Pol. J. Environ. Stud. Vol. XX, No. X (XXXX), 1-11
DOI: *10.15244/pjoes/195743* ONLINE PUBLICATION DATE: 2025-04-09

Original Research

In-Silico Analysis of R-Genes in Rice for New Insights of Partial and Complete Resistance Against Bacterial Leaf Blight Disease

Areeqa Shamshad[1], Muhammad Rashid[1], Somia Noor[1], Salman Alamery[2], Kotb A. Attia[2], Yaser M Hafez[3], Shahwar Fatima[1], Muzna Shahid[4], Aamir Riaz[5], Sajid Fiaz[6]*, Qamar Uz Zaman[4]**

[1]Nuclear Institute for Agriculture & Biology (NIAB-C), PIEAS, Pakistan
[2]Department of Biochemistry, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia
[3]EPCRS Excellence Center, Plant Pathology and Biotechnology Lab.; Agric. Botany Dept.; Fac. Agric.; Kafrelsheikh University; 33516, Egypt
[4]The University of Lahore, Department of Environmental Sciences, Lahore, 54590, Pakistan
[5]National Key Laboratory of Green Pesticide, Key Laboratory of Green Pesticide and Agricultural Bioengineering, Ministry of Education, Center for R&D of Fine Chemicals of Guizhou University, Guiyang 550025, China
[6]Institute of Molecular Biology and Biotechnology, The University of Lahore, 54590 Lahore, Pakistan

Received: 9 June 2024
Accepted: 8 November 2024

Abstract

Bacterial leaf blight (BLB) disease is a rice disease caused by *Xanthomonas oryzae* pv.*oryzae* (*Xoo*). Under variable climate conditions, the new races of *Xanthomonas oryzae* pv. *oryzae* (Xoo) have a massive impact on rice crop yield. The focus of the study has been finding disease-resistant genes that will provide resistance against *Xoo*. Different strains of Xoo attack rice crops all over the world; the kinds that are resistant to these strains can differ depending on where in the world you live.

Currently, 45 *Xa*-genes have been identified that possess resistance against the *Xoo* pathogen and are being used in screening rice-diverse cultivars. In this project, amino acid sequences of *Xa23*, *Xa23Ni*, *Xa10,* and *X10Ni* were mined and used for their motif, domain identification, and chromosome localization. Almost 113 amino acids in both the *Xa23* and *Xa23-Ni* were observed. However, 126 and 134 amino acids were identified in *XaXa10* and *X10-Ni*. The amino acid-conserved regions of *Xa23* contribute 50% of the identity to the known executor R protein of *Xa10*. In addition, trans-membrane helices of *Xa23* were also present in *Xa10*. However, the activation of transcription in *Xa23* observed by Avr*Xa23* was distinct from that in *Xa10*. During *in-silico* analysis, two novel *Xa10*-like genes (*Xa10-Ni* and *Xa23*-Ni) were recognized. The structural analysis (physical and chemical characterization) of the *Xa23*, *Xa23*Ni, *Xa10,* and *X10Ni* proteins showed that, except for all the genes, *Xa10*-Ni has α-helix and was slightly acidic in nature. Due to their hydrophobic attributes, all were found to be

*e-mail: sajidfiaz50@yahoo.com
**e-mail: qamar.zaman1@envs.uol.edu.pk

stable proteins. In Ramachandran plot analysis, over 90% of amino acid residues fall in the favored regions. In homology modeling, except for *Xa10*-Ni, the predicted models of *Xa23*, *Xa23*-Ni, and *Xa10* possess the ligands. During chromosomal localization, all the genes were found on Chromosome No. 11. It is concluded that the project study will be helpful for the selection of R-genes to build resistance in rice varieties against particular isolates of the Xoo strain to achieve sustainable rice crop production under variable climate conditions.

Keywords*:* rice, Phylogenetic analysis, comparative modeling, Ramachandran plot

Introduction

Rice belongs to the genus *Oryza* (Family: Poaceae) and is a semiaquatic annual grass. It has 22 species and originated about 130 million years ago in the supercontinent Gondwanaland. In 2000, the shotgun sequencing method was used to sequence the whole genome of rice. This method concluded that rice has 12 chromosomes (2n = 24) and about 40,000 genes [1, 2]. About 40,000 rice varieties have been reported and are being cultivated in more than a hundred countries, covering approximately 158 million hectares with a production of 503.17 million metric tons in 2021 [3]. According to USDA reports, an increase of 6.77 million tons, or 1.36%, in rice production around the globe was reported in 2021. About 90% of the global production of rice is harvested in Asia [4, 5]. In Pakistan, rice is the 2nd cereal crop after wheat. Due to good gain quality attributes like aroma and cooking quality, the Super Basmati rice variety is famous among consumers, farmers, traders, and exporters [6]. Under changing climate conditions, there is a high chance of an attack from different diseases on rice, which can affect its global production by approximately 10-29% [7, 8]. Among biotic stresses, bacterial blight (BB) is the major disease in rice caused by a bacterium named *Xanthomonas oryzae* pv. oryzae, a vascular pathogen [9], which results in yield loss that may vary from 20-80% depending upon the severity of conditions, and this loss is more visible in hybrid rice [10]. The symptoms of BB (leaf blight, or "Kresek", or the acute wilting of young plants) can occur at different growth stages. *Zoysia japonica*, *Panicum maximum*, *Cynodon dactylon*, *Leersia oryzoides*, *Cyperus difformis*, and *Cyperus rotundus,* etc., act as hosts of Xoo other than *Oryza* [11]. *Xanthomonas* releases a type-III secretion system (T3SS) directly into plant nuclei and binds a particular DNA sequence called transcription activator-like effectors (TALEs) on either strand of host target genes [12].

To control this disease, the development of BB-resistant varieties is the most cost-effective strategy. To develop resilient rice varieties, the exploration of novel resistant germplasm and the use of excellent gene resources are critically important [13]. Currently, 45 resistance *Xa* genes that possess resistance to BB have been identified and are in use in the screening of diverse rice [14]. The earlier workers reported that the cloning and pyramiding of 11 R genes exhibited multiple mechanisms of resistance against Xoo isolates. Since the 1980s, a few R genes, such as *Xa3*, *Xa4*, *Xa7*, and *Xa21,* have been used widely in rice breeding programs [15]. However, the resistance provided by these genes has decreased under variable climate conditions, and it's no longer tolerant of strong or evolved pathogenic variation [16]. The Avr genes induced resistance to Xoo and may be affected by the genetic makeup of rice or the insufficient dominance of R genes [17, 18]. However, the resistance gene *Xa23,* which is a new executor and derived from wild rice, exhibited wide dominance and broad-spectrum resistance against BB of rice at all stages of growth [19].

In Pakistan, all the cultivated varieties are susceptible to BLB disease, and farmers use different cultural practices to save the rice crop [20]. Besides management practices, the only sustainable solution is the development of resistant cultivars. Under the current scenario of climate change, the use of integrated tools like omics tools, breeding, and bioinformatics approaches may play an important role in the development of rice cultivars exhibiting resistance to the most virulent pathogens of BLB disease (*Xoo*) [21].

In recent years, different NILs and pyramids were released by the *IRRI* [22]. The *Xa* genes encode particular proteins and have diverse domains in their sequences that are useful for bacterial resistance. Because of both sequence and functional similarities, the majority of the gene's regions are identical and possess the same types of domains. Researchers used various ways to explain *Xa* genes, but the full characterization of numerous genes is still incomplete. Earlier workers used wet lab techniques; now, there is a demand for dry lab, i.e., in-silico analysis, to fully explain the *Xa* genes using various computational tools to escalate the resistance ability of these genes up to 100% [23].

The objective of the study is the sequence and structural analysis of BLB resistance *Xa* genes to understand and increase its effect on rice plants and make useful information that will be supportive of future research. It includes motifs and domain analysis of conserved regions, secondary structure analysis, phylogenetic analysis, homology modeling, Ramachandran plot analysis, comparative analysis, and chromosome localization using different window-based and web-based tools. The project study will strengthen the knowledge and skills of researchers by having new resistant *Xa* genes. Incorporating new genes into the

rice breeding program may be used to screen early generations against BLB disease.

Materials and Methods

Sequence Retrieval and Purification

The amino acid sequences of the rice resistance *Xa* gene were derived from the UniProtKB database (www.uniprot.org/help/uniprotkb/). The Rice Genome Annotation Project Database and Resource (http://rice.plantbiology.msu.edu) was used to browse sequences of *Xa* genes for conformation, genomic linkages, and annotation with other *Xa* genes. The BLAST Aligning Tool for Multiple Protein Sequences was used (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) to remove redundant sequences based on their similarity to other amino acid sequences. The EMBOSS Transeq tool (https://www.ebi.ac.uk/Tools/st/emboss_transeq/) was used to translate nucleotide sequences to amino acid sequences.

Sequence Analysis

Motif and Domain Identification

Multiple Sequence Alignment (MSA) of *Xa* amino acid sequences was accomplished by CLUSTAL-X, which works by applying a progressive alignment method with a heuristic approach [24]. The MEME suite tool (http://memesuite.org/tools/meme) was utilized to find the corresponding motifs, which led to their location identification and graphical representation by the sequence logo [25]. The Hidden Markov Models for protein domain identification (SMART) were used to generate detailed results for the domains (http://smart.embl-heidelberg.de), and CDD was used for domain identification (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and to purify protein sequences [26]. The Pfam database (https://pfam.xfam.org) was used to confirm the presence of the resultant domains in the sequences. InterProScan was applied to classify protein families based on (https://www.ebi.ac.uk/interpro/search/sequence) motifs, conservancy, and domain presence in the sequences.

Phylogenetic Analysis

The MEGA-X (https://www.megasoftware.net/) was used to achieve the MSA of *Xa* gene amino acids. A phylogenetic tree was generated with the FastTree program using (http://www.microbesonline.org/fasttree/#:~:text=FastTree%20infers%20approximately%2Dmaximum%2Dlikelihood, PhyML%203.0%20or%20RAxML%207). Finally, the ITol (https://itol.embl.de/) was used for phylogenetic tree annotation, editing of tree data such as cluster names, text font, tree style, bootstrap value, evolutionary distance, merging clusters into clades, and relevant manipulations.

Primary Structure Analysis

The ProtParam tool https://web.expassy.org/protparam was used to compute the physical and chemical parameters beneficial for the structural analysis of *Xa* proteins. The computed parameters comprise the number of amino acids, molecular weight formula, theoretical pI, number of negatively charged residues, number of positively charged residues, total number of atoms, extinction coefficient, aliphatic index, GRAVY, and instability index. Trans-membrane helices in *Xa* amino acid sequences were calculated by applying the TMHMM tool (http://www.cbs.dtu.dk/services/TMHMM), which provides an approach to the hidden Markov model using membrane protein topology identification.

Secondary Structure Analysis

The SOPMA tool was used (https://npsa-prabi.ibcp.fr/cgi-bin/npsa-automate.pl?page=/NPSA/npsa_sopma.html) to predict secondary structures in the *Xa* proteins. The alpha helix, beta-turn, and random coil are among the key components of projected secondary structures. The discovery of signal peptides with locations of their cleavage sites in *Xa* amino acid sequences was determined using the SignalP tool that is based on (http://www.cbs.dtu.dk/services/SignalP) recurrent and concurrent architectures and convolutional neural networks with a random option of condition.

Comparative Modeling

Template Searching and Modeling

The residues of the *Xa23*, *Xa23*-Ni, *Xa10*, and *Xa10-Ni* proteins were searched in BLASTP against the protein database (http://www.rcsb.pdb.org) to determine suitable templates for homology modeling. These templates are designed to predict the 3D structures and functions of these *Xa* proteins. For the homology modeling of *Xa* proteins, suitable templates are searched in the I-TASSER (http://zhanglab.ccmb.med.umich.edu/I-TASSER) tool. The templates of *Xa* proteins were selected based on various parameters of BLASTP, such as the highest alignment score and percentage identity, maximum query coverage, and lowest E-value. All of the above steps predicted that 4he8.1.C, a hetero oligomer (monomer) whose C chain is a well-suited template for *Xa23* and *Xa23*-Ni, for *Xa10* and 7npu.1.E, a hetero-6-6-12-mer whose E chain is suitable for *Xa10-Ni* proteins. Comparative modeling is a technique where the sequence alignment of a target is done using a template whose structure existed in the PDB database, as reported earlier [27]. Models were constructed using

the Swiss model (https://swissmodel.expasy.org/). The UCSF Chimera 1.14 was used to visualize these *Xa* models more extensively.

Model Quality Assessment

Different software, such as GNUPLOT Ramachandran plot analysis, was implemented for quality evaluation, solidity, and internal consistency of the *XaXa23*, *Xa23*-Ni, *Xa10*, and *Xa10-Ni* protein models. The phi and psi angles of proteins were used to check the stereochemical quality of the model, and the Swiss model tool was applied to determine the number of residues in favorable and unfavorable regions of the Ramachandran plot [28]. Overall quality-factor determination of *Xa* proteins was done using the ERRAT (https://servicesn.mbi.ucla.edu/ERRAT) tool. Statistics of unbounded interactions in the various kinds of atoms were calculated by ERRAT. The SAVES tool (http://nihserver.mbi.ucla.edu/SAVES) was applied to figure out the standard bond angles and bond lengths of *Xa* models. The MolProbity web server (http://molprobity.biochem.duke.edu) was used to determine the quality validation of predicted 3D models. Templates from the I-TASSER tool were used to compute the root mean square deviation (RMSD) among the atoms of the predicted models. The Ramachandran plot was made to diagnose the updated dihedral angle. For the last validation and purification of the computed 3D structures of *Xa* models, the ProSA tool (https://prosa.services.came.sbg.ac.at/prosa.php) was used to find out the native protein folding energy of the computed. Besides, the comparison of native protein folding energy with the known structural models' energy field was made.

Results

Motif and Domain Prediction Using Sequential Analysis

The previously published amino acid sequences of *Xa23* (113aa), *Xa23*-Ni (113aa), *Xa10* (126aa), and *Xa10-Ni* (134aa) proteins from *Oryza sativa* and *Oryza indica*, responsible for resistance against BLB disease in rice, were downloaded from the UniProt database (https://www.uniprot.org). The domain search using the SMART tool revealed that all four *Xa* proteins possess trans-membrane helices and low-complexity regions. Three unique motifs are computed from these *Xa* protein sequences through a tool called MEME suite. A 0-order model of the sequence is used as a background model to construct the motifs. Sequence logos are used to represent the conserved regions (PWGWMLEFPY) in *Xa23* and *Xa23*-Ni, as shown in Fig. 1a). The conserved region LAAAYLCRLL was found in *Xa10,* as shown in Fig. 1b), whereas the conserved region FLRFRKVLYLSLFY was found in *Xa10*Ni Fig. 1c).

Primary Structure Analysis

The ProtParam tool was applied to calculate the primary structural parameters of *Xa* proteins, like

Fig. 1. Consensus sequences in *Xa23*, *Xa23*-Ni, *Xa10,* and *Xa*21, *Xa10-Ni,* and their positions. The color shows the motif of a specific sequence, and the sequence against that color motif is defined at the bottom of the image.

molecular weight, which is 13081.97 for *Xa23*, 13081.97 for *Xa23*-Ni, 13870.88 for *Xa10*, and 14751.91 for *XaXa10*-Ni protein, respectively. The isoelectric point (pI) refers to the pH at which the net charge on a protein is zero and the surface of the protein is covered with charge. At pI, proteins behave as compact and stable [29]. The protein *Xa23* and *Xa23*Ni have a pI of 9.34, whereas *Xa10* Ni has a pI of 8.37, indicating its basic nature (pI>7.0), and *Xa10* has a pI of 6.11, indicating its slightly acidic nature (pI ~ 7.0). The amount of space taken up by aliphatic side chains like alanine, valine, isoleucine, and leucine in a protein is known as the aliphatic index (AI) [30]. The aliphatic index of rice *XaXa23* is 141.42, *Xa23*Ni is 141.42, *Xa10* protein is 133.17, and *Xa10-Ni* is 136.94. The instability index provides an approximation of the stability of the protein in an in-vitro environment. It was calculated that the instability index of these proteins is about 45.89, 45.89, 35.59, and 42.09, respectively. The GRAVY index of *Xa23* is 1.011; for *Xa23*-Ni, it is 1.011; for *Xa10,* it is 1.157; and for *Xa10 -Ni*, it is 1.128.

Secondary Structure

In order to find the structural features of *Xa* proteins, their secondary structures were predicted using the SOPMA tool. A high proportion of random coils and alpha helices were found in the protein structures. The α-helices are a core component of formed trans-membrane bundles and represent 30% of the structure of the average globular protein [31]. The TMHMM tool predicted the presence of four trans-membrane helices of 63.67 amino acids in *Xa23*, three trans-membrane helices of 63.67 amino acids in *Xa23*-Ni, four trans-membrane helices of 79.60 amino acids in *Xa10*, and four trans-membrane helices of 85.60 amino acids in *Xa10-Ni*.

Comparative Analysis of Xa Genes

The phylogenies revealed the relationship between *Xa* proteins. The results showed that *Xa23*, *Xa23*-Ni, *Xa10*, and *Xa10-Ni* fall in the same clade because they are more closely related to each other as they have sequence similarity.

Comparative Modeling

Comparative modeling, also known as homology modeling, is considered the most precise and accurate technique for predicting the 3-D structure of a protein sequence. The main purpose of comparative modeling is to build 3-D models for proteins whose structures are not known. It predicts models on the basis of sequence similarity to proteins of known structure. The model prediction process consists of fold assignment, target-template alignment, model building, and model evaluation [32].

It gives us the idea that if there is sequence similarity between two sequences, they also possess structural

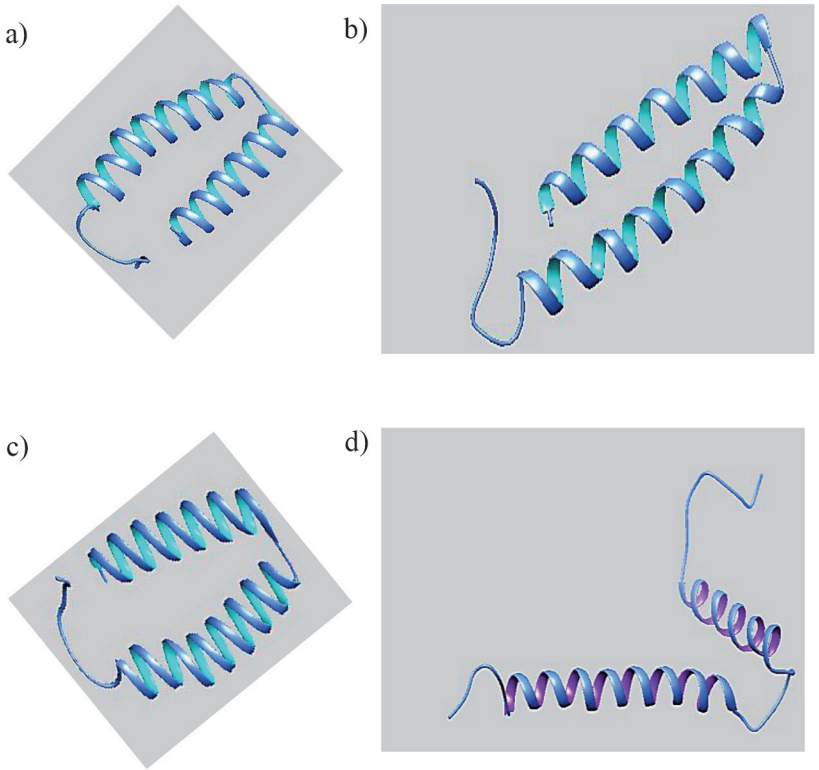

Fig. 2. 3D view of secondary structures of homology-based predicted *Xa* protein models: a) Snip of *Xa23* model, b) Snip of *Xa23*-Ni model, c) Snip of *Xa10* model, and d) Snip of *Xa10-Ni* model using Chimera.

Table 1. The Ramachandran plot result elucidation.

Proteins	Model length	Bad bonds	Bad angles	MolP score
Xa23	113	0	12	2.02
Xa23-Ni	113	0	12	2.01
Xa10	126	0	9	2.05
Xa10 -Ni	134	0	1	0.5
4he8.1.	113	0	12	2.02
7npu.1	134	0	1	0.5

similarity. The parameter for selecting the best template for a model that has maximum query coverage is defined. 4he8.1.C, 6i0d.2.L, 6ziy.1.K, 7npu.1.E, and 7np7.1D were predicted as the top templates of the Swiss model. 4he8.1.C has the highest similarity with *Xa23*, *aXa23*-Ni, *Xa10*, and 7npu.1E, with *Xa10*-Ni as shown in (Fig. 2) [33]. The well-suited templates for *Xa* protein were 4he8.1.C. Pairwise sequence alignment of these templates was done with these *Xa* proteins through Mega X, and for visualization, the Sea-viewer tool (http://doua.prabi.fr/software/seaview) was utilized. On the basis of the alignment of the template and target, the Swiss model created 10 rough models of query protein sequences using 4he8.1.C (for *Xa23*, *Xa23*Ni, and *Xa10*) and 7npu.1.E for *Xa10*-Ni as templates (Table 1). Out of these 10 models, the visualization of secondary structures of selected templates and predicted models was done using Chimera software, and a comparison of these results showed that there were some conserved regions between templates and predicted models (Fig. 3) (Table 2).

Model Assessment and Validation

GNU plot is used for graphical visualization of the alignment of the target and template to study the conservancy of the secondary structure of protein calculated by the Swiss model tool. The ExPasy online server (https://swissmodel.expasy.org/interactive) is utilized to visualize amino acids that fall in the available region using the Ramachandran plot. The Ramachandran plot gives us the reliable value of the amino acid in the model, having Psi clashes and rotatable angles, or dihedral angles. According to Ramachandran plot analysis of the models, it revealed that 88.52% of residues of *XA23*, 88.52% of residues of *Xa23*Ni, 91.94% of residues of *Xa10,* and 98.00% of residues of *Xa10* Ni fall in the favored regions, while 6.56% of residues of *Xa23*, 6.56% of residues of *Xa23*Ni, and 4.84% of residues of *Xa10* fall in Ramachandran outliers (Fig. 4 and Fig. 5). The VERIFY-3D (https://servicesn.mbi.ucla.edu/Verify3D) tool was used to check the model compatibility with its own amino acid according to the sequence, and the score value of over 80% residue indicates the model is reliable [34].

Chromosomal Localization

The rice genome consists of 12 chromosomes [35]. Chromosomal mapping of these proteins was constructed using the MapChart tool (https://www.wur.nl/en/show/mapchart.htm). Fifteen genes, including *Xa*3/26, *Xa*4, *Xa10*, *Xa*21, *Xa*22, *Xa23*, *Xa*26, *Xa*30, *Xa*32, *Xa*35, *Xa*36, *Xa*39, *Xa*40, *Xa*43, and *Xa*44, have been reported to be located on chromosome 11. Results revealed that all four genes were located on chromosome 11, shown in Fig. 6. The length of the Japonica cultivar of rice is estimated

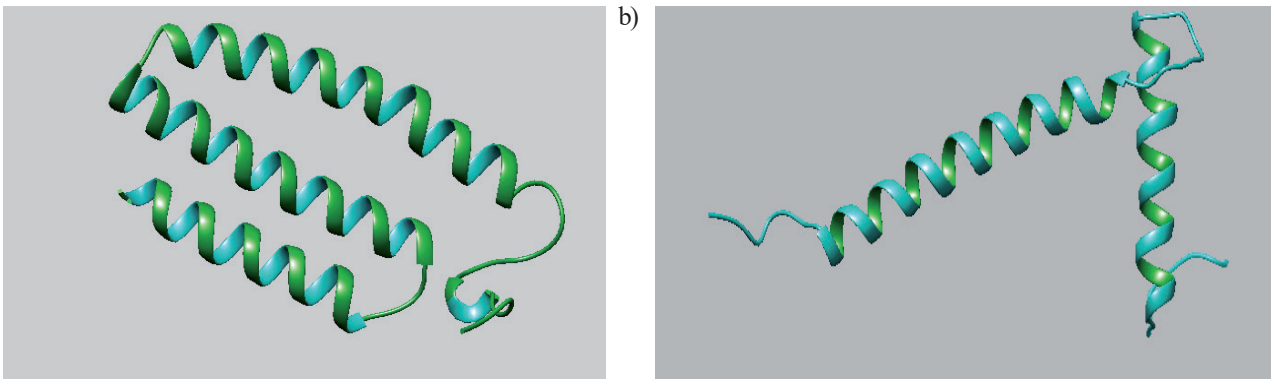

Fig. 3. 3D view of secondary structures of the proteins selected as templates for comparison-based homology modeling: a) Image of 4he8.1 (Chain-C), and b) Image of 7npu.1 (Chain-E).

Table 2. Selected template for modeling from Swiss model against PDB database.

Sr.	Templates	Query coverage	Identity %	GMQE
1	4he8.1.C (for *Xa23*)	56	24	0.26
2	6i0d.2.L	56	24	0.25
3	6ziy.1.K	56	24	024
4	4he8.1.C (for *Xa23*Ni)	56	24	0.26
5	6i0d.2.L	56	24	0.25
6	6ziy.1.	56	24	0.24
7	4he8.1.C (for *Xa10*)	51	25	0.24
8	6i0d.2.L	51	25	0.23
9	6ziy.1.K	51	25	0.23
10	7npu.1.E (for *Xa10* Ni)	39	24	0.12
11	7np7.1.1D	39	24	012

to be 118.6 cm. The long arm of rice chromosome 11 (11L) comprises almost twice (132 genes) the number of R-like genes compared to the short arm (68 genes) [36]. It also possesses 17 downstream defense response genes, including glucanases, chitinases, and thaumatin-like proteins. The amino acid-conserved regions of *Xa23* contribute 50% of the identity to the known executor R protein of *Xa10*. In addition, trans-membrane helices of *Xa23* were also present in *Xa10*. However, the activation of transcription in *Xa23* observed by Avr*Xa23* was distinct from *Xa10* [37].

Discussion

The cause of bacterial leaf blight (BLB), Xanthomonas oryzae pv. oryzae (Xoo), is rapidly spreading throughout the world and causing significant

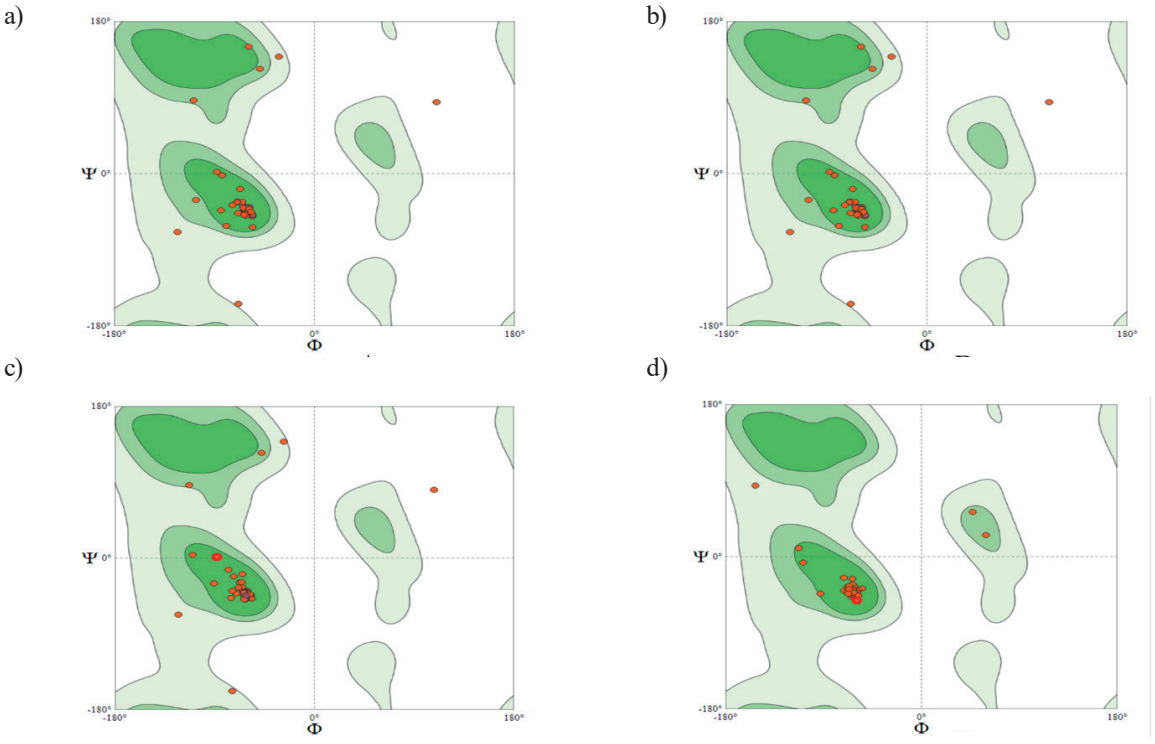

Fig. 4. The Ramachandran plot of predicted 3D structures of proteins: In (A) the Ramachandran plot of *Xa23* protein; (B) the Ramachandran plot of *Xa23*Ni; (C) the Ramachandran plot of *Xa10*; (D) the Ramachandran plot of *Xa10* Ni.

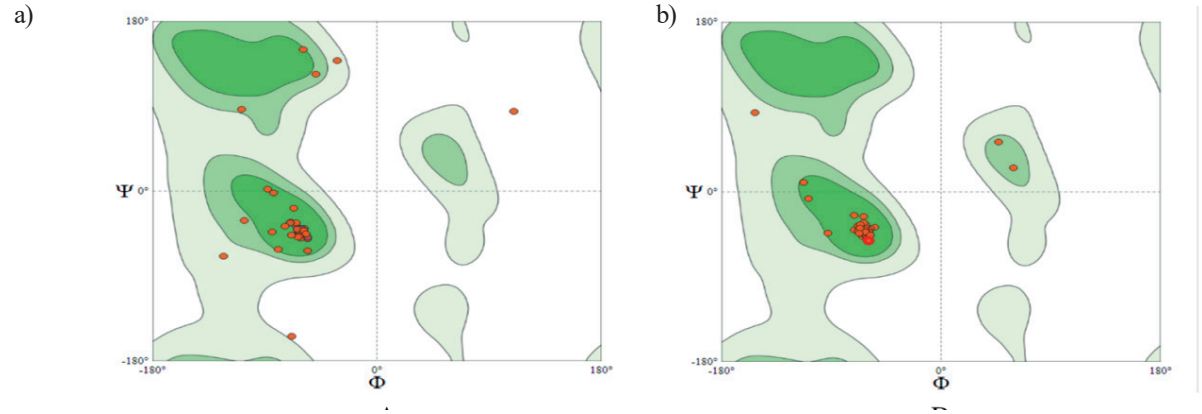

Fig. 5. The Ramachandran plot of selected templates (4he8.1.C and 7npu.1.E).

losses in paddy crops. Furthermore, in this region, due to the monsoon climate, chemical control of this disease is unsuccessful. Since BLB reduces rice quality, it has been suggested that the best way to prevent it is to cultivate resistant rice types. The most efficient and economical way to control the disease is to produce and breed resistant types that possess the resistance gene. The genomic sequence that is contained within this QTL was systematically analyzed in this work. The genes were identified and categorized into resistance and defense-related genes. These genes were then utilized to forecast how these genes will interact and contribute to rice resistance to BLB. The discovery of important genes may be used to improve rice through transformation and marker creation. The rice genome database and every accessible online bioinformatic tool will be used for all the aforementioned analyses. Our research team will use the information from this study and numerous other comparable QTL analyses to create an integrated map that scientists and breeders may use.

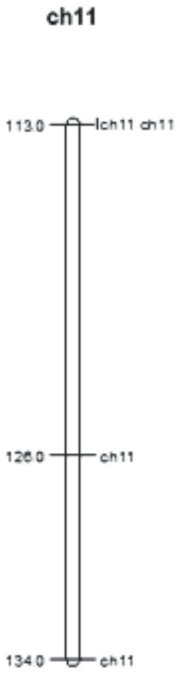

Fig. 6. Chromosomal mapping of *Xa23*, *Xa23*-Ni, *Xa10*, *0Xa10-Ni.*

During motifs and domain prediction, *Xa23* (113aa), *Xa23*-Ni (113aa), *Xa10* (126aa), and *Xa10-Ni* (134aa) proteins possessed trans-membrane helices that performed a diversity of roles such as enzyme catalysis, transport across membranes, transducing signals as receptors of hormones and growth factors, and transfer of energy in ATP synthesis [38]. Particular sequence motifs generally facilitate a common function, like protein binding or targeting a specific subcellular position, in a diversity of proteins [39]. The conservancy of residues showed that *Xa23*, *Xa23*-Ni, *Xa10,* and *Xa10*-Ni have a resemblance to each other and phylogenetically make their own separate cluster. R-genes are a crucial part of a plant's defense mechanism against pathogen invasion. Resistance genes have been categorized into eight types based on the presence of domains, including the well-known NBS-LRR, protein kinase, and other domain configurations that are present in these R genes.

During primary structure analysis, the amount of space taken up by aliphatic side chains is considered a positive factor for the increase in thermal stability of globular proteins. The aliphatic index indicated that these proteins may be stable for a wide range of temperatures [40]. A protein possessing an instability index smaller than 40 is considered stable, while a value greater than 40 is considered potentially unstable [41]. These were on the positive side and showed their slightly hydrophobic nature (GRAVY>0), indicating their low affinity for water. The results of predicted secondary structures showed the proportion of extended strands (β folds), α helices, β turns, and random coils accounted for *Xa23*; 23.01, 56.64, 1.77, and 18.58%, for *Xa23*-Ni; 23.01, 56.64, 1.77, and 18.58%, for *Xa10*; 18.25, 54.76, 3.96, and 23.02%, for *Xa10-Ni*; 18.66, 48.51, 2.99, and 29.85% of the secondary structure, respectively. The comparative analysis showed that these *Xa* genes are executor genes

and work as promoter traps. They have a very close relationship with *Xa*27 [29]. Forty antimicrobial genes have been found, the majority of which have been cloned and thoroughly mapped. Yasmin et al. (2017) report that a Teqing x Lemont hybrid produced a QTL, qBBR11-1, that showed resistance to BLB illness caused by three different forms of Xoo: C2, C4, and C510 for two years in a row.

During comparative modeling, it was noted that the models having maximum query coverage and the highest percent identity were considered thermodynamically stable and selected for the next validation. The information provided by Ramachandran is that the model has the same percentage of allowed and disallowed amino acids as the template, so it was concluded that the model and template were mostly conserved, and similarly, VERIFY-3D indicated that the model is reliable. In chromosomal localization, the fast-evolving disease resistance and defense response genes found on rice chromosome 11 were created by tandem duplication, followed by divergence under the selection pressure of rice pathogens.

Conclusions

These genes belong to the *Xa* family of rice and act as a code for executor proteins that have a strong hypersensitive response and possess complete resistance against pathogens, especially Xoo. Analysis showed that these genes hold similarities and have low-complexity regions and trans-membrane helices. Because the resistance and defense response genes, enriched on 11 chromosomes relative to the whole genome, also occur in clusters, they offer a favored target for the isolation of their allelic variations and the development of long-lasting disease resistance in rice. On chromosome 11, there were 4,286 identified genes, 3,148 of which may be classified as non-transposable elements. A similar percentage of genes on the chromosome could be assigned a putative function or annotated as encoding an expressed protein, leaving a similar percentage annotated as encoding a hypothetical protein. Based on their encoded protein products, disease resistance genes (R-like genes) that confer resistance to viral, bacterial, fungal, and nematode pathogens have been divided into five groups. For instance, a genome-wide study identified typical protein domains like LRR, CC, NBS, etc., in the R-like genes of Arabidopsis. A previous report on the analysis of R-like genes in the entire rice genome revealed that most of the R-like genes (24.98%) are present on chromosome 11. The domains' identification and characterization will lead us to a specific domain that involves resistance biologically. The Xoo races are changing continuously, and we can use these identified specific regions to predict new genes and pairing arrangements of genes for the upcoming races.

Authors' Contributions

Conceptualization, M.S.; R.I.; methodology, R.I.; M.Z.M.; software, K.A.; K.I.; formal analysis, R.A.; investigation, M.S.; resources, M.S., and R.I.; data curation, R.I.; writing – original draft preparation, M.Z.M., S.A., K.A.A., Y.M.H., I.K., U.T., and A-R.Z.G.; writing – review and editing, M.S., Q.Z., K.A., M.H.S., and M.S.H.; supervision, M.S.; project administration, S.A., K.A.A., Y.M.H., and M.S.; funding acquisition, M.H.S., A-R.Z.G., and M.S.H. All authors read and approved the final manuscript.

Acknowledgments

The authors extend their appreciation to the Researchers Supporting Project number (RSP-2025R241), King Saud University, Riyadh, Saudi Arabia. The authors would like to extend their sincere appreciation to the Researchers Supporting Project number (IAEA CRP-23080), IAEA, Vienna.

Conflict of Interests

The authors declare that there are no conflicts of interest related to this article.

References

1. ZHOU X., SHAFIQUE K., SAJID M., ALI Q., KHALILI E., JAVED M., HAIDER M., ZHOU G., ZHU G. Era-like GTP protein gene expression in rice. Brazilian Journal of Biology, **82**, e250700, **2021**.
2. BONNAVE M., BERTIN P. Evaluation of nitrogen use efficiency (NUE) of rice (*O. sativa*) under saline conditions. **2018**.
3. EL ARJA S. Neuromorphic perception for greenhouse technology using event-based sensors. **2022**.
4. KUMAR A., KUMAR V., RODRÍGUEZ-SEIJO A., SETIA R., SINGH S., KUMAR A., SETH C.S., SOMMA R. Appraisal of heavy metal (loid) s contamination in rice grain and associated health risks. Journal of Food Composition and Analysis, 106215, **2024**.
5. RUBIN B., MAGLASANG R., MAYO F. Modeling the Rice Production and Consumption in the Philippines: A System Dynamics Approach. Available at SSRN 4247739. **2022**.
6. KRISHNAN S.G., VINOD K., BHOWMICK P.K., BOLLINEDI H., ELLUR R.K., SETH R., SINGH A. Rice Breeding 3. Fundamentals of Field Crop Breeding, **192** (966.88), 113, **2022**.
7. SIMKHADA K., THAPA R. Rice blast, a major threat to the rice production and its various management techniques. Turkish Journal of Agriculture-Food Science and Technology, **10** (2), 147, **2022**.
8. IQBAL S. Insect, pest and disease management in rice. Austin Publication, Irving, TX, **85**, **2020**.
9. KUMAR A., KUMAR R., SENGUPTA D., DAS S.N., PANDEY M.K., BOHRA A., SHARMA N.K., SINHA

P., SK H., GHAZI I.A. Deployment of genetic and genomic tools toward gaining a better understanding of rice-Xanthomonas oryzae pv. oryzae interactions for development of durable bacterial blight resistant rice. Frontiers in Plant Science, **11**, 1152, **2020**.
10. SANYA D.R.A., SYED-AB-RAHMAN S.F., JIA A., ONÉSIME D., KIM K.-M., AHOHUENDO B.C., ROHR J.R. A review of approaches to control bacterial leaf blight in rice. World Journal of Microbiology and Biotechnology, **38** (7), 113, **2022**.
11. KHARAYAT B.S. Diseases of Field and Horticultural Crops and their Management Volume–I. PHI Learning Pvt. Ltd., **2023**.
12. TEPER D., WHITE F.F., WANG N. The dynamic transcription activator-like effector family of Xanthomonas. Phytopathology®, **113** (4), 651, **2023**.
13. MONDAL R., KUMAR A., GNANESH B.N. Crop germplasm: Current challenges, physiological-molecular perspective, and advance strategies towards development of climate-resilient crops. Heliyon, **9** (1), **2023**.
14. KHANNETAH K., RAMCHANDER S., LEON M.A.P., SHOBA D., SARAVANAN S., KANNAN R., YASIN J.K., PILLAI M.A. Genetic diversity analysis in indigenous rice (*Oryza sativa* L.) germplasm for bacterial leaf blight (*Xanthomonas oryzae* pv. oryzae)(BB) using resistance genes-linked markers. Euphytica, **217** (7), 145, **2021**.
15. WANG S., LIU W., LU D., LU Z., WANG X., XUE J., HE X. Distribution of bacterial blight resistance genes in the main cultivars and application of Xa23 in rice breeding. Frontiers in Plant Science, **11**, 555228, **2020**.
16. SINGH B.K., DELGADO-BAQUERIZO M., EGIDI E., GUIRADO E., LEACH J.E., LIU H., TRIVEDI P. Climate change impacts on plant pathogens, food security and paths forward. Nature Reviews Microbiology, **21** (10), 640, **2023**.
17. JI Z., GUO W., CHEN X., WANG C., ZHAO K. Plant executor genes. International Journal of Molecular Sciences, **23** (3), 1524, **2022**.
18. NGUYEN M.H.R. Host-induced adaptation of Xanthomonas oryzae pv. oryzae. University of Southern Queensland, **2021**.
19. YANG Y., ZHOU Y., SUN J., LIANG W., CHEN X., WANG X., ZHOU J., YU C., WANG J., WU S. Research progress on cloning and function of Xa genes against rice bacterial blight. Frontiers in Plant Science, **13**, 847199, **2022**.
20. AFTAB U., ALI S., GHANI M.U., SAJID M., ZESHAN M.A., AHMED N., MAHMOOD R. Epidemiological Studies of Bacterial Leaf Blight of Rice and its Management. Basrah Journal of Agricultural Sciences, **35** (1), 106, **2022**.
21. RAMASETTY B.T., KUMAR R.M., UDAYSHANKAR A.C., PRAKASH H.S. Chapter-2 Phyto-Endophytes for the Sustainable Management of Bacterial Blight Disease in Rice Caused by Xanthomonas oryzae pv. oryzae (Xoo); Recent Advances and Future Challenges.
22. ISNAINI I., NUGRAHA Y., BAISAKH N., CARSONO N. Toward Food Security in 2050: Gene Pyramiding for Climate-Smart Rice. Sustainability, **15** (19), 14253, **2023**.
23. SCHNEIBLE J.D. A Material Toolbox for Advanced Therapeutics. North Carolina State University, **2020**.
24. ALTWAIJRY N., ALMASOUD M., ALMALKI A., AL-TURAIKI I. Multiple sequence alignment using a multiobjective artificial bee colony algorithm. IEEE, **2020**.
25. SHAMSHAD A., RASHID M., ZAMAN Q.U. In-silico analysis of heat shock transcription factor (OsHSF) gene family in rice (Oryza sativa L.). BMC Plant Biology, **23** (1), 395, **2023**.
26. YANG M., DERBYSHIRE M.K., YAMASHITA R.A., MARCHLER-BAUER A. NCBI's conserved domain database and tools for protein domain analysis. Current Protocols in Bioinformatics, **69** (1), e90, **2020**.
27. HAVLIK L.P., DAS A., MIETZSCH M., OH D.K., ARK J., MCKENNA R., AGBANDJE-MCKENNA M., ASOKAN A. Receptor switching in newly evolved adeno-associated viruses. Journal of Virology, **95** (19), 10.1128/jvi. 00587, **2021**.
28. SINGH A., PANDEY A.K., DUBEY S.K. Genome sequencing and in silico analysis of isoprene degrading monooxygenase enzymes of Sphingobium sp. BHU LFT2. Journal of Biomolecular Structure and Dynamics, **41** (9), 3821, **2023**.
29. CHEN X., QIU Q., CHEN K., LI D., LIANG L. Water-soluble myofibrillar protein–pectin complex for enhanced physical stability near the isoelectric point: Fabrication, rheology and thermal property. International Journal of Biological Macromolecules, **142**, 615, **2020**.
30. YD M., DS M., KANNARATH A. In Silico Drug Search for Better Treatment for Cancer: L-Asparaginase. International Journal of Research, **2** (11), 75, **2014**.
31. BEVACQUA A., BAKSHI S., XIA Y. Principal component analysis of alpha-helix deformations in transmembrane proteins. Plos One, **16** (9), e0257318, **2021**.
32. WEBB B., SALI A. Comparative protein structure modeling using MODELLER. Current Protocols in Bioinformatics, **54** (1), 5.6. 1, **2016**.
33. ORBÁN-NÉMETH Z., BEVERIDGE R., HOLLENSTEIN D.M., RAMPLER E., STRANZL T., HUDECZ O., DOBLMANN J., SCHLÖGELHOFER P., MECHTLER K. Structural prediction of protein models using distance restraints derived from cross-linking mass spectrometry data. Nature Protocols, **13** (3), 478, **2018**.
34. HASHEMZADEH P., GHORBANZADEH V., LASHGARIAN H.E., KHEIRANDISH F., DARIUSHNEJAD H. Harnessing bioinformatic approaches to design novel multi-epitope subunit vaccine against Leishmania infantum. International Journal of Peptide Research and Therapeutics, **26** (3), 1417, **2020**.
35. SHAMSHAD A., RASHID M., HAMEED A., IMRAN ARSHAD H.M. Identification of biochemical indices for brown spot (*Bipolaris oryzae*) disease resistance in rice mutants and hybrids. Plos One, **19** (4), e0300760, **2024**.
36. HAN X., YANG Y., WANG X., ZHOU J., ZHANG W., YU C., CHENG C., CHENG Y., YAN C., CHEN J. Quantitative trait loci mapping for bacterial blight resistance in rice using bulked segregant analysis. International Journal of Molecular Sciences, **15** (7), 11847, **2014**.
37. PANDA S., BUSATTO N., HUSSAIN K., KAMBLE A. Piriformospora indica-primed transcriptional reprogramming induces defense response against early blight in tomato. Scientia Horticulturae, **255**, 209, **2019**.
38. RAMASARMA T. Transmembrane domains participate in functions of integral membrane proteins. Indian Journal of Biochemistry & Biophysics, **33** (1), 20, **1996**.
39. TIRINCSI A., SICKING M., HADZIBEGANOVIC D., HAßDENTEUFEL S., LANG S. The molecular biodiversity of protein targeting and protein transport related to the endoplasmic reticulum. International Journal of Molecular Sciences, **23** (1), 143, **2021**.

40. SAINI P., YASHVEER S., REDHU N. S., CHAUDHARY S., KAMBOJ A., HEMBADE V., SHARMA K., SANGWAN S. Opportunities for Bioinformatics Tools for the Management of Rice Bacterial Diseases. Apple Academic Press, **2023**.

41. TONG X., CAO J., SUN M., LIAO P., DAI S., CUI W., CHENG X., LI Y., JIANG L., WANG H. Physical and oxidative stability of oil-in-water (O/W) emulsions in the presence of protein (peptide): Characteristics analysis and bioinformatics prediction. Lwt. **149**, 111782, **2021**.