Int.J.Curr.Microbiol.App.Sci (2019) *8*(1): 790-798

International Journal of Current Microbiology and Applied Sciences
***ISSN: 2319-7706* Volume 8 Number 01 (2019)**
Journal homepage: http://www.ijcmas.com

Original Research Article https://doi.org/10.20546/ijcmas.2019.801.086

Characterization and Molecular Modelling of Pi56 Ortholog from *Oryza rufipogon*

Deepak V. Pawar[1*], Pawan Mainkar[1], Ashish Marathe[2], Rakesh Kumar Prajapat[1], Tilak R. Sharma[3] and Nagendra K. Singh[1]

[1]ICAR-National Research Centre on Plant Biotechnology, Pusa Campus, New Delhi-110012, India
[2]ICAR-ICAR- National Institute of Biotic Stress Management, Raipur-493225, India
[3]National Agri-Food Biotechnology Institute, Mohali, Punjab-140306, India

**Corresponding author*

Keywords

Oryza rufipogon, Ortholog, *Pi56or*, Rice blast, Phylogeny, NBS-LRR domain

Article Info

Accepted:
07 December 2018
Available Online:
10 January 2019

A B S T R A C T

Rice genome harbours many resistance genes (*R*-genes) with tremendous allelic diversity, constituting a robust immune system effective against microbial pathogens like rice blast fungus *M. oryzae.* Nevertheless, few functional R-genes have been identified for rice blast resistance. Wild species of cultivated plants are treasure trove for important agronomic traits. The wild rice *Oryza rufipogon* is resistant to many virulent strains of *Magnaporthe oryzae.* Although considerable research on characterizing genes involved in biotic stress resistance is accomplished at genomic and transcript level, characterization at proteins level is yet to be explored. In the present study, we report the amplification, sequencing and protein sequence analysis of *Pi56*ortholog (*Pi56or*) in *O. rufipogon* accession WRA21. The *Pi56or* encodes 746 amino acid protein with an isoelectric point of 5.69.Sequence analysis revealed that *Pi56or* shared highest similarity (80%) with *Oryza meridionalis* ortholog. The predicted 3D model confirmed 17 α helices and 18β pleated sheets with ATP-binding site close to β sheet present towards the N-terminus of the protein molecule. The present study using various molecular and bio-computational tools could, therefore, help in improving our understanding of this key resistance protein and prove to be a potential target towards developing resistance to *M. oryzae* in rice.

Introduction

Rice blast disease, caused by the fungus *Magnaportheoryzae*, is one of the most devastating diseases of rice worldwide (Kush and Jena 2009; Liu *et al.*, 2010). The yield losses in rice account for about 20–50 % in the absence of adequate resistance (Savary *et al.*, 2000). Because of the effectiveness of plant *R*-genes in preventing diseases, the incorporation of blast resistance genes into high yielding cultivars has been the most favoured strategy to minimize the yield losses. A majority of the major resistance genes with steady broad-spectrum resistance follow a model of gene-for-gene interaction (Jia *et al.*, 2000).

However, blast resistant varieties of rice when introduced in the disease infected areas succumb to disease within 2-3 years, which necessitates need for genes with broad-spectrum and stable resistance (Bonman *et al.*, 1992).In some cases, the donors of these *R*-genes have not been extensively evaluated in agronomically relevant conditions. In other cases, even when the donors have been extensively tested, *R*-genes such as, Pi3(t), Pi5 and Pi9 fail to confer broad-spectrum resistance to*M. oryzae* when deployed individually (Variar *et al.*, 2009). For practical breeding, increasing emphasis has been placed on identifying sources of broad-spectrum resistance to blast based on various criteria (Jeung *et al.*, 2006).

Molecular cloning came into picture when first disease resistance gene HM1 from maize was isolated (Johal and Briggs, 1992). Till date, more than 100 *R*-genes have been identified in the rice genome but only 24 genes have been cloned and well characterized (Sharma *et al.*, 2012). These cloned and characterized genes include *Pib, Pita, Pi54, Piz-t, Pi5, Pish, Pi-k, Pikm, Pi-9, Pid3, Pid2, pi21, Pit, Pb1, NLS1, Pi25, Pi54rh, Pi2, Pi-37, Pia, Pi-36, Pik-pPid3-A4* and *Pi54of* (Devanna *et al.*, 2014). Pid2 is an exception as it encodes extracellular β-lectin receptor kinase while all other cloned *R*-genes encode intracellular proteins having nucleotide binding site-leucinerich repeat (NBS-LRR) domains that play an important role in imparting disease resistance. The N terminal NBS domain is involved in ATP binding and hydrolysis, while the C terminal LRR is involved in protein-protein interactions (Takken and Tameling, 2009).

Wild species of rice can be a potential target for broad-spectrum resistance genes. Orthologs of major resistance genes can be explored and assayed for their resistance towards *M. oryzae*. For example, the *Pi54*orthologcloned from wild rice *O. rhizomatis* (*Pi54rh)* and *O. officinalis (Pi54of)* confers broad-spectrum resistance against *M. oryzae* (Das *et al.*, 2012). *Pi56* gene characterised from resistant variety Sanhuangzhan No 2 (SHZ-2) confers broad-spectrum resistance to *M. oryzae* (Liu *et al.*, 2013). Wild species of rice like *O. rufipogon*is known to be resistant to *M. oryaze.* Molecular basis of resistance to *M. oryzae* have been well characterised for 24 genes (Devanna *et al.*, 2014). But various physio-chemical properties like size, shape, hydrophilicity and structural features like 3-dimensional configuration, molecular flexibility of a protein determine its functional behaviour in vivo. Physical and enzymatic alterations have been a conventional tool in improving the functionality of a protein and therefore understanding the structural features through various bio-computational tools could provide new avenues to enhance the functionality of a protein at molecular level.

Materials and Methods

Pi56or gene amplification, sequencing and analysis

Genomic DNA was isolated from the leaves of wild species of rice, *Oryza rufipogon* accession WRA21. High quality DNA (100ng/µl) was used in PCR amplification of Pi56or. Two primer pairs were designed to amplify the *Pi56or* region (Table 1).PCR was carried out in thermocycler in a 25 µL reaction volume containing 1X Taq Buffer, 0.4 Units Phusion High-Fidelity DNA polymerase, 2.5 mM MgCl2, 0.2mMdNTP in each tube.

The PCR conditions were as follows: initial denaturation at 95 °C for 5 min, 35 cycles of 95 °C for 1 min, 60.6 °C for 45 s, 72 °C for 90 s; an additional extension at 72 °C for 10 min. The amplicons were gel eluted and sequenced by primer walking. The trace files were base

called, checked for quality of the sequence and trimmed for primer sequences using Phred and assembled to generate consensus sequence using Phrap software tools (Ewing and Green, 1998 and Ewing *et al.,* 1998). Sequences containing at least 100 continuous nucleotides with a Phred score greater than 30 were clustered by Phrap with a minimum consensus Phrap score of 80. The assembled contigs were viewed and edited by using Consed (Gordon *et al.,* 1998). Gene prediction was carried out using FGENESH (http://linux1. softberry.com). The functional domains of lectin were determined using the InterPro tool available on the EBI web page (www.ebi.ac.uk/interpro/).

*Pi56*ortholog sequences were obtained by performing BLAST search against Ensembl genome browser (http://plants.ensembl.org) database using Pi56 sequence as query sequence. 11 orthologs (*O. sativa japonica, O. sativaindica, O. punctata, O_rufipogon, O. nivara, O. meridionalis, O. longistaminata, O, glumipatula, O. glaberrima, and O. barthii*) were obtained. The amino acid sequences all *Pi56*orthologs were used for phylogenetic studies. MEGA (Molecular Evolutionary Genetic Analysis) version 6 software (http://mega soft ware.net/) was implemented for constructing the phylogeny treeusing the Neighbour Joining method.

The physico-chemical properties like amino acid composition, pI, molecular weight, half-life and instability index were determined using Protparam (http://web.expasy.org/ protparam/). Probability of protein disorder was determined by the PrDOS (Protein disorder prediction server) tool (http://prdo s.hgc.jp). The subcellular location and molecular functions of protein were predicted by using CELLO2GO (http://cello.life.nctu.edu.tw/ cello2go/) web server.

Structural analysis and homology-based modelling

The secondary structure and solvent accessibility of Pi56or was determined by the RaptorX protein structure server (http://raptorx.uchicago.edu/StructurePredictio n/predict/). The 3D structure of the target protein Pi56or was generated using SWISS Model tool (https://swissmodel.expasy.org/). The authenticity of the predicted models was further validated employing RAMPAGE tool (http://mordred. bioc.cam.ac.uk/~rapper/ram page.php).

Active site mapping, cleft analysis and molecular docking

The amino acid residues present in the ligand-binding sites were analyzed using FunFold2 server (http://www.read ing.ac.uk/bioinf/FunF OLD/) and I-TASSER (http://zhanglab. ccmb.med.umic h.edu/I-TASS ER/). The cleft analysis to detect the ligand-binding domains of the protein was done using FTSite Server (http://ftsi te.bu.edu/). Docking studies were executed to investigate the probable binding modes of the substrates to the active site of Pi56or, for which, PDB file of the modelled Pi56or was imported into SwissDock module (http://www.swis sdoc k.ch). The docking results were viewed using UCSF Chimera 1.11rc package (www.cgl.ucsf.edu/chim era).

Results and Discussion

Sequence analysis and characterization

Pi56 gene is reported to confer broad spectrum resistance to *M. oryzae* (Liu *et al.,* 2013). We amplified the corresponding *Pi56or*(where "or" stands for *oryzarufipogon*) ortholog from *Oryzarufipogon* accession WRA21. The two amplicons of size 2286 bp and 1544 bp were obtained (Fig. 1). The amplicons were sequenced by primer walking, and gene

prediction was carried out in the assembled contig sequence. Gene prediction revealed that the Open Reading Frame (ORF) of *Pi56or* is 3078 bp which codes for 743 amino acids. Phylogenetic analysis of Pi56or with other orthologs was performed with 11 orthologs of Pi56, four main clusters were observed for the Pi56 orthologs in Cluster I contains *O. nivara, O. Sativaindica, O. glumipatula, O rufipogon,O. meridionalis* and *O. sativa japonica*orthologs, in cluster II *O. rufipogon* WRA21 and *O. longidtaminata* orthologs, in cluster III *O. barthii* and *O.glaberrima*. Cluster IV contained single *O. punctate* ortholog was clustered (Fig. 2).

The functional domain of Pi56or protein were defined using InterPro tool (Fig. 3). The Pi56or contains P-loop containing nucleoside triphosphate hydrolase domain (from 117th to 330th amino acid), Leucine-rich repeat (LRR) domain (from 498th to 743rd amino acid), and nucleotide binding Domain (NB-ARC) from 123rd to 310th amino acid. The two domains NB-ARC and LRR are typical characteristic of *R*-genes. Out of 24 genes cloned and characterized proteins of blast resistance genes, nine proteins have been predicted to belong to the NBS-LRR type whereas thirteen proteins are of CC-NBS-LRR class. The *Pid-2* protein is a unique type of β-lectin receptor having Serine Threonine Kinase (STK) type domain and *pi21* is a non NBS-LRR protein, and encodes a proline rich heavy metal binding protein and a protein-protein interaction motif (Fukuoka *et al.*, 2009).

Table.1 Primers used for Pi56or gene amplification

Primer	Forward primer (5' to 3')	Reverse primer (5' to 3')	Amplicon size (bp)
Pi56_Seq_ 1	ATGGCGGGGAAAGCGACCGC	CAAGTTTCCATGTCTTGATT	2286
Pi56_Seq_ 2	AATCAAGACATGGAAACTTG	CTATGAGTTCACTATGTGGAGGC	1544

Fig.1 PCR amplification of Pi56or gene. Lane nos. 1–2 show amplicons 2286 and 1544 bp respectively; M-Molecular weight marker (1 kbDNA ladder)

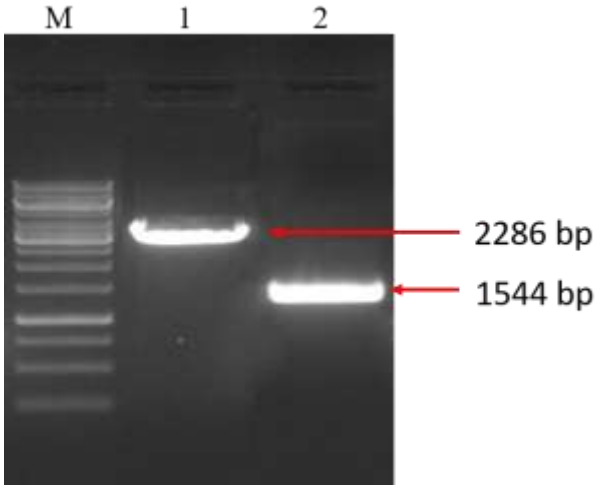


Fig.2 Phylogenetic tree showing the evolutionary relatedness of Pi56or with other Pi56orthologs

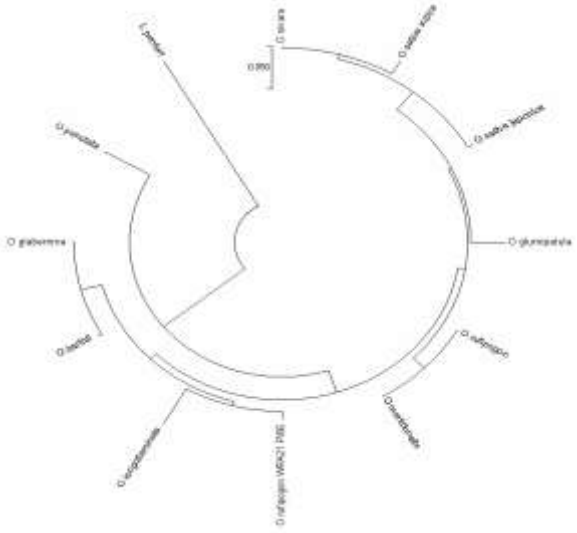


Fig.3 Functional domains analysis of Pi56or protein sequence

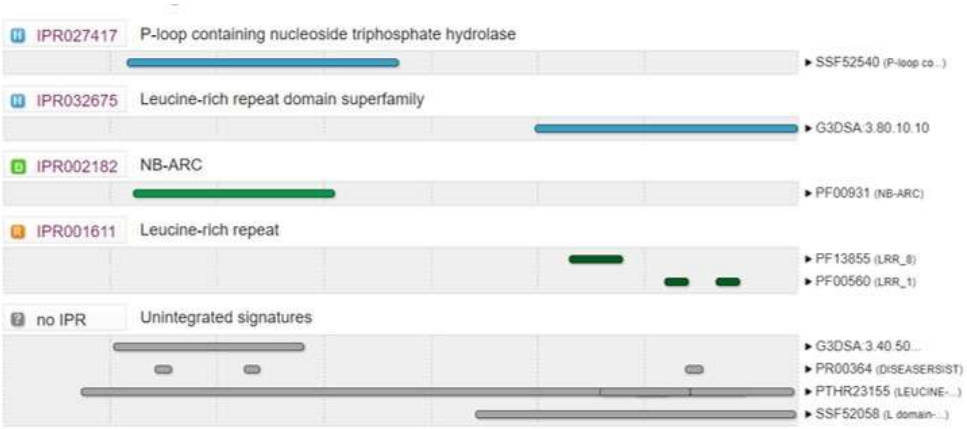


Fig.4 Prediction of the disordered amino acid residues present in Pi56or protein (shown in red) using PrDOS tool

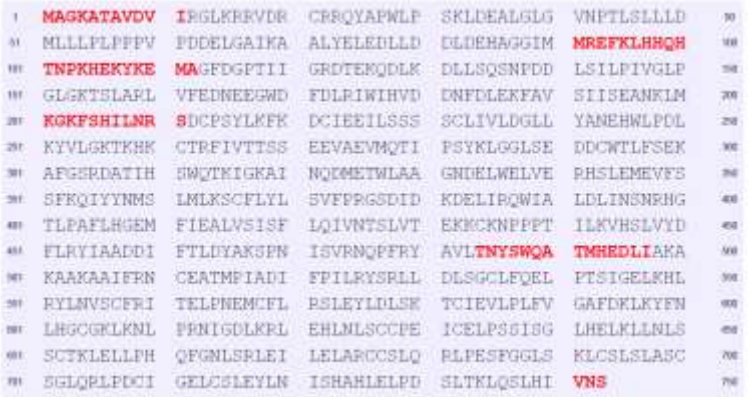


Fig.5 3D model of Pi56or generated via homology-based modelling using SWISS MODEL depicting various secondary structures—α helices, β pleated sheets and random coils

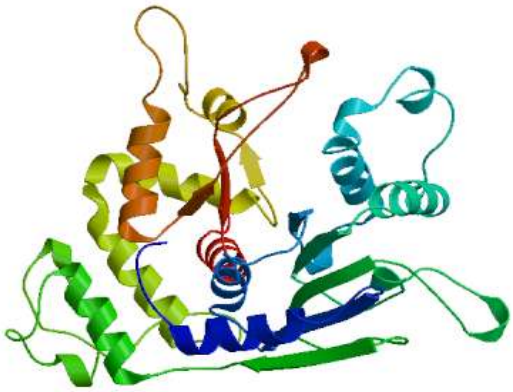


Fig.6 Validation of 3D predicted structure using RAMPAGE

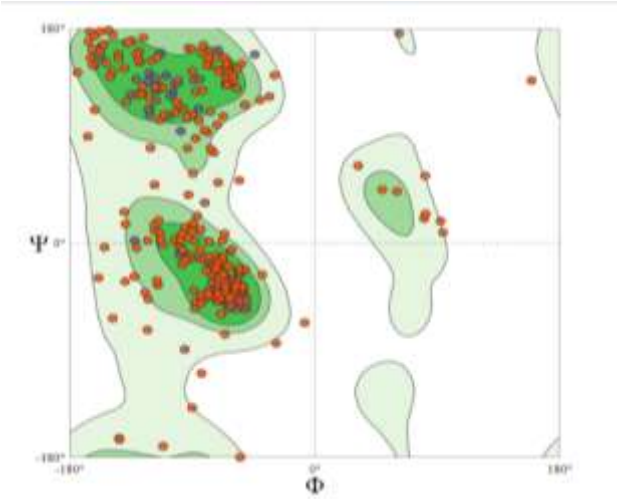


Fig.7 Schematic representation of the secondary structure prediction of Pi56or using PDBSumtool. Arrows (Pink) indicating the β pleated sheets and Barrels (Red) indicating the α Helices

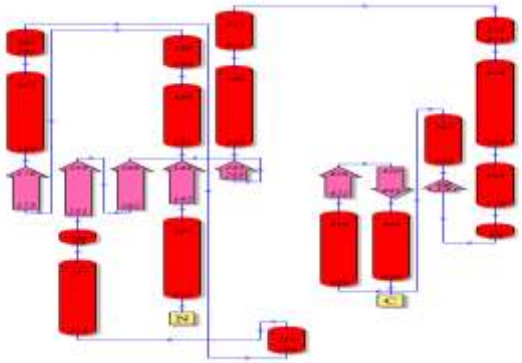
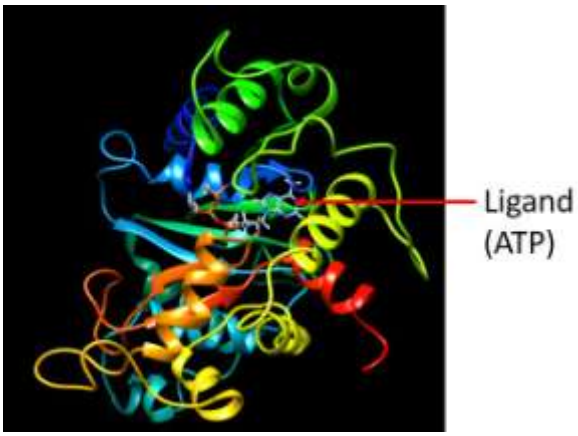


Fig.8 Representative binding mode of ATP at the active site of Pi56or subsequent to docking simulation using Swiss-Dock



The Pi56or is characterized as acidic protein based on computed pI value 5.69 (pI<7). Protparam analysis indicated that the molecular weight of Pi56or is 84063.36 Da. The analysis revealed Leucine as the most abundant amino acid in Pi56or, accounting for about 16.2%, while Tryptophan, Tyrosine and Glutamine were the least abundant. The CELLO2GO tools revealed that the Pi56or is localised in cytoplasm with a reliability score of 1.450. Previously three blast resistance genes (*Pi37, Pi21 and Pita*) have been reported to be localised in the cytoplasm. The Pi56or protein plays important molecular functions in conferring resistance to *M. oryzae*, by binding with AVR in gene for gene manner. The instability index of Pi56or was 44.45, classifying it as an unstable protein. Five disordered regions were predicted in the protein sequence, of which the longest disordered region was found between Met90 to Ala112 comprising 22 amino acid residues (Fig. 4). GRAVY indices for Pi56or was -0.064, indicates the possibility of better interaction with water i.e. hydrophilic nature of the protein which is attributed to charged amino acid residues present in the protein sequence (98 negatively charged and 79 positively charged), suggesting that Pi56or might be present in cytoplasm. The estimated half-life of Pi56or is about 30 hours. The Aliphatic index of the Pi56 in the present study was103.97 to 105.1. High Ai of these proteins indicates a higher thermostability of the protein and is predicted to play a role in response to various biotic and abiotic stresses. This attribute can possibly be explored in studies pertaining to cell signalling under biotic and abiotic stress.

Homology modelling and structural characterization

The homology model of Pi56or was generated employing, SWISSMODEL server (Fig. 5). Six models were generated in total, of which the best model chosen was based on high resolution and per cent coverage. The sequence identity score was 38.95 with a resolution of 2.80 A°. The 3D model generated was further validated using RAMPAGE program (Fig. 6). The torsion angles, ψ and ϕ were examined to access the reliability of the protein model. The results obtained in the validation, 88.0% of the amino acid residues were found in the most favoured region, while 9.0 and 3.0% of the amino acid residues were found in the allowed region and the outlier region, respectively. The secondary structure generated with PDBsum predicted a total of 17 α helices (34.29%), 8 β pleated sheets (23.17%) (Fig. 7).

Active site mapping and molecular docking

Active site mapping for determining the residues involved in binding to the ATP ligand was done using I-TASSER and Funfold server. The residues Glu125, Arg159, Glu195, Ala242, Gln248, Val253, Val273, Ile280, Glu299, Phe382, and Ile309 bind to ATP with highest C-score of 0.75 (Fig. 8), the ATP-binding site is located in the β sheet close to the N-terminus (amino acid residue 108–330).

The presence of His in the substrate-binding site among the proteins is predicted to play a role in inter-substrate phosphate transfer (Chamberlian *et al.,* 2007). The presence of conserved amino acid residues in the nucleotide-binding and *Avr* (avirulence) protein binding domain of LRRs indicates the potential for protein engineering and altering the protein activity via targeted mutagenesis. Rice, being a target for functional genomics, such in silico protein models can prove beneficial in predicting their role prior to protein engineering.

Acknowledgement

DVP, RKP and PM are thankful to DBT (Department of Biotechnology), Govt. of India, AM is thankful to DBT (Department of Science and Technology) Govt. of India, and TRS and NKS are thankful to ICAR-NPTC for providing the financial assistance.

References

Bonman, J. M., 1992. Durable resistance to rice blast disease-environmental influences. *Euphytica* 63: 115–123.

Chamberlian, P, P., Qian, X., Stiles, A, R., Cho, J., Jones, D. H., Lesley, S. A., Grabau, E. A., Shears, S. B., Spraggon, G., 2007. Integration of inositol phosphate signalling pathways via Human ITPK1.*J BiolChem* 21:28117–28125.

Das, A., Soubam, D., Singh, P. K., Thakur, S., Singh, N. K., Sharma, T. R., 2012. A novel blast resistance gene, Pi54rh cloned from wild species of rice, *Oryza rhizomatis* confers broad spectrum resistance to *Magnaporthe oryzae*. *Functional and Integrative Genomics*, 12, 215–228. doi:10.1007/s10142-012-0284-1.

Devanna, N. B., Vijayan, J., Sharma, T. R., 2014. The blast resistance gene Pi54of cloned from *Oryza officinalis* interacts with Avr-Pi54 through its novel non-LRR domains. *PLoS ONE*, 9.doi: 10.1371/journal.pone.0104840.

Ewing, B. and Green, P. 1998. Base calling sequencer traces using *Phred*II. Errorprobabilities. *Genome Res*. 8: 186-194.

Ewing, B., Hillier, L., Wendl, M. and Green, P. 1998. Base calling sequencer tracesusing *Phred* I. Accuracy assessment. *Genome Res.* 8: 175-185.

Fukuoka, S., Saka, N., Koga, H., Ono, K., Shimizu, T., Ebana, K., Yano, M., 2009. Loss of function of a proline-containing protein confers durable disease resistance in rice. *Science* (New York, N.Y.), 325, 998–1001. doi:10.1126/science.1175550.

Gordon, D., Abajian, C. and Green, P. 1998. *Consed*: A graphical tool for sequence finishing. *Genome Res*. 8: 195-202.

Jeung, J. U., Heu, S. G., Shin, M. S., Vera Cruz C. M., Jena, K. K., 2006. Dynamics of *Xanthomonasoryzae*pv. *oryzae* populations in Korea and their relationship to known bacterial blight resistance genes. *Phytopathology*. 96(8): 867-75. doi: 10.1094/PHYTO-96-0867.

Jia, Y. L., McAdams, S. A., Bryan, G. T., Hershey, H. P., Valent, B., 2000. Direct interaction of resistance gene and

*avirulence*gene products confers rice blast resistance. *EMBO J* 19:4004–4014.

Johal, G. S., and Briggs, S. P., 1992. Reductase activity encoded by the HM1 disease resistance gene in maize. *Science* (New York, N.Y.), 258, 985–987. doi:10.1126/science.1359642.

Kush, G. S. and Jena, K. K., 2009. Current status and future prospects for research on blast resistance in rice (*oryza sativa* L.). In: Wang GL, Valent B (eds) Advances in genetics, genomics and control of rice blast disease. *Springe*r, Dordrencht, pp 1–10.

Liu, J. L., Wang, X. J., Thomas. M., Hu, Y. J., Liu, X. L., Dai, L.Y., Wang, G. L., 2010.Recent progress and understanding of the molecular mechanisms of the rice—*Magnaporthe* interaction. *Mol Plant Pathol*, 11(3):419–427

Liu., Y., Liu, B., Zhu, X., Yang, J., Bordeos, A., Wang, G., Leach, J. E., Leung, H., 2013. Fine-mapping and molecular marker development for Pi56(t), a NBS-LRR gene conferring broad-spectrum resistance to *Magnaportheoryzae* in rice. *TheorAppl Genet*, 126:985–998.

Savary, S., Willocquet, L., Elazegui, F. A., Castilla, N. P., Teng, P. S., 2000. Rice pest constraints in tropical Asia: quantification of yield losses due to rice pests in a range of production situations. *Plant Dis*, 84:357–369.

Sharma, T. R., Rai, A. K., Gupta, S. K., Vijayan, J., Devanna, B. N., Ray, S., 2012. Rice Blast Management through Host-Plant Resistance: Retrospect and Prospects. *Agricultural Research*, 37–52. doi:10.1007/s40003-011-0003-5.

Takken, F. L. W. and Tameling, W. I. L., 2009. To nibble at plant resistance proteins. *Science* (New York, N.Y.), 324, 744–746. doi:10.1126/science. 1171666.

Variar, M. Vera. C. C. M., Carrillo, M. G., Bhatt, J. C., Sangar, R. B.S. 2009.Rice blast in India and strategies to develop durably resistant cultivars. Advances in genetics, genomics and control of rice blast disease. *Springer*; New York: 2009. pp. 359–374.

How to cite this article:

Deepak V. Pawar, Pawan Mainkar, Ashish Marathe, Rakesh Kumar Prajapat, Tilak R. Sharma and Nagendra K. Singh. 2019. Characterization and Molecular Modelling of Pi56 Ortholog from *Oryza rufipogon. Int.J.Curr.Microbiol.App.Sci.* 8(01): 790-798.
doi: https://doi.org/10.20546/ijcmas.2019.801.086