

Phylogenetic and Sequence Analysis of a Putative Powdery Mildew Resistance Protein from *Cucumis melo* L.

Majjiga Komala^{1*}, K. Gopal², P. Latha³ and N. Premalatha⁴

¹College of Horticulture, Chinalataripi, Dr. Y.S.R. Horticultural University, Andhra Pradesh, India

²College of Horticulture, Anantharajupeta, Dr. Y.S.R. Horticultural University, Andhra Pradesh, India

³Coconut Research Station, Tamil Nadu Agricultural University, Aliyar Nagar, Tamil Nadu, India

⁴Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

*Corresponding author: komalayadav89@gmail.com (ORCID ID: 0000-0001-8647-4482)

Paper No. 1067

Received: 17-11-2022

Revised: 26-02-2023

Accepted: 04-03-2023

ABSTRACT

Cucumis melo L. is a widely cultivated horticultural crop for its delicious fruits. Melon crop is highly susceptible to various fungal diseases. Powdery mildew is one of the common fungal diseases that severely affects the plant growth thereby compromises the fruit yield and quality. Currently applying chemical fungicides by farmers poses serious threat to human health in the long-term use and also impacts the soil quality as well as leads to adaptation of the causative agent, *Podosphaera xanthii* into fungicide resistant variants. Therefore, at present, numerous studies are being conducted worldwide to develop the melon cultivars resistant to fungal diseases by the use of various resistance breeding approaches as well as stimulating the plant innate immune response pathways. In order to target the plant's innate immunity, prior knowledge on functions of such disease resistance proteins is paramount. Therefore, in this study, here we analyzed the known powdery mildew resistance protein sequence from wheat in order to predict the putative disease resistance candidate protein encoded by *C. melo* genome. We predicted a putative disease resistance protein available in NCBI GenBank database as a putative powdery mildew disease resistance protein based on its sequence similarity to the characterized powdery mildew resistance protein from wheat. We performed its phylogenetic and sequence analysis in relation to the homologous disease resistance proteins from other members of the Cucurbitaceae family and found its evolutionary relationship and high conservation. In addition, its homology model built with the SWISS-MODEL program revealed the presence of a protein fold called Leucine-Rich Repeat (LRR), which is a signature property of proteins conferring innate immunity in plants. Altogether, this study provides valuable insights into the understanding of the conservation of powdery mildew resistance proteins in melon and would help future studies aimed at exploration of the powdery mildew disease resistance mechanisms in melon.

HIGHLIGHTS

- A putative disease resistance protein from *Cucumis melo* L. is predicted to be mediating resistance to powdery mildew.
- Phylogenetic and sequence analysis revealed its conservation and 3D structure model confirmed its role in the plant immune system.

Keywords: Phylogeny, Powdery mildew, Homology model, *Cucumis melo*, Plant immune system

Melon (*Cucumis melo* L.; $2n = 2x = 24$; Cucurbitaceae) is one of the economically important and widely cultivated horticultural crop worldwide. The melon fruits are very delicious in taste with a great aroma and rich in sugars and minerals, thus

How to cite this article: Komala, M., Gopal, K., Latha, P. and Premalatha, N. (2023). Phylogenetic and Sequence Analysis of a Putative Powdery Mildew Resistance Protein from *Cucumis melo* L. *Int. J. Ag. Env. Biotech.*, 16(01): 05-12.

Source of Support: None; **Conflict of Interest:** None





widely consumed as raw fruit. Nonetheless, it is largely used in the food, pharmaceutical and cosmetic industries for the extraction of value-added products of food and health importance.

The fruit yield of melon is severely affected by its susceptibility to various fungal diseases. Among the major fungal diseases, Powdery mildew is probably the most common and widespread disease that drastically affects the melon fruit yield and quality. Powdery mildew in melon is caused by the fungi, *Podosphaera xanthii* (an obligate biotrophic ectoparasite) which is easily recognized by the visual examination of leaf surfaces, petioles, and young stems for the presence of white powdery fungal mass (Zitter *et al.* 1996). Due to the loss of chlorophyll content, the photosynthetic efficiency is severely affected in the diseased plants (Mieslerova *et al.* 2022). Although, the currently applying fungicides could able to control the powdery mildew, the large and long-term use of such fungicides poses serious problems to human health as well as increases the selective pressure on *P. xanthii* to acquire the fungicide resistance (Alengebawy *et al.* 2021; Liang *et al.* 2022). Therefore, the most efficient attractive alternative to the fungicides would be the development of powdery mildew resistant melon cultivars either by plant breeding approaches to introduce the resistance genes or by stimulating the already existing disease resistance genes of plant immune system to fight against the disease (Bentham *et al.* 2020; Maharjan *et al.* 2020; Kesh and Kaushik 2021; Komala and Kuni 2022).

The plant immune system constitutes cell surface and intracellular receptor proteins that recognize the own cell damage associated as well as pathogen associated molecular patterns and activate the downstream effector genes to confer immunity to the plant to tolerate or to resist the biotic stress. Cell surface receptor proteins contain three domains: (i) extracellular ectodomain that recognizes the pathogen ligands, (ii) transmembrane domain that passes through the cell membrane, (iii) intracellular kinase domain that transmit the signals to the downstream effector proteins in the cell. For example, Leucine-Rich Repeat Receptor-Like Kinases (LRR-RLKs) and LRR-Receptor-Like Proteins (LRR-RLPs). LRR-RLPs are devoid of intracellular kinase domain, thus requiring a partner

co-receptor to elicit the immune response (Jones and Dangl, 2006). Whereas intracellular receptor proteins include nucleotide binding, leucine-rich repeat receptor proteins (NLRs), which are of two types: CC-NLRs (CC: coiled coil domain) and TIR-NLRs (TIR: Toll–interleukin-1 receptor) (Jones *et al.* 2016). Unlike animals, plants lack adaptive immunity, thus all above receptors confer only the innate immunity to plants as similar to innate immune proteins of animals. Plant innate immune receptors are encoded by the resistance or 'R' genes which were found to be largely expressed in various disease resistant plant cultivars developed by the plant breeders (Bentham *et al.* 2020).

In this current research article, we explored the melon plant immune system proteins with possible roles in conferring powdery mildew resistance. We predicted a putative disease resistance protein available in NCBI GenBank (accession number: KAA0043493) as a possible candidate protein that might be mediating resistance to powdery mildew in *Cucumis melo* L., based on its sequence similarity to the already known powdery mildew resistance proteins from wheat. Our phylogenetic and sequence alignment analysis of this *C. melo* putative powdery mildew resistance protein (*CmPPMRP*) in relation to its homologous proteins from other important members of Cucurbitaceae family revealed that *CmPPMRP* shares the high conservation of protein sequence with the other family members. In addition, its homology model that we built in this study, showed the structural conservation of protein fold of ectodomain LRR as similar to the known LRR-RLPs. Overall, this study would shed light on the possible role of *CmPPMRP* in conferring powdery mildew resistance in *C. melo* and would help future understandings of its biochemical properties and disease resistance mechanisms.

MATERIALS AND METHODS

Protein BLAST

In order to search for the putative powdery mildew resistance proteins present in *Cucumis melo* L., we used previously well characterized powdery mildew resistance protein, PM3F from wheat (*Triticum aestivum*; NCBI GenBank ID: AAZ23115; Srichumpa *et al.* 2005) as a query sequence to



search specifically against the organism, *C. melo* L. (taxid:3656) in the protein databases with the help of protein BLAST online server (<https://blast.ncbi.nlm.nih.gov>; Altschul *et al.* 1997). The protein sequence of the putative disease resistance protein from *C. melo* var. *makuwa* (NCBI GenBank ID: KAA0043493) was showed highest sequence similarity to PM3F from wheat, thus it predicted to be the putative powdery mildew resistance protein encoded by its corresponding gene existing in the genome of *C. melo*.

Retrieving Protein Sequence Information

The protein sequence of the putative powdery mildew disease resistance protein from *C. melo* (*CmPPMRP*) was retrieved from the NCBI GenBank (KAA0043493). This protein sequence was then used as a query sequence to search for its homologous (similar function) proteins in other members of Cucurbitaceae family by the protein BLAST. Using the BLAST information, the similar proteins present in other species such as *Cucumis sativus* (XP_031741269), *Cucurbita pepo* (XP_023550412), *Cucurbita moschata* (XP_022922246), *Cucurbita argyrosperma* (KAG7016874), *Cucurbita maxima* (XP_022972974) and *Momordica charantia* (XP_022153528) were also retrieved from the NCBI database (<https://www.ncbi.nlm.nih.gov/genbank/>; Clark *et al.* 2016) using their respective accession numbers (given in parenthesis).

Phylogenetic Analysis

Phylogenetic analysis was performed for the *CmPPMRP* and its homologous proteins from other members of Cucurbitaceae using the online server, EMBL-EBI ClustalW2 –Phylogeny (<http://www.ebi.ac.uk/Tools>; Larkin *et al.* 2007, Goujon *et al.* 2010).

Multiple Sequence Alignment

In order to analyze the conservation of amino acid residues among *CmPPMRP* and its homologous proteins, we performed the multiple sequence alignment using the online server, EMBL-EBI Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>; Madeira *et al.* 2022) and it further rendered using the online server, ESPript 3.0 (<https://esprict.ibcp.fr/ESPript/ESPript/>; Robert and Gouet, 2014)

Homology Modelling

In order to build a 3D structure model of *CmPPMRP* using its protein sequence, we first searched for templates and then used the high sequence similarity template, the crystal structure of the chain A of LRR ectodomain of the plant membrane receptor kinase GASSHO1/SCHENGEN3 from *Arabidopsis thaliana* (Protein Data Bank ID: 6S6Q; Okuda *et al.* 2020) with the help of the protein homology modelling online server, SWISS-MODEL (<https://swissmodel.expasy.org>; Waterhouse *et al.* 2018).

RESULTS AND DISCUSSION

CmPPMRP protein sequence identity

We chose a characterized powdery mildew resistance protein from wheat as a query sequence to search for putative powdery mildew resistance proteins in melon, because wheat is not closely related plant species to melon thus it helps in actual prediction of very highly evolutionary conserved proteins of melon plant immune system. The protein BLAST results for the query sequence of known powdery mildew resistance protein from wheat, PM3F (NCBI GenBank: AAZ23115; Srichumpa *et al.* 2005) showed a list of melon proteins with the various levels of sequence identity in a range of 31-46%. Among these, the previously characterized melon proteins such as RGH12, RGA2 showed the highest sequence identity (46%) to the wheat PM3F, whereas some of the putative disease resistance proteins showed the lowest sequence identity of 31%. Among the putative disease resistance proteins, we chose one with highest sequence identity (34%) to wheat PM3F, that is from *C. melo* var. *makuwa* (NCBI GenBank: KAA0043493) for this study. Due to its high sequence identity to wheat PM3F, we anticipated that it could be involved in conferring resistance to powdery mildew in *C. melo*. Thus, we predicted it as a putative powdery mildew resistance protein from *C. melo* (*CmPPMRP*). A recent study evaluated the subcellular localization of proteins encoded by the melon powdery mildew resistance genes such as *CmPMRI* and *CmPMrs* and showed that these proteins are localized in the cell membrane, thus may play key roles in recognition of invading fungal pathogen (Cui *et al.* 2022). Therefore, we further predicted that *CmPPMRP*



might be localized in the cell membrane based on its sequence identity to other membrane proteins of the melon plant immune system.

Evolutionary relationship of *CmPPMRP*

The phylogenetic analysis carried out for *CmPPMRP* in relation to the other homologous proteins from the Cucurbitaceae family shows the divergence of the protein sequences into three independent major clusters (Fig. 1). The first cluster of the phylogenetic tree contains the protein sequences from *C. melo*, *C. sativus*, *M. Charantia* and *C. pepo*, thus indicating these sequences are evolved independently and among these *C. melo* and *C. sativus* shares the same lineage, thus indicating these might be very closely related. The second cluster contains sequences of *C. moschata* and *C. argyrosperma* which are evolved together from the same lineage whereas the sequence of *C. maxima* alone falls into the third independent cluster (Fig. 1). Overall, when compared *CmPPMRP* (*C. melo*) with its homologous proteins, it was found to diverge from the other members of the Cucurbitaceae family, with the exception of *C. sativus* (Fig. 1).

Conservation of *CmPPMRP* protein sequence

The multiple sequence alignment performed for *CmPPMRP* and its homologous proteins from other members of Cucurbitaceae family revealed the interesting insights into the protein sequence conservation among these proteins. The very highly

conserved regions as well as non-conserved regions were observed. Most of the alignment contains highly conserved regions than the non-conserved regions, exhibiting 70-80% similarity between all the aligned protein sequences. For example, the regions spanning the amino acid residues, 420-440, 720-740 and 1000-1020 were found to be very highly conserved in all the protein sequences. Whereas, the regions such as 1-20, 80-100, and 1070-1090 were observed to be non-conserved among these aligned sequences (Fig. 2). When compared between the termini, the N-terminus was conserved more than the C-terminus (Fig. 3). Nevertheless, the region between 940-950 showed variable lengths between the aligned sequences, and the protein sequences of *C. pepo*, *C. moschata*, *C. argyrosperma*, and *C. maxima* were found to be longer in length of about 30 residues than the sequences of *C. melo*, *C. sativus* and *M. charantia* at this region (Fig. 3). This observation was consistent with the phylogenetic analysis, where it was observed that the protein sequences of *C. melo*, *C. sativus* and *M. charantia* were found to be grouped into a sub cluster within the first major cluster (Fig. 1).

3D structure model of *CmPPMRP*

In order to understand the significance of conservation of protein sequence of *CmPPMRP* among the homologous proteins from the Cucurbitaceae family at the structural level, we attempted to build its 3D structure model. We

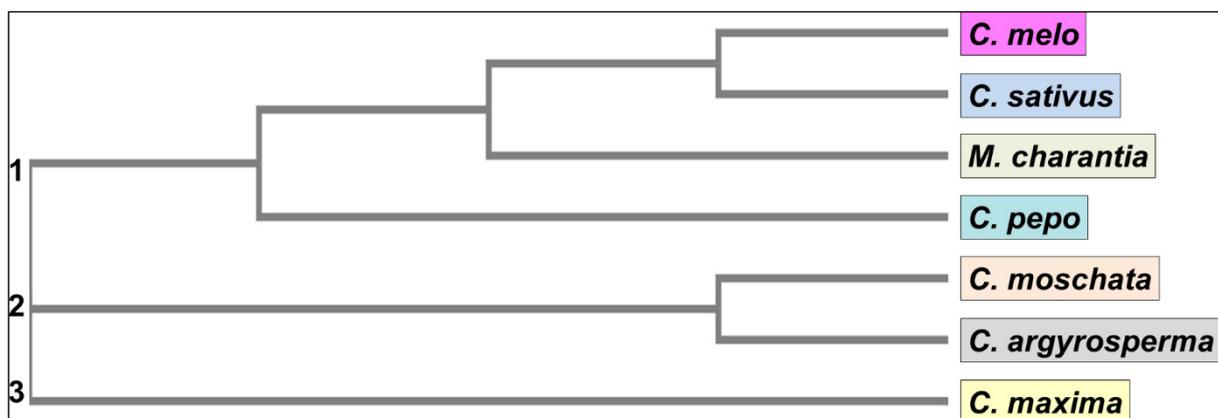


Fig. 1: Phylogenetic tree for putative powdery mildew resistance protein from *Cucumis melo* L. and its homologous proteins from other members of the family, Cucurbitaceae. NCBI GenBank accession numbers for these protein sequences: KAA0043493 (*Cucumis melo* var. *makuwa*), XP_031741269 (*Cucumis sativus*), XP_022153528 (*Momordica charantia*), XP_023550412 (*Cucurbita pepo*), XP_022922246 (*Cucurbita moschata*), KAG7016874 (*Cucurbita argyrosperma*), XP_022972974 (*Cucurbita maxima*). Numbers 1, 2, and 3 indicate each major cluster. ClustalW2 –Phylogeny was used for the construction of the phylogenetic tree

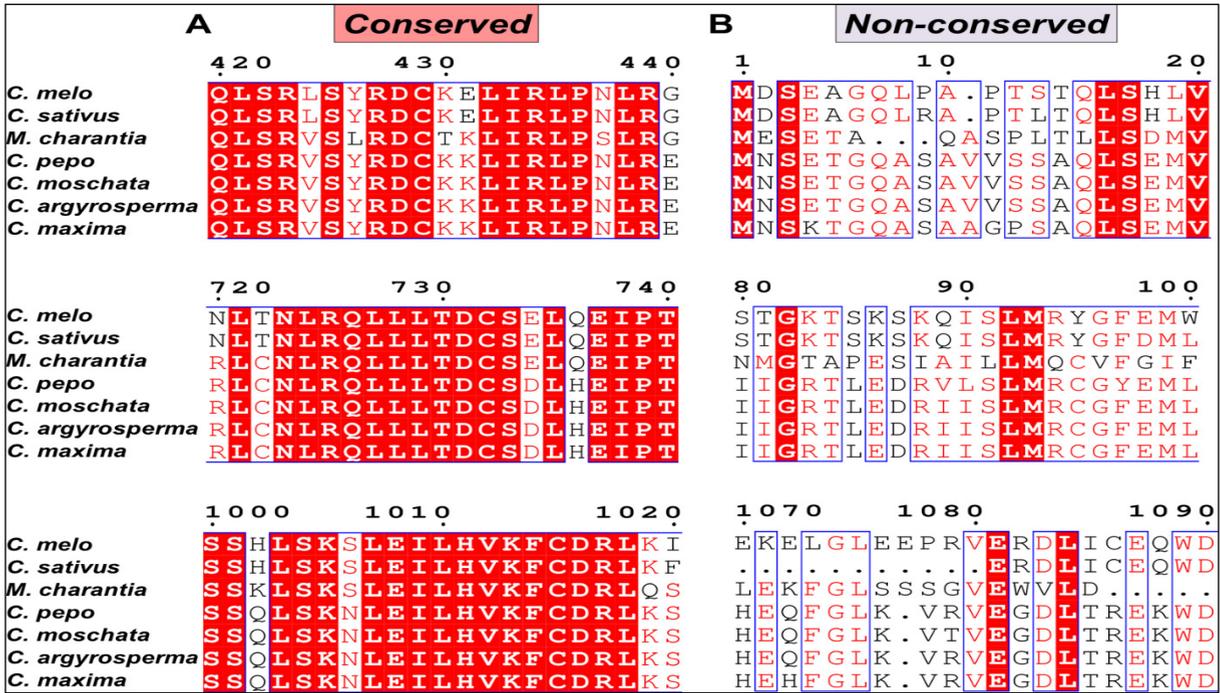


Fig. 2: Multiple sequence alignment for putative powdery mildew resistance protein from *Cucumis melo* L. and its homologous proteins from other members of the family, Cucurbitaceae. The protein sequences were retrieved from the NCBI GenBank (Accession numbers are given in Fig. 1). Only some of the highly conserved (A), and non-conserved regions (B) are represented in this figure due to the large size of the whole alignment. The amino acid residue numbers indicated on top are corresponding to the protein sequence of *Cucumis melo* var. *makuwa* and the sequence similarity is highlighted in red. Clustal Omega was used for the initial generation of the alignment which was further rendered by ESPript for visualization

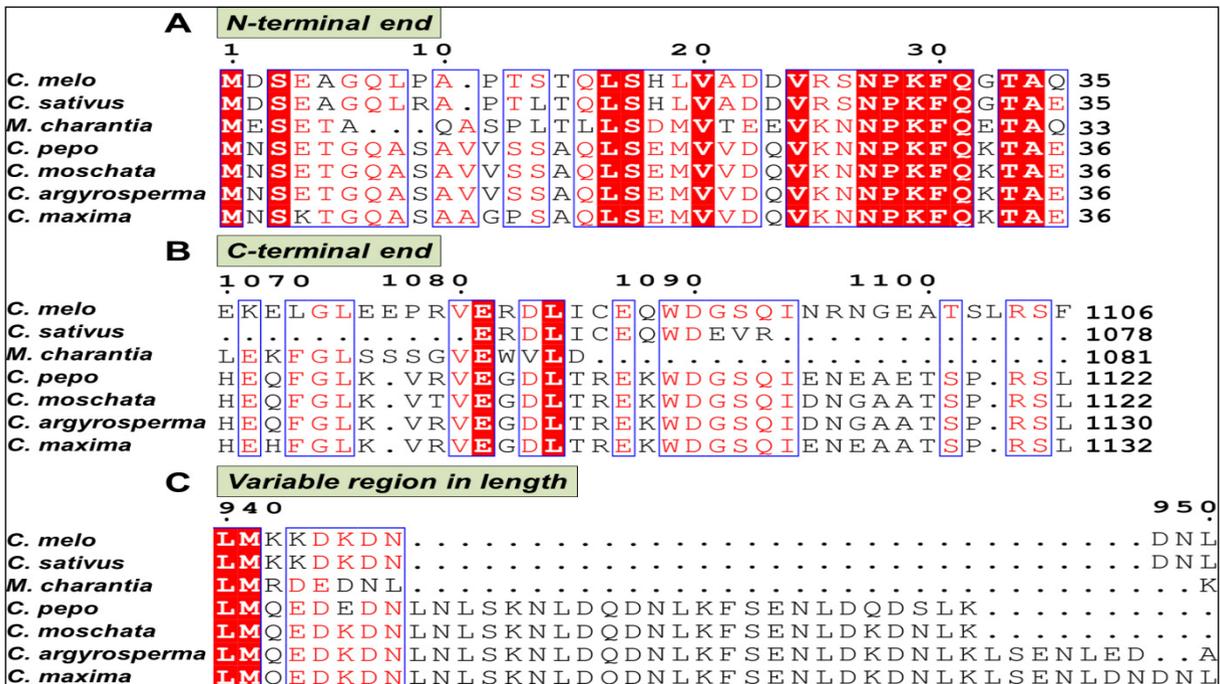


Fig. 3: Multiple sequence alignment at the N-terminal end (A), C-terminal end (B) and variable region in length (C), for putative powdery mildew resistance protein from *Cucumis melo* L. and its homologous proteins from other members of the family, Cucurbitaceae. The protein sequences were retrieved from the NCBI GenBank (Accession numbers are given in Fig. 1). The amino acid residue numbers indicated on top are corresponding to the protein sequence of *Cucumis melo* var. *makuwa* whereas numbers at the side of the alignment in A and B are corresponding to the given respective member of the Cucurbitaceae family. The sequence similarity is highlighted in red. Clustal Omega was used for the initial generation of the alignment which was further rendered by ESPript for visualization

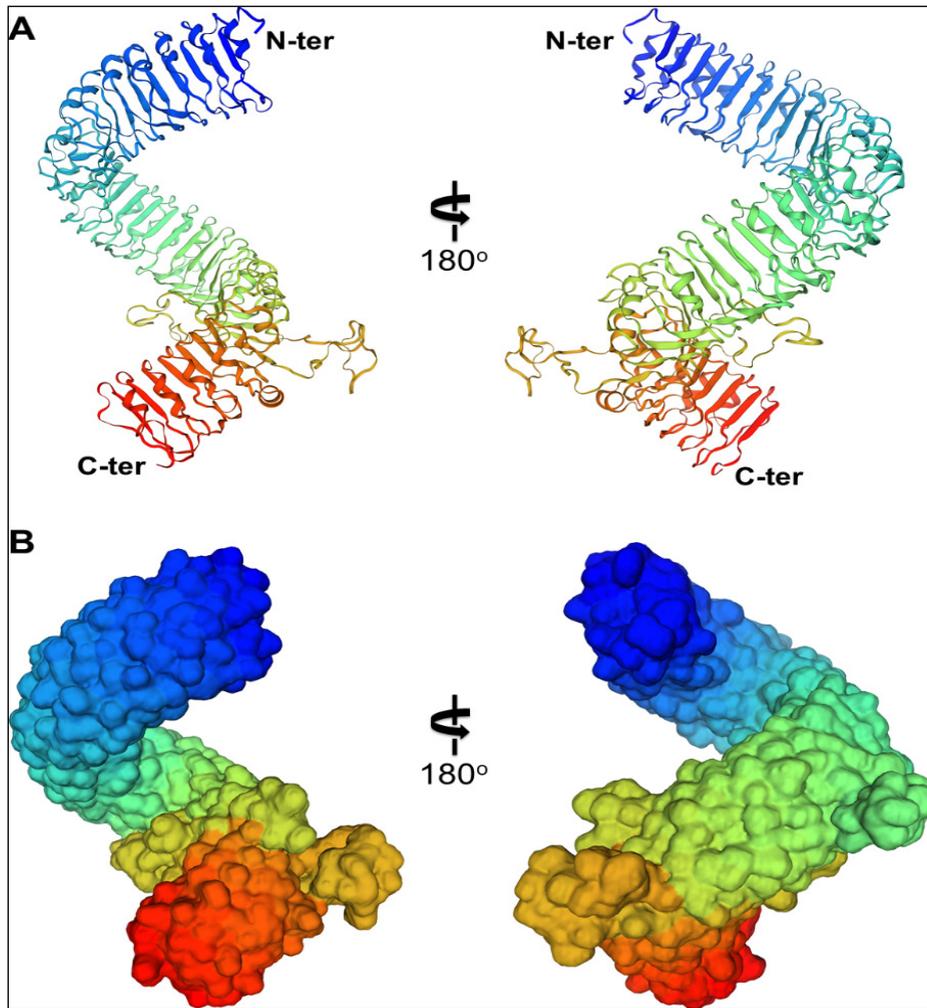


Fig. 4: Homology model for the three-dimensional structure of putative powdery mildew resistance protein from *Cucumis melo* var. *makuwa* (NCBI GenBank accession: KAA0043493). SWISS-MODEL was used to build this homology model against the predicted high sequence similarity template, the X-ray crystal structure of the chain A of Leucine-Rich Repeat (LRR) ectodomain of the plant membrane receptor kinase GASSHO1/SCHENGEN3 from *Arabidopsis thaliana* (Protein Data Bank ID: 6S6Q). Cartoon (A) and surface (B) representations of the homology model with the rainbow color scheme (N to C-terminal end) were generated using the NGL viewer of the SWISS-MODEL web interface

used the SWISS-MODEL program to search for appropriate templates for generating the *CmPPMRP* model. The program displayed the chain A of crystal structure of LRR ectodomain of the plant membrane receptor kinase GASSHO1/SCHENGEN3 from *Arabidopsis thaliana* (Protein Data Bank ID: 6S6Q; Okuda *et al.* 2020) with the sequence identity of 22%, a better value than the other predicted templates by the program. Therefore, we then used this template to build the homology model for *CmPPMRP*. Surprisingly, the 3D structure model of *CmPPMRP* showed the existence of Leucine-Rich Repeat (LRR), a protein fold that is highly conserved among the cell surface receptors of the plant immune system (Fig. 4). Among the known cell surface receptors, we anticipate that *CmPPMRP* is probably structurally

similar to LRR-RLKs. The LRR-RLKs involve in direct recognition of external pathogens and elicit the immune response in association with a co-receptor for signal transmission (Bentham *et al.* 2020).

The *CmPPMRP* homology model covers the amino acid residues from Glu227 at the N-terminus to Lys1064 at the C-terminus. These residues fold into an LRR domain of 31 repeats with four short insertion segments (Fig. 4). Overall, it represents a super helical assembly as previously seen for known plant LRR-RLKs (Hohmann *et al.* 2017).

CONCLUSION

This study carried out the prediction of possible powdery mildew resistance protein existing in



Cucumis melo L. based on the protein sequence information of known powdery mildew resistance protein from wheat, and then performed the phylogenetic and sequence analyses of *C. melo* putative powdery mildew resistance protein. The results of this study showed that the *C. melo* putative powdery mildew resistance protein was diverged from the other members of the Cucurbitaceae family as seen in the phylogenetic tree. Multiple sequence analysis revealed the conserved and non-conserved amino acid residues among the homologous proteins. Homology model showed the conservation of Leucine-Rich Repeat (LRR), a key structural feature of cell surface proteins that belong to the plant immune system. Overall, this study provides valuable insights into the conservation of a putative powdery mildew resistance protein from *C. melo* and helps its future investigations aimed at biochemical and functional characterization.

ACKNOWLEDGEMENTS

Dr. M. Komala and Dr. K. Gopal acknowledge Dr. Y.S.R. Horticultural University, Andhra Pradesh for providing all the infrastructure and research facilities. Dr. M. Komala would like to thank Dr. Tolety Janakiram (Vice-Chancellor, Dr. YSRHU), Dr. K. Gopal (Associate Dean, COH Anantharajupeta) and Dr. S. Sree Vijaya Padma (Associate Dean, COH Chinalataripi) for their support and encouragement. Dr. K. Gopal (Associate Dean, COH Anantharajupeta) would like to thank Dr. Tolety Janakiram (Vice-Chancellor, Dr. YSRHU). Dr. P. Latha and Dr. N. Premalatha acknowledge Tamil Nadu Agricultural University for providing all research facilities and also thank Dr. V. Geethalakshmi (Vice-Chancellor, TNAU).

REFERENCES

Alengebawy, A., Abdelkhalik, S.T., Qureshi, S.R. and Wang, M.Q. 2021. Heavy metals and pesticides toxicity in agricultural soil and plants: ecological risks and human health implications. *Toxics*, **9**(3): 42.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**: 3389–3402.

Bentham, A.R., De la Concepcion, J.C., Mukhi, N., Zdrzałek, R. Draeger, M., Gorenkin, D., Hughes, R.K. and Banfield, M.J. 2020. A molecular roadmap to the plant immune system. *J. Biol. Chem.*, **295**(44): 14916–14935.

Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J. and Sayers, E.W. 2015. GenBank. *Nucleic Acids Res.*, **44**(D1): D67-72.

Cui, H., Fan, C., Ding, Z., Wang, X., Tang, L., Bi, Y., Luan, F. and Gao, P. 2022. *CmPMR1* and *CmPMrs* are responsible for resistance to powdery mildew caused by *Podosphaera xanthii* race 1 in melon. *Theor. Appl. Genet.*, **135**: 1209–1222.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J. and Lopez, R. 2010. A new bioinformatics analysis tools framework at EMBL-EMBL-EBI. *Nucleic Acids Res.*, **38**: 695–699.

Hohmann, U., Lau, K. and Hothorn, M. 2017. The structural basis of ligand perception and signal activation by receptor kinases. *Annu. Rev. Plant Biol.*, **68**: 109-137.

Jones, J.D.G. and Dangl, J.L. 2006. The plant immune system. *Nature*, **444**: 323–329.

Jones, J.D.G., Vance, R.E. and Dangl, J.L. 2016. Intracellular innate immune surveillance devices in plants and animals. *Science*, **354**: aaf6395.

Kesh, H. and Kaushik, P. 2021. Advances in melon (*Cucumis melo* L.) breeding: An update. *Sci. Hort.*, **282**: 110045.

Komala, M. and Kuni, P. 2022. Genetic diversity and molecular breeding of melon (*Cucumis melo* L.): A Review. *Curr. Agric. Res. J.*, **10**(3): 181–192.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. and Higgins, D.G. 2007. ClustalW and ClustalX version 2. *Bioinformatics*, **23**(21): 2947–2948.

Liang, Q., Wei, L., Xu, B., Liu, J., Zhang, S. *et al.* 2022. Induction of resistance of *Podosphaera xanthii* (hull-less pumpkin powdery mildew) to triazole fungicides and its resistance mechanism. *PLoS One*, **17**(2): e0263068.

Madeira, F., Pearce, M., Tivey, A.R.N. *et al.* 2022. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.*, 2022 Apr: gkac240.

Maharjan, S., Upadhyay, N. and Bhattarai, A.M. 2020. Pesticides use pattern and handling practices among cucumber (*Cucumis sativus* L.) growing farmers of Parsa district of Nepal. *Int. J. Agric. Environ. Biotechnol.*, **13**(2): 155–162.

Mieslerova, B., Cook, R.T.A., Wheeler, C.P. and Lebeda, A. 2022. Ecology of powdery mildews – influence of abiotic factors on their development and epidemiology. *Crit. Rev. Plant Sci.*, **41**: 365–390.

Okuda, S., Fujita, S., Moretti, A., Hohmann, U., Doblaz, V.G., Ma, Y., Pfister, A., Brandt, B., Geldner, N. and Hothorn, M. 2020. Molecular mechanism for the recognition of sequence-divergent CIF peptides by the plant receptor kinases GSO1/SGN3 and GSO2. *Proc. Natl. Acad. Sci. U S A.*, **117**(5): 2693–2703.

Robert, X. and Gouet, P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.*, **42**(W1): 320–324.



Srichumpa, P., Brunner, S., Keller, B. and Yahiaoui, N. 2005. Allelic series of four powdery mildew resistance genes at the *Pm3* locus in hexaploid bread wheat. *Plant Physiol.*, **139**(2): 885–895.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R. and Schwede, T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.*, **46**: 296-303.

Zitter, T.A., Hopkins, D.L. and Thomas, C.E. 1996. Compendium of cucurbits diseases. APS Press, Saint Paul.