; WU, X.; YANG, W.; YORK, D. M.; KARPLUS, M. CHARMM: The biomolecular simulation program. Journal of Computational Chemistry, [*s. l.*], v. 30, n. 10, p. 1545–1614, 2009. https://doi.org/10.1002/jcc.21287.

BRUNETTI, S. C.; ARSENEAULT, M. K. M.; GULICK, P. J. The caleosin CLO7 and its role in the heterotrimeric G-protein signalling network. Journal of Plant Physiology, [*s. l.*], v. 279, p. 153841, dez. 2022. https://doi.org/10.1016/j.jplph.2022.153841.

BRUNETTI, S. C.; ARSENEAULT, M. K. M.; WRIGHT, J. A.; WANG, Z.; EHDAEIVAND, M.-R.; LOWDEN, M. J.; RIVOAL, J.; KHALIL, H. B.; GARG, G.; GULICK, P. J. The stress induced caleosin, RD20/CLO3, acts as a negative regulator of GPA1 in Arabidopsis. Plant Molecular Biology, [*s. l.*], v. 107, n. 3, p. 159–175, out. 2021. https://doi.org/10.1007/s11103-021-01189-x.

CAO, Y.; ZHAN, Y.; LIU, J.; TANG, T.; LI, J.; ZHAO, R.; ZHANG, Q.; HU, S.; CAO, W.; GAO, Y. Identification and Analysis of Expression Patterns of the Caleosin Genes in Hickory (Carya cathayensis Sarg.). Forests, [*s. l.*], v. 15, n. 4, p. 609, abr. 2024. https://doi.org/10.3390/f15040609.

CHAPMAN, K. D.; DYER, J. M.; MULLEN, R. T. Biogenesis and functions of lipid droplets in plants: Thematic Review Series: Lipid Droplet Synthesis and Metabolism: from Yeast to Man. Journal of Lipid Research, [*s. l.*], v. 53, n. 2, p. 215–226, fev. 2012. https://doi.org/10.1194/jlr.R021436.

CULVER, K. D.; SADECKI, P. W.; JACKSON, J. K.; BROWN, Z. A.; HNILICA, M. E.; WU, J.; SHAW, L. N.; WOMMACK, A. J.; HICKS, L. M. Identification and Characterization of CC-AMP1-like and CC-AMP2-like Peptides in Capsicum spp. Journal of Proteome Research, [*s. l.*], v. 23, n. 8, p. 2948–2960, 2 ago. 2024. https://doi.org/10.1021/acs.jproteome.3c00597.

FENG, P.; HAI-XIA, Y. U.; KUN, Z.; YING-YING, L. I. U.; GUI-YU, T. a. N. Review on the Regulation of Caleosin on Plant Lipid Droplet. Biotechnology Bulletin, [*s. l.*], v. 40, n. 4, p. 33, 26 abr. 2024. https://doi.org/10.13560/j.cnki.biotech.bull.1985.2023-1051.

GARCIA, T.; DUITAMA, J.; ZULLO, S. S.; GIL, J.; ARIANI, A.; DOHLE, S.; PALKOVIC, A.; SKEEN, P.; BERMUDEZ-SANTANA, C. I.; DEBOUCK, D. G.; MARTÍNEZ-CASTILLO, J.; GEPTS, P.; CHACÓN-SÁNCHEZ, M. I. Comprehensive genomic resources related to domestication and crop improvement traits in Lima bean. Nature Communications, [*s. l.*], v. 12, n. 1, p. 702, 29 jan. 2021. https://doi.org/10.1038/s41467-021-20921-1.

HANANO, A.; BLÉE, E.; MURPHY, D. J. Caleosin/peroxygenases: multifunctional proteins in plants. Annals of Botany, [*s. l.*], v. 131, n. 3, p. 387–409, 16 fev. 2023. https://doi.org/10.1093/aob/mcad001.

HEREDIA-PECH, M.; CHÁVEZ-PESQUEIRA, M.; ORTIZ-GARCÍA, M. M.; ANDUEZA-NOH, R. H.; CHACÓN-SÁNCHEZ, M. I.; MARTÍNEZ-CASTILLO, J. Consequences of introgression and gene flow on the genetic structure and diversity of Lima bean (Phaseolus lunatus L.) in its Mesoamerican diversity area. PeerJ, [*s. l.*], v. 10, p. e13690, 2022. https://doi.org/10.7717/peerj.13690.

JIANG, P.-L.; TZEN, J. T. C. Caleosin serves as the major structural protein as efficient as

oleosin on the surface of seed oil bodies. Plant Signaling & Behavior, [*s. l.*], v. 5, n. 4, p. 447–449, abr. 2010. https://doi.org/10.4161/psb.5.4.10874.

KIM, Y. Y.; JUNG, K. W.; YOO, K. S.; JEUNG, J. U.; SHIN, J. S. A stress-responsive caleosin-like protein, AtCLO4, acts as a negative regulator of ABA responses in Arabidopsis. Plant & Cell Physiology, [*s. l.*], v. 52, n. 5, p. 874–884, maio 2011. https://doi.org/10.1093/pcp/pcr039.

KUMAR, S.; STECHER, G.; TAMURA, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular Biology and Evolution, [*s. l.*], v. 33, n. 7, p. 1870–1874, 1 jul. 2016. https://doi.org/10.1093/molbev/msw054.

LASKOWSKI, R. A.; JABŁOŃSKA, J.; PRAVDA, L.; VAŘEKOVÁ, R. S.; THORNTON, J. M. PDBsum: Structural summaries of PDB entries. Protein Science, [*s. l.*], v. 27, n. 1, p. 129–134, 2018. https://doi.org/10.1002/pro.3289.

LETUNIC, I.; BORK, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Research, [*s. l.*], v. 44, n. W1, p. W242–W245, 8 jul. 2016. https://doi.org/10.1093/nar/gkw290.

LIU, X.; YANG, Z.; WANG, Y.; SHEN, Y.; JIA, Q.; ZHAO, C.; ZHANG, M. Multiple caleosins have overlapping functions in oil accumulation and embryo development. Journal of Experimental Botany, [*s. l.*], v. 73, n. 12, p. 3946–3962, 24 jun. 2022. https://doi.org/10.1093/jxb/erac153.

LU, Y.; CHI, M.; LI, L.; LI, H.; NOMAN, M.; YANG, Y.; JI, K.; LAN, X.; QIANG, W.; DU, L.; LI, H.; YANG, J. Genome-Wide Identification, Expression Profiling, and Functional Validation of Oleosin Gene Family in Carthamus tinctorius L. Frontiers in Plant Science, [*s. l.*], v. 9, p. 1393, 2018. https://doi.org/10.3389/fpls.2018.01393.

MIKLASZEWSKA, M.; ZIENKIEWICZ, K.; KLUGIER-BOROWSKA, E.; RYGIELSKI, M.; FEUSSNER, I.; ZIENKIEWICZ, A. CALEOSIN 1 interaction with AUTOPHAGY-RELATED PROTEIN 8 facilitates lipid droplet microautophagy in seedlings. Plant Physiology, [*s. l.*], v. 193, n. 4, p. 2361–2380, 1 dez. 2023. https://doi.org/10.1093/plphys/kiad471.

PALUPI, H. T.; ESTIASIH, T.; YUNIANTA; SUTRISNO, A. Characterization of nutritional and functional properties of Lima bean flour (Phaseolus Lunatus L.). IOP Conference Series: Earth and Environmental Science, [*s. l.*], v. 924, n. 1, p. 012033, nov. 2021. https://doi.org/10.1088/1755-1315/924/1/012033.

RAHMAN, F.; HASSAN, M.; ROSLI, R.; ALMOUSALLY, I.; HANANO, A.; MURPHY, D. J. Evolutionary and genomic analysis of the caleosin/peroxygenase (CLO/PXG) gene/protein families in the Viridiplantae. PLOS ONE, [*s. l.*], v. 13, n. 5, p. e0196669, 17 maio 2018. https://doi.org/10.1371/journal.pone.0196669.

SAADAT, F. A computational study on the structure–function relationships of plant caleosins. Scientific Reports, [*s. l.*], v. 13, n. 1, p. 72, 2 jan. 2023. https://doi.org/10.1038/s41598-022-26936-y.

SHEN, M.; SALI, A. Statistical potential for assessment and prediction of protein structures. Protein Science, [*s. l.*], v. 15, n. 11, p. 2507–2524, 2006. https://doi.org/10.1110/ps.062416606.

SHEN, Y.; LIU, M.; WANG, L.; LI, Z.; TAYLOR, D. C.; LI, Z.; ZHANG, M. Identification, duplication, evolution and expression analyses of caleosins in Brassica plants and Arabidopsis subspecies. Molecular genetics and genomics: MGG, [*s. l.*], v. 291, n. 2, p. 971–988, abr. 2016. https://doi.org/10.1007/s00438-015-1156-x.

SHEN, Y.; XIE, J.; LIU, R.-D.; NI, X.-F.; WANG, X.-H.; LI, Z.-X.; ZHANG, M. Genomic analysis and expression investigation of caleosin gene family in Arabidopsis. Biochemical and Biophysical Research Communications, [*s. l.*], v. 448, n. 4, p. 365–371, 13 jun. 2014. https://doi.org/10.1016/j.bbrc.2014.04.115.

SIEVERS, F.; HIGGINS, D. G. Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences. *In*: RUSSELL, D. J. (org.). Multiple Sequence Alignment Methods. Totowa, NJ: Humana Press, 2014. p. 105–116. DOI 10.1007/978-1-62703-646-7_6. Disponível em: https://doi.org/10.1007/978-1-62703-646-7_6. Acesso em: 25 jun. 2024.

SILVA ALVES, M. C.; CABRAL DA SILVA, R. C.; QUARESMA, S. Characterization of α-hairpinin in the genome of lima bean (Phaseolus lunatus). International Journal of Molecular Biology Open Access, [*s. l.*], v. 7, n. 1, p. 112–116, 28 ago. 2024. https://doi.org/10.15406/ijmboa.2024.07.00179.

SONG, W.; QIN, Y.; ZHU, Y.; YIN, G.; WU, N.; LI, Y.; HU, Y. Delineation of plant caleosin residues critical for functional divergence, positive selection and coevolution. BMC Evolutionary Biology, [*s. l.*], v. 14, n. 1, p. 124, 9 jun. 2014. https://doi.org/10.1186/1471-2148-14-124.

STUDER, G.; REMPFER, C.; WATERHOUSE, A. M.; GUMIENNY, R.; HAAS, J.; SCHWEDE, T. QMEANDisCo—distance constraints applied on model quality estimation. Bioinformatics, [*s. l.*], v.

36, n. 6, p. 1765–1771, 15 mar. 2020. https://doi.org/10.1093/bioinformatics/btz828.

VARADI, M.; ANYANGO, S.; DESHPANDE, M.; NAIR, S.; NATASSIA, C.; YORDANOVA, G.; YUAN, D.; STROE, O.; WOOD, G.; LAYDON, A.; ŽÍDEK, A.; GREEN, T.; TUNYASUVUNAKOOL, K.; PETERSEN, S.; JUMPER, J.; CLANCY, E.; GREEN, R.; VORA, A.; LUTFI, M.; FIGURNOV, M.; COWIE, A.; HOBBS, N.; KOHLI, P.; KLEYWEGT, G.; BIRNEY, E.; HASSABIS, D.; VELANKAR, S. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Research, [*s. l.*], v. 50, n. D1, p. D439–D444, 7 jan. 2022. https://doi.org/10.1093/nar/gkab1061.

WANG, J.; CHITSAZ, F.; DERBYSHIRE, M. K.; GONZALES, N. R.; GWADZ, M.; LU, S.; MARCHLER, G. H.; SONG, J. S.; THANKI, N.; YAMASHITA, R. A.; YANG, M.; ZHANG, D.; ZHENG, C.; LANCZYCKI, C. J.; MARCHLER-BAUER, A. The conserved domain database in 2023. Nucleic Acids Research, [*s. l.*], v. 51, n. D1, p. D384–D388, 6 jan. 2023. https://doi.org/10.1093/nar/gkac1096.

WASĄG, P.; SUWIŃSKA, A.; RICHERT, A.; LENARTOWSKA, M.; LENARTOWSKI, R. Plant-specific calreticulin is localized in the nuclei of highly specialized cells in the pistil—new observations for an old hypothesis. Protoplasma, [*s. l.*], v. 261, n. 6, p. 1171–1184, 1 nov. 2024. https://doi.org/10.1007/s00709-024-01961-y.

WEBB, B.; SALI, A. Comparative Protein Structure Modeling Using MODELLER. Current Protocols in Bioinformatics, [*s. l.*], v. 47, n. 1, p. 5.6.1-5.6.32, 2014. https://doi.org/10.1002/0471250953.bi0506s47.

WIEDERSTEIN, M.; SIPPL, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Research, [*s. l.*], v. 35, n. Web Server issue, p. W407-410, jul. 2007. https://doi.org/10.1093/nar/gkm290.

WISSER, R. J.; OPPENHEIM, S. J.; ERNEST, E. G.; MHORA, T. T.; DUMAS, M. D.; GREGORY, N. F.; EVANS, T. A.; DONOFRIO, N. M. Genome assembly of a Mesoamerican derived variety of lima bean: a foundational cultivar in the Mid-Atlantic USA. G3 Genes|Genomes|Genetics, [*s. l.*], v. 11, n. 11, p. jkab207, 1 nov. 2021. https://doi.org/10.1093/g3journal/jkab207.

ZHU, Y.; WANG, Y.; WEI, Z.; ZHANG, X.; JIAO, B.; TIAN, Y.; YAN, F.; LI, J.; LIU, Y.; YANG, X.; ZHANG, J.; WANG, X.; MU, Z.; WANG, Q. Analysis of oil synthesis pathway in *Cyperus esculentus* tubers and identification of oleosin and caleosin genes. Journal of Plant Physiology, [*s. l.*], v. 284, p. 153961, 1 maio 2023. https://doi.org/10.1016/j.jplph.2023.153961.