

Article

Not peer-reviewed version

---

# Genome-Wide in silico Analysis Expanding the Potential Allergen Repertoire of Mango (*Mangifera indica* L.)

---

Amit Singh , [Annelise N Huynh](#) , Aayan Zarif , [Zhibo Yang](#) , [Nagib Ahsan](#) \*

Posted Date: 9 June 2025

doi: 10.20944/preprints202506.0684.v1

Keywords: Allergome; Computational modeling; Mango; Genome; Proteome



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

## Article

# Genome-Wide in silico Analysis Expanding the Potential Allergen Repertoire of Mango (*Mangifera indica* L.)

Amit Singh <sup>1</sup>, Annelise N Huynh <sup>2</sup>, Aayan Zarif <sup>3</sup>, Zhibo Yang <sup>1,4</sup> and Nagib Ahsan <sup>1,5,\*</sup>

<sup>1</sup> Department of Chemistry and Biochemistry, The University of Oklahoma, Norman, OK, USA

<sup>2</sup> Dodge Family College of Arts and Sciences, University of Oklahoma, Norman, OK, USA

<sup>3</sup> Mercy School Institute, Edmond, OK, USA

<sup>4</sup> Department of Biochemistry and Physiology, The University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

<sup>5</sup> Mass Spectrometry, Proteomics and Metabolomics Core Facility, Stephenson Life Sciences Research Center, The University of Oklahoma, Norman, OK, USA

\* Correspondence: nahsan@ou.edu; Tel.: +1 405-325-2852

**Abstract:** The potential of a protein to cause an allergic reaction is often assessed using a variety of computational techniques. Leveraging advances in high-throughput protein sequence data coupled with in silico or computational methods can be used to systematically analyze large proteomes for allergenic potential. Despite its widespread consumption and growing clinical reports of hypersensitivity, the full extent of their allergenicity is yet unknown. In this study, for the first time, we conducted a genome-wide in silico analysis by analyzing a total of 54,010 protein sequences to identify the complete spectrum of potential mango allergens. These proteins were analyzed using various bioinformatics tools to predict their allergenic potential based on sequence similarity, structural features, and known allergen databases. In addition to the known mango allergens, including Man i1, Man i2, and Man i3, our findings demonstrated that several isoforms of cysteine protease, non-specific lipid-transfer protein (LTP), legumin B-like, 11S globulin, vicilin, thaumatin-like protein, and ervatamin-B family proteins exhibited strong allergenic potential, with >80% 3D epitope identity, >70% linear 80 aa window identity, and matching with >80 known allergens. Thus, a genome-wide in silico study provided a comprehensive profile of the possible mango allergome, which could help identify the low allergen-containing mango cultivars and aid in the development of accurate assays for variety-specific allergic reactions.

**Keywords:** allergome; computational modeling; mango; genome; proteome

## 1. Introduction

The identification and characterization of allergens and allergen-like proteins from food sources is a central concern in not only allergy research but is also of growing importance of the food industry, where safety and risk assessment are crucial for regulatory approval, product development, and consumer protection [1–3]. On the other hand, identification of a complete allergen profile from any food species is challenging and time-consuming because of the complexity of the proteins involved and the individual variability in immune responses. However, computational methods and the availability of complete genome sequences can streamline this process, leveraging bioinformatics tools to analyze protein structures and identify potential allergenic components more efficiently. These computational approaches hinge on key features, such as sequence similarity, conserved structural motifs, and immunologically relevant domains that are commonly associated with known allergenic proteins [4].

Although *in silico* and computational tools cannot distinguish between the sensitization and elicitation phases of allergy development, international guidelines, including those from the FAO/WHO and the European Food Safety Authority (EFSA), endorse their use, defining a protein as potentially allergenic if it shares more than 35% identity over 80 amino acids, or contains exact matches of 6-8 contiguous amino acids with a known allergen [5]. These thresholds recognize *in-silico* predictions as a first-pass assessment in allergen risk evaluation, form the basis of many allergenicity prediction platforms, and are widely adopted to streamline regulatory evaluations in both academic and industry contexts [4].

Fruit allergies, which impact 0.03% to 8% of the global population, are a growing health concern [2,6]. Despite the incidence rate, many fruits remain underrepresented in allergen databases. In particular, tropical fruits have seen a surge in global consumption yet remain poorly characterized with respect to their allergenic profiles [4,7]. Mango (*Mangifera indica* L.), a widely consumed tropical fruit valued for its flavor and nutritional benefits, exemplifies the growing divide between growing global dietary expansion and the gap in allergen knowledge. In recent decades, mango has experienced an evident rise in popularity beyond its native growing regions. Namely, in the United States, mango imports have quadrupled over the past two decades to meet consumer demands, from smoothies to skincare products and therefore, contributing to a broader and more frequent exposure across all age groups (USDA Economic Research Service). A recent meta-analysis of studies on fruit allergies found that mangos are one of the top five tropical fruits that trigger fruit allergies [6]. Multiple studies have implicated mango as a significant allergen, with reported prevalence ranging from 0.3% in Switzerland to 16% in Thailand, and skin test positivity rates as high as 42.3% in some Chinese cohorts [4,8].

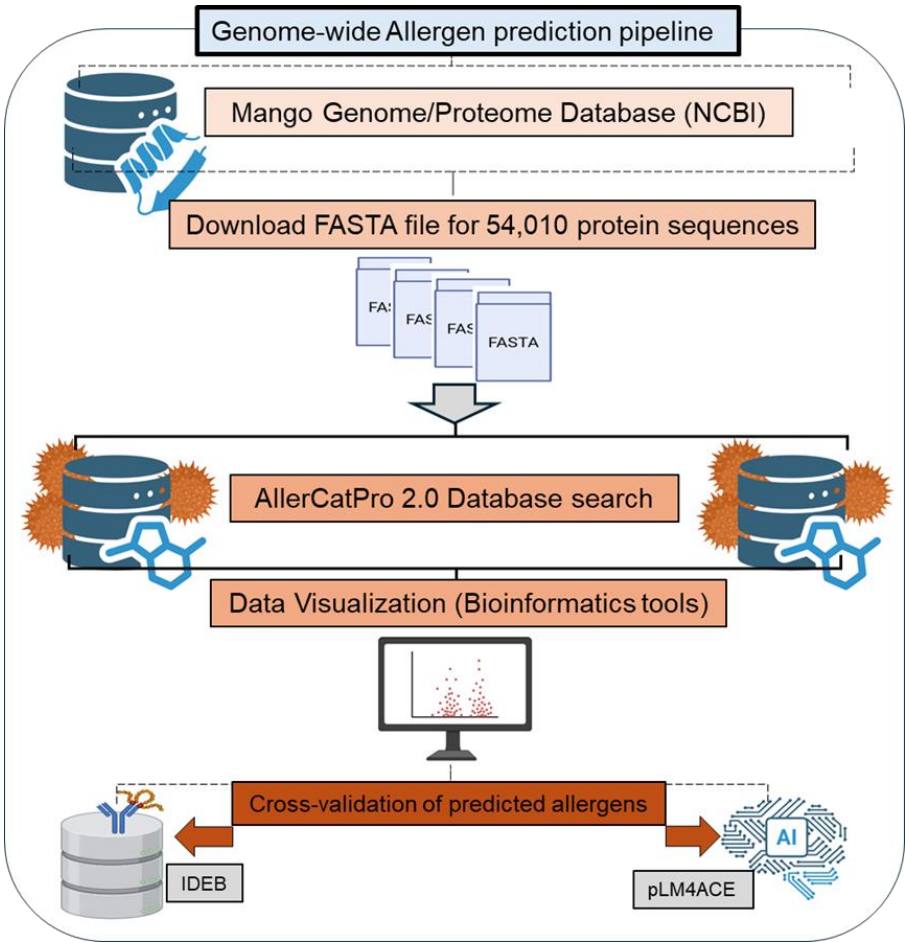
Mango allergy manifests primarily in two immunological forms: immediate (Type I) and delayed (Type IV) hypersensitivity [8]. These reactions often manifest as contact dermatitis, rash, eczema, and blistering. As documented by Ukleja-Sokolowska et al, a 30-year-old woman developed generalized urticaria, facial oedema, strong stomach pain, and watery diarrhea within several minutes of eating a mango [9]. Despite its clinical significance, the allergenic profile of mango remains incompletely characterized. According to the WHO/IUIS Allergen Nomenclature Sub-committee (<https://allergen.org/>), only three allergenic proteins have been formally identified in mango, including chitinase (Man i1), pathogenesis-related (PR) protein (Man i2), and profilin (Man i4), representing only a fraction of the mango's total proteome. A recent study by Guo and Cong (2024) listed a couple more proteins [4], including glyceraldehyde-3-phosphate dehydrogenase (GAPDH) [10], lipoxygenase and glucanase [11], which showed cross-reactivity with allergenic patient sera. Furthermore, mango is not only antigenic on its own but also exhibits extensive cross-reactivity with other food and inhalant allergens [12]. These existing findings underscore the broader immunological risk that mango poses to individuals with existing sensitizations, emphasizing the need for a deeper and wider understanding of mango's allergenic potential.

*In silico* approaches have been successfully employed to identify the potential allergenicity of many proteins in various plant species [13–19]. However, to our knowledge, no genome-wide analysis has been conducted to profile the complete allergome of any tropical fruits including mango. With the majority of the proteome remaining uncharacterized for allergenicity, a genome-wide evaluation is both timely and necessary. In this study, we conducted a genome-wide *in-silico* screening of 54,010 protein sequences from mango. Using a combination of allergenicity prediction algorithms, such as AllergenOnline, AllerCatPro, epitope-based scanning, and structural homology assessments, we aimed to systematically identify and catalog potential allergens in mango. By expanding the known allergen repertoire of mango, our analysis provides a valuable foundation for future research in allergy diagnostics, immunotherapy design, and food safety assessment, especially pertinent as tropical fruits like mango become increasingly integrated into global diets and diverse consumer products [20].

## 2. Materials and Methods

2.1. Collection of Mango Protein Sequences:

The mango genome (CATAS\_Mindica\_2.1, assembled in 2020 by the Beijing Institute of Genomics, Chinese Academy of Sciences, using the widely cultivated Alphonso variety (RefSeq Accession: GCF\_011075055.1) currently consists of 54,010 protein entries, which are publicly accessible under the NCBI Protein database (Taxonomy ID: 29780). To assess the allergenic potential of mango proteins, these protein sequences were downloaded into FASTA format and curated to include essential metadata such as NCBI accession numbers, protein names, sequence lengths, and source information. Figure 1 outlines the stepwise bioinformatics pipeline used to assess the allergenic potential of *Mangifera indica* proteins.



**Figure 1.** Bioinformatics pipeline for allergen prediction in *Mangifera indica* proteins using database screening, data visualization, and AI-based hypersensitivity analysis.

2.2. In Silico Platform for Identification of Potential Mango Allergens

The curated protein sequence dataset served as the foundation for our genome-wide allergen prediction screening and was subjected to AllerCatPro 2.0, a well-established prediction tool that integrates sequence similarity, epitope analysis, and structural comparisons [21,22]. Since AllerCatPro can only analyze 50 protein sequences at a time, we split the mango proteome dataset (54010 protein sequences) into 1081 FASTA files and submitted each one for analysis. We additionally implemented a filter with two criteria: (i)  $\geq 70\%$  sequence identity over an 80-amino-acid linear window, and/or (ii) best match with at least 80 known allergen entries, to ensure the strong consistency and specificity of the allergen prediction results (strong and weak evidence).

2.3. Bioinformatics Tools for Data Processing

Scatter and bubble plots were generated using Python. Multiple amino acid sequence alignment was performed, and a phylogenetic tree was generated using an open-source sequence alignment platform, Clustal Omega (<https://www.ebi.ac.uk>). Heatmaps on percentage (%) identity were generated using Excel. B-cell epitope prediction was performed using Immune Epitope Database (IEDB) (<http://tools.iedb.org/bcell/>) following the Kolaskar and Tongaonkar method [23].

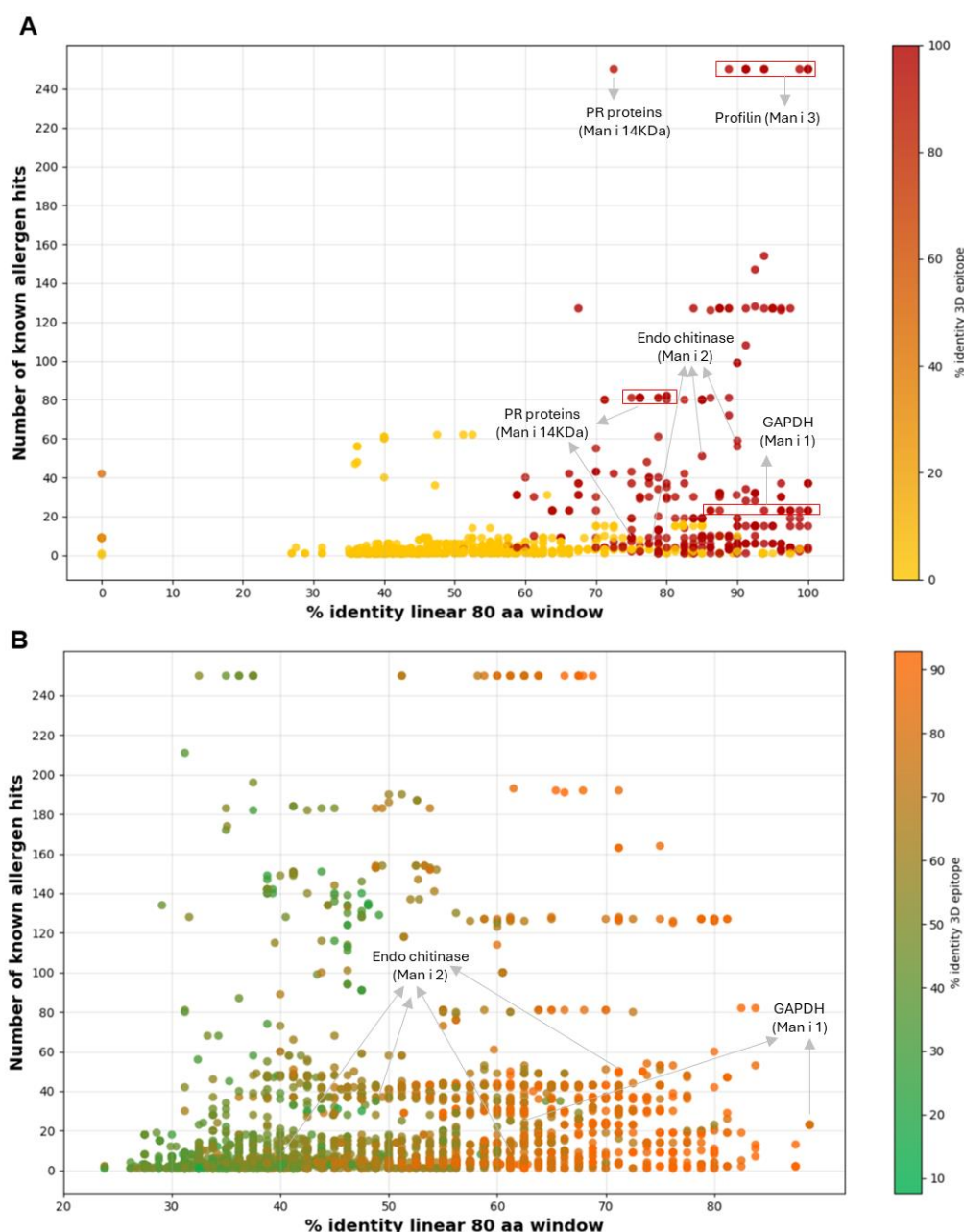
### 3. Results and Discussion

The identification of allergenic proteins in food sources is a growing priority for researchers and regulatory agencies alike, especially as global diets diversify to include a wider range of under-characterized foods. In this regard, *in silico* approaches in conjunction with high-quality genome sequence data can enable scalable, quick evaluation of allergen identification, structural motifs, epitope similarity, and sequence homology to clinically verified allergens.

#### 3.1. Identification of Known Mango Allergens

Our results indicated that five of the six proteins and/or protein families that were previously identified as mango allergens [4] were successfully identified and classified as having substantial allergy potential, demonstrating the effectiveness of the AllerCatPro as a powerful allergen prediction tool. These proteins include glyceraldehyde 3-phosphate dehydrogenase (GAPDH), profilin,  $\beta$ -1,3-glucanase, Bet v 1-like homologous protein/pathogenesis-related proteins (PR proteins), and Type I chitinases (Figure 2, Table S1).

It is important to note that members of these families were identified in both the high and weak allergen evidence categories, indicating that they may have clinical significance and that additional experimental validation is necessary. For example, out of the 39 chitinase isoforms that were found, three had strong evidence and 36 had weak evidence. Comparably, PR proteins have 25 isoforms, 8 of which are strong and 17 of which are weak; GAPDH has 19 isoforms, 9 of which are strong and 10 of which are weak; and profilins have 13 isoforms, all of which exhibit strong evidence. These findings highlight the varying levels of evidence supporting the role of different protein isoforms in biological processes. Further research is essential to clarify the functions of the weaker isoforms and to explore their potential as mango allergens. Importantly, GAPDH was recognized as Man i 1 in a recent publication by Guo and Cong [4], although chitinase proteins were listed as Man i 1 in the WHO/IUIS Allergen Nomenclature Sub-committee database. Similarly, chitinases were named as Man i 2 and Man I chitinase, profilin as Man i 3.01 and Man i 3.02, and Bet v 1-like homologous protein/pathogenesis-related proteins (PR proteins) as Man i 14 kDa. As a result, we decided to use the name of the mango allergens in this investigation in line with the most recent publication [4].

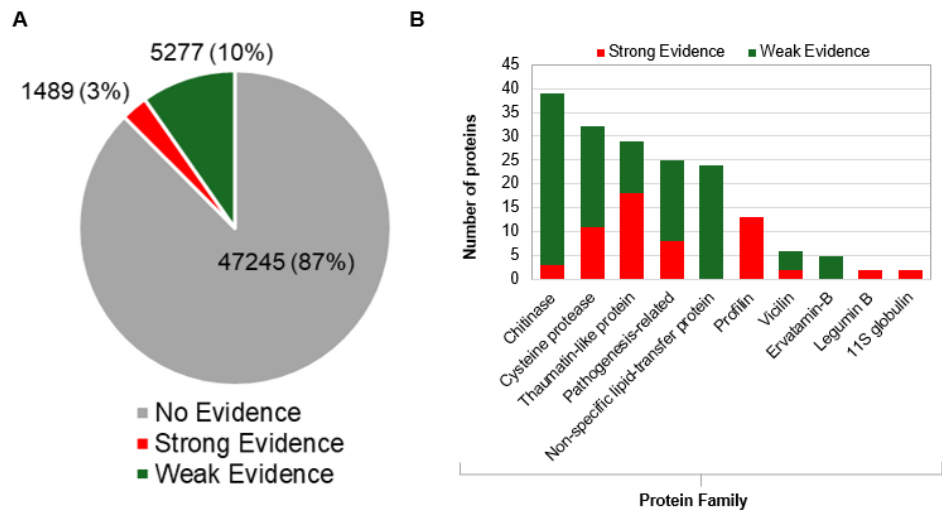


**Figure 2.** Scatter plot analysis of mango proteins with predicted allergenicity by AllerCatPro 2.0. (A) Proteins with strong allergen evidence, plotted by % identity over an 80 amino acid window (x-axis) and number of known allergen hits (y-axis). The dot color indicates 3D epitope identity (yellow to red gradient). (B) Proteins with weak allergen evidence, shown using the same axes; dot color represents 3D epitope identity (green to orange gradient).

### 3.2. Potential Allergens in Mango Genome

In addition to the recognized mango allergens, AllerCatPro discovered hundreds more proteins belonging to different protein families that have a high probability of being allergens. In particular, a total of 1,489 (3%) and 5,277 proteins (10%) out of the 54,010 mango protein sequences had strong and weak allergenicity potential, respectively (Figure 3A, Table S1). The discovery of multiple allergenic protein families that are not yet considered as mango allergens but show a high degree of resemblance to known allergens from other fruit species further validated the efficacy of this *in silico* workflow. For instance, a number of proteins from cysteine proteases, legumin, non-specific lipid transfer proteins (LTP), thaumatin-like protein, vicilin, ervatamin-like proteins, and globulins are among these families that showed strong allergenic potency (Figure 3B). These findings expand the

known allergome of mango well beyond what has been previously reported in allergen databases and underscore the potential of in silico approaches to uncover novel candidate allergens based on structural and sequence homology with clinically established allergens from other sources (Figure 3B).



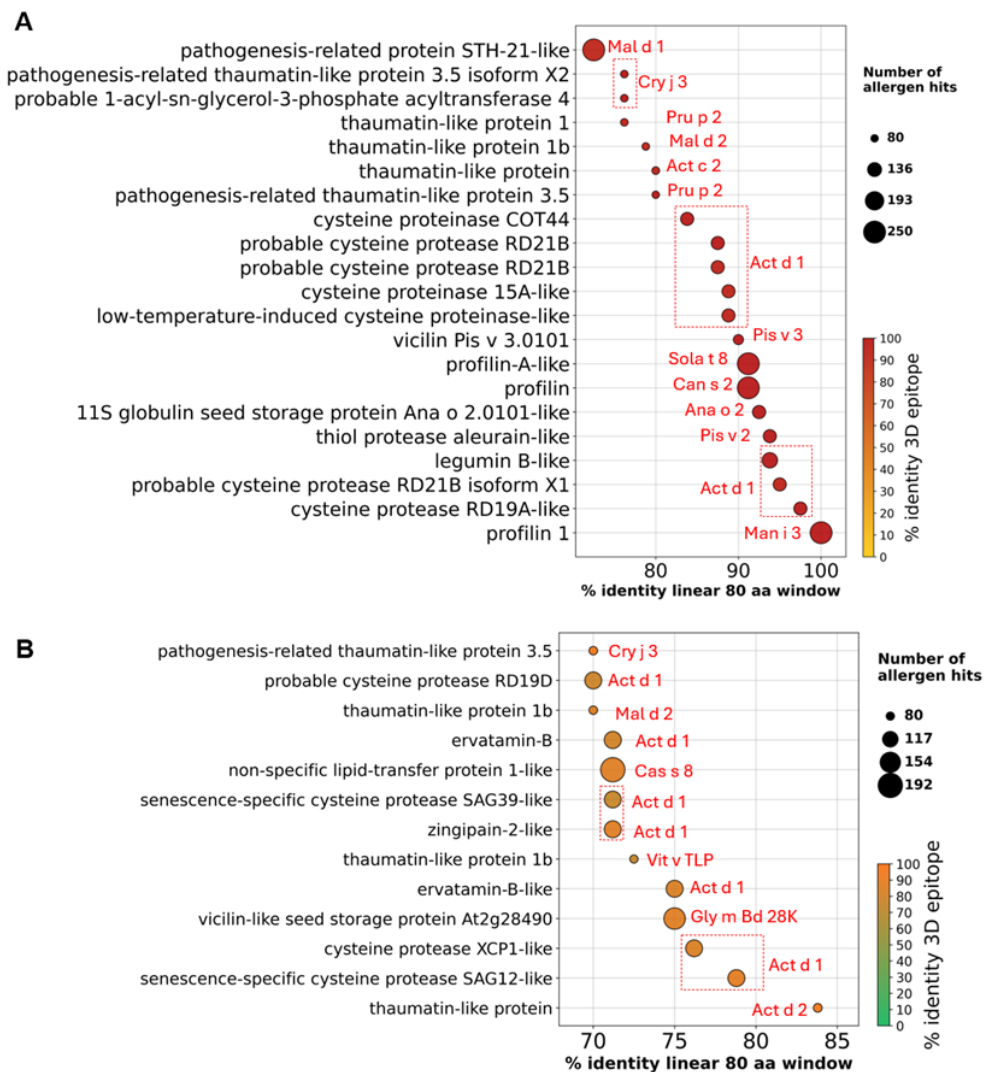
**Figure 3.** In silico allergenicity classification of 54,010 mango proteins. (A) Pie chart showing strong (n = 1,489; 3%), weak (n = 5,277; 10%), and no (87%) allergen evidence. (B) Bar graph showing isoform-level classification of five known mango allergens into strong and weak categories.

3.3. Identification of High-Confidence Allergen Protein Families in Mango

While approximately 1,500 proteins demonstrated strong evidence of allergenicity and over 5,200 proteins were considered to pose a lower allergenic risk, we applied an additional filter ( $\geq 70\%$  identity across an 80-amino-acid linear window and  $\geq 80$  known allergen hits) to refine the candidate list and better categorize protein families. This approach reduced the dataset to 63 high-confidence allergens and 185 proteins with moderate evidence (Figure 4, Table S2).

We discovered that the second-highest category of proteins in the mango genome indicated a significant risk of allergies were those belonging to the cysteine protease family (Figure 3B). Several cysteine protease protein isoforms were significantly matched with kiwi (*Actinidia deliciosa*) Act d1 allergen (Figure 4A). The top three cysteine protease proteins, which are accession XP\_044488586.1, XP\_044511549.1, and XP\_044509143.1, were aligned with Act d1 using multiple sequence alignment (Figure S1) and showed  $>82\%$  amino acid sequence identity with Act d1 (Figure 5A).

Furthermore, amino acid sequence alignment reveals a highly conserved B-Cell epitope region (Figure S2), and the epitope scores of these cysteine proteases were higher than those of the known Act d1 (Figure S3). A newly created AI-based allergy prediction tool called pLM4AIg [24], further cross-validated the allergen potential for these cysteine proteases and confirmed the significant allergenicity comparable to Act d1.

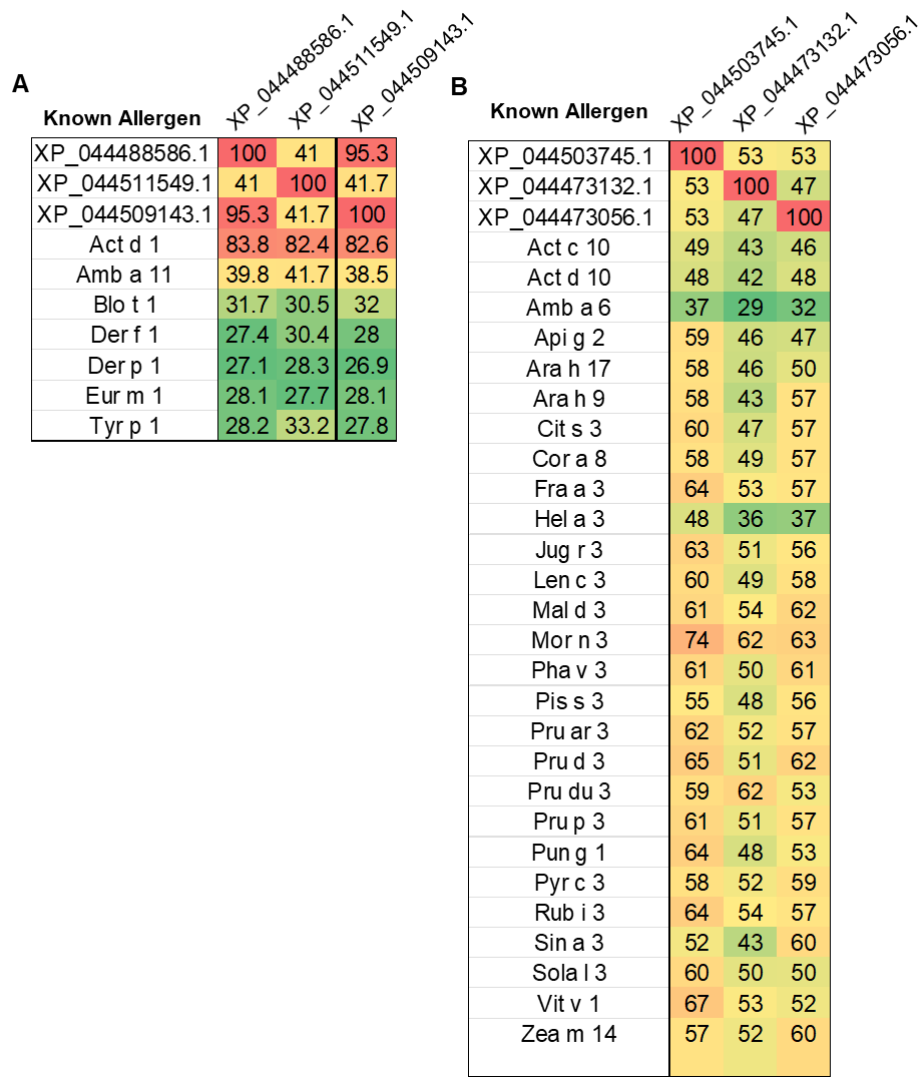


**Figure 4.** High-confidence mango allergen candidates identified using refined thresholds. (A) Proteins predicted as strong allergens (n = 63). Dot size indicates number of allergen hits; color shows % identity to 3D epitope. Several cysteine protease isoforms showed strong matches with kiwi allergen Act d 1. (B) Proteins predicted as weak allergens (n = 185) using the same thresholds.

Cysteine proteases are increasingly recognized as clinically important plant food allergens, with several members of this protein family known to trigger IgE-mediated hypersensitivity reactions [25,26]. These enzymes, which play key roles in plant defense and ripening, are structurally stable and resistant to gastrointestinal digestion—features that contribute to their heightened allergenic potential [25,27]. Well-characterized examples include Act d 1 from kiwi, Ana c 2 from pineapple, and Cari p 1 from papaya, all of which have been associated with severe systemic allergic responses [25]. Their conserved structure and proteolytic function are believed to support both direct immune sensitization and disruption of epithelial barriers, increasing the likelihood of allergen exposure [28,29]. The high sequence identity and epitope similarity observed between mango cysteine proteases and Act d 1 point to a significant risk of cross-reactivity in sensitized individuals, aligning with the broader clinical significance of this allergen family.

Lipid transfer proteins (LTPs), which are primarily low molecular weight heat-resistant proteins, are among the most well-characterized plant allergens that can be found in substantial quantities in the fruits [30]. Interestingly, our findings demonstrated that only two non-specific lipid-transfer proteins (LTPs), which correspond to accession numbers XP\_044477456.1 and XP\_044465367.1 in the mango genome, were identified as allergens with strong evidence. Despite being matched to known allergens Jug r 8 and Cry j LTP, however, these LTPs did not meet our high confidence criteria. On

the other hand, among the weakly evident allergen list, a total of 24 LTPs met our high confidence criteria (Table S2). Among these LTPs, a non-specific lipid-transfer protein 1-like (XP\_044503745.1) protein matched with 192 known LTPs, showing 71.2% identity in a linear 80 aa window and 85.7% identity in a 3D epitope. Notably, this mango LTP showed the greatest match with the chestnut (*Castanea sativa*) Cas s8 allergen (Table S2, Figure 4B).



**Figure 5.** Heatmap showing sequence identity between mango proteins and known allergens. (A) Cysteine protease proteins XP\_044488586.1, XP\_044511549.1, and XP\_044509143.1 were aligned with cysteine protease allergens from various species. (B) Non-specific lipid-transfer proteins (LTPs) XP\_044503745.1, XP\_044473132.1, and XP\_044473056.1 were aligned with allergens across multiple plant species. Color scale reflects percentage identity: red = high, yellow = moderate, green = low.

To further demonstrate the similarities among the LTP isoforms in the mango genome, we selected two more LTPs (XP\_044473132.1 and XP\_044473056.1) and performed multiple sequence alignment in addition to a pairwise comparison of the known LTP allergens listed in the WHO/IUIS Allergen Nomenclature Sub-committee database. Our findings demonstrated that mango LTPs exhibited a conserved epitope binding location among plant LTPs and shared >50% amino acid sequence identity with all well-characterized fruit LTPs (Figure 5B, Figure S4). Cross-reactivity among LTPs from botanically distant fruits is well documented, with shared IgE-binding epitopes and overlapping T-cell reactivity contributing to the clinical phenomenon of LTP syndrome, where

sensitization to one LTP (e.g., Pru p 3 from peach) can trigger allergic responses to multiple plant foods [31–33].

4. Conclusions

In conclusion, our study highlights the strength of genome-wide in silico approaches in identifying potential allergens in food sources that have not been fully explored before. We were able to examine over 54,000 protein sequences of mango genome as a case study, revealing a more comprehensive allergen profile that has been previously documented. In addition to confirming the known allergens from mangos, we also found new protein families that have substantial similarities to recognized allergens from other plant sources, such as lipid transfer proteins and cysteine proteases. These proteins showed good matches in allergy databases, conserved epitope areas, and high sequence identity, suggesting a high chance of cross-reactivity in sensitive people. The increasing number of clinical reports that associate mango with allergic reactions can be explained by this improved understanding of the mango proteome. Our findings also raise the possibility that there are a number of additional proteins with potential for allergies that have not been identified because of the limited scope of previous research on mango allergy. In order to determine which proteins, require further investigation in clinical or experimental settings, researchers may now effectively and affordably screen entire genomes using thorough in silico methods.

Overall, this work contributes to filling the gap in allergen knowledge for tropical fruits like mango and sets the stage for future research in allergy diagnostics and food safety. These findings may also help breeders and food scientists develop low-allergen cultivars and safer food products, ultimately benefiting individuals with food sensitivities around the world.

**Supplementary Materials:** The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

**Author contributions:** Conceptualization, N.A; Data curation and analysis, A.S, A.Z, A.N.H, N.A; Bioinformatics, N.A, A.S; Primary draft of the manuscript, N.A, A.S; writing, review, and editing, N.A, A.S; Z.Y; All authors have read and agreed to the published version of the manuscript.

**Data Availability Statement:** All data are presented in the main text, as well as the Supplementary Materials.

**Acknowledgments:** N.A. gratefully acknowledges the initial funding support from the OU VPRP Office for the establishment of the Proteomics Core Facility.

Abbreviations

The following abbreviations are used in this manuscript:

LTP	Lipid Transfer Protein
PR	Pathogenesis-Related
GPI	Glycosylphosphatidylinositol
EFSA	European Food Safety Authority
IEDB	Immune Epitope Database
pLm4Alg	Plant-based Language Model for Allergenicity
CATAS	Chinese Academy of Tropical Agricultural

References

1.

Liu W, Yang Q, Wang Z, Wang J, Min F, Yuan J, et al. Quantitative food allergen risk assessment: Evolving concepts, modern approaches, and industry implications. *Compr Rev Food Sci Food Saf* [Internet]. 2025 [cited 2025 May 23];24(2):e70132. doi: 10.1111/1541-4337.70132

2.

Präger L, Simon JC, Treudler R. Food allergy - New risks through vegan diet? Overview of new allergen sources and current data on the potential risk of anaphylaxis. *JDDG J Dtsch Dermatol Ges* [Internet]. 2023 [cited 2025 May 23];21(11):1308–13. doi: 10.1111/ddg.15157

3. Shaheen N, Hossen MdS, Akhter KT, Halima O, Hasan MdK, Wahab A, et al. Comparative Seed Proteome Profile Reveals No Alternation of Major Allergens in High-Yielding Mung Bean Cultivars. *J Agric Food Chem* [Internet]. 2024 Jun 19 [cited 2025 May 23];72(24):13957–65. doi: 10.1021/acs.jafc.4c01054
4. Guo H, Cong Y. Recent advances in the study of epitopes, allergens and immunologic cross-reactivity of edible mango. *Food Sci Hum Wellness* [Internet]. 2024 May 25 [cited 2025 May 23];13(3):1186–94. doi: 10.26599/FSHW.2022.9250098
5. Maurer-Stroh S, Krutz NL, Kern PS, Gunalan V, Nguyen MN, Limvipuvadh V, et al. AllerCatPro: prediction of protein allergenicity potential from the protein sequence. *Bioinforma Oxf Engl*. 2019 Sep 1;35(17):3020–7. doi: 10.1093/bioinformatics/btz029
6. Krikeerati T, Rodsaward P, Nawiboonwong J, Pinyopornpanish K, Phusawang S, Sompornrattanaphan M. Revisiting Fruit Allergy: Prevalence across the Globe, Diagnosis, and Current Management. *Foods Basel Switz*. 2023 Nov 10;12(22):4083. doi: 10.3390/foods12224083
7. Jeong KY, Lopata AL. Editorial: Spotlight on allergy research in Asia. *Front Allergy* [Internet]. 2024 Mar 6 [cited 2025 May 23];5. doi: 10.3389/falgy.2024.1371795
8. Berghea EC, Craiu M, Ali S, Corcea SL, Bumbacea RS. Contact Allergy Induced by Mango (*Mangifera indica*): A Relevant Topic? *Medicina (Mex)* [Internet]. 2021 Nov [cited 2025 May 23];57(11):1240. doi: 10.3390/medicina57111240
9. Ukleja-Sokołowska N, Gawrońska-Ukleja E, Lis K, Żbikowska-Gotz M, Sokołowski Ł, Bartuzi Z. Anaphylactic reaction in patient allergic to mango. *Allergy Asthma Clin Immunol* [Internet]. 2018 Oct 31 [cited 2025 May 23];14(1):78. doi: 10.1186/s13223-018-0294-1
10. Paschke A, Kinder H, Zunker K, Wigotzki M, Weißbecher R, Vieluf D, et al. Characterization of Allergens in Mango Fruit and Ripening Dependence of the Allergenic Potency. *Food Agric Immunol* [Internet]. 2001 Mar 1 [cited 2025 May 23];13(1):51–61. doi: 10.1080/09540100051074220
11. Cardona EEG, Heathcote K, Teran LM, Righetti PG, Boschetti E, D'Amato A. Novel low-abundance allergens from mango via combinatorial peptide libraries treatment: A proteomics study. *Food Chem* [Internet]. 2018 Dec 15 [cited 2025 May 23];269:652–60. doi: 10.1016/j.foodchem.2018.06.113
12. Recent advances in the study of epitopes, allergens and immunologic cross-reactivity of edible mango [Internet]. [cited 2025 May 23].
13. Garino C, Coisson JD, Arlorio M. In silico allergenicity prediction of several lipid transfer proteins. *Comput Biol Chem*. 2016 Feb;60:32–42. doi: 10.1016/j.compbiolchem.2015.11.006
14. Kulkarni A, Ananthanarayan L, Raman K. Identification of putative and potential cross-reactive chickpea (*Cicer arietinum*) allergens through an in silico approach. *Comput Biol Chem*. 2013 Dec;47:149–55. doi: 10.1016/j.compbiolchem.2013.08.003
15. Halima O, Najar FZ, Wahab A, Gamagedara S, Chowdhury AI, Foster SB, et al. Lentil allergens identification and quantification: An update from omics perspective. *Food Chem Mol Sci*. 2022 Jul 30;4:100109. doi: 10.1016/j.fochms.2022.100109
16. Bastiaan-Net S, Pina-Pérez MC, Dekkers BJW, Westphal AH, America AHP, Ariëns RMC, et al. Identification and in silico bioinformatics analysis of PR10 proteins in cashew nut. *Protein Sci Publ Protein Soc*. 2020 Jul;29(7):1581–95. doi: 10.1002/pro.3856
17. Jamakhani M, Lele SS, Rekadwad B. In silico assessment data of allergenicity and cross-reactivity of NP24 epitopes from *Solanum lycopersicum* (Tomato) fruit. *Data Brief*. 2018 Dec;21:660–74. doi: 10.1016/j.dib.2018.09.074
18. Zhou W, Bias K, Lenczewski-Jowers D, Henderson J, Cupp V, Ananga A, et al. Analysis of Protein Sequence Identity, Binding Sites, and 3D Structures Identifies Eight Pollen Species and Ten Fruit Species with High Risk of Cross-Reactive Allergies. *Genes*. 2022 Aug 17;13(8):1464. doi: 10.3390/genes13081464
19. Hernández-Lao T, Rodríguez-Pérez R, Labella-Ortega M, Muñoz Triviño M, Pedrosa M, Rey MD, et al. Proteomic identification of allergenic proteins in holm oak (*Quercus ilex*) seeds. *Food Chem*. 2025 Feb 1;464(Pt 1):141667. doi: 10.1016/j.foodchem.2024.141667
20. Berghea EC, Craiu M, Ali S, Corcea SL, Bumbacea RS. Contact Allergy Induced by Mango (*Mangifera indica*): A Relevant Topic? *Med Kaunas Lith*. 2021 Nov 13;57(11):1240. doi: 10.3390/medicina57111240

21. Nguyen MN, Krutz NL, Limviphuvadh V, Lopata AL, Gerberick GF, Maurer-Stroh S. AllerCatPro 2.0: a web server for predicting protein allergenicity potential. *Nucleic Acids Res.* 2022 Jul 5;50(W1):W36–43. doi: 10.1093/nar/gkac446
22. Krutz NL, Kimber I, Winget J, Nguyen MN, Limviphuvadh V, Maurer-Stroh S, et al. Application of AllerCatPro 2.0 for protein safety assessments of consumer products. *Front Allergy.* 2023;4:1209495. doi: 10.3389/falgy.2023.1209495
23. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett [Internet].* 1990 Dec 10 [cited 2025 May 23];276(1):172–4. doi: 10.1016/0014-5793(90)80535-Q
24. Du Z, Ding X, Hsu W, Munir A, Xu Y, Li Y. pLM4ACE: A protein language model based predictor for antihypertensive peptide screening. *Food Chem [Internet].* 2024 Jan 15 [cited 2025 May 23];431:137162. doi: 10.1016/j.foodchem.2023.137162
25. Giangrieco I, Ciardiello MA, Tamburrini M, Tuppo L, Rafaiani C, Mari A, et al. Comparative Analysis of the Immune Response and the Clinical Allergic Reaction to Papain-like Cysteine Proteases from Fig, Kiwifruit, Papaya, Pineapple and Mites in an Italian Population. *Foods [Internet].* 2023 Jan [cited 2025 May 24];12(15):2852. doi: 10.3390/foods12152852
26. Suzuki M, Itoh ,Makoto, Ohta ,Nobuo, Nakamura ,Yoshihisa, Moriyama ,Akihiko, Matsumoto ,Tamami, et al. Blocking of protease allergens with inhibitors reduces allergic responses in allergic rhinitis and other allergic diseases. *Acta Otolaryngol (Stockh) [Internet].* 2006 Jul 1 [cited 2025 May 24];126(7):746–51. doi: 10.1080/00016480500475625
27. Szezińska J, Simińska J, Bielawski W. The roles of cysteine proteases and phytocystatins in development and germination of cereal seeds. *J Plant Physiol [Internet].* 2016 Dec 1 [cited 2025 May 24];207:10–21. doi: 10.1016/j.jplph.2016.09.008
28. Sharma A, Vashisht S, Mishra R, Gaur SN, Prasad N, Lavasa S, et al. Molecular and immunological characterization of cysteine protease from *Phaseolus vulgaris* and evolutionary cross-reactivity. *J Food Biochem [Internet].* 2022 [cited 2025 May 24];46(9):e14232. doi: 10.1111/jfbc.14232
29. Soh WT, Zhang J, Hollenberg MD, Vliagoftis H, Rothenberg ME, Sokol CL, et al. Protease allergens as initiators–regulators of allergic inflammation. *Allergy [Internet].* 2023 [cited 2025 May 24];78(5):1148–68. doi: 10.1111/all.15678
30. Anagnostou A. Lipid transfer protein allergy: An emerging allergy and a diagnostic challenge. *Ann Allergy Asthma Immunol Off Publ Am Coll Allergy Asthma Immunol.* 2023 Apr;130(4):413–4. doi: 10.1016/j.anai.2023.01.033
31. Andersen MBS, Hall S, Dragsted LO. Identification of european allergy patterns to the allergen families PR-10, LTP, and profilin from Rosaceae fruits. *Clin Rev Allergy Immunol.* 2011 Aug;41(1):4–19. doi: 10.1007/s12016-009-8177-3
32. Costa J, Mafra I. Rosaceae food allergy: a review. *Crit Rev Food Sci Nutr.* 2023;63(25):7423–60. doi: 10.1080/10408398.2022.2045897
33. Skypala IJ, Asero R, Barber D, Cecchi L, Diaz Perales A, Hoffmann-Sommergruber K, et al. Non-specific lipid-transfer proteins: Allergen structure and function, cross-reactivity, sensitization, and epidemiology. *Clin Transl Allergy.* 2021 May;11(3):e12010. doi: 10.1002/clt2.12010

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.