RESEARCH
ISSN: 2639-5274
Volume 09 Issue 01

In-silico Diversity Analysis of Disease resistant NBS-LRR protein variants in *Oryza sativa*

Fatima Muccee[1*], Naeem Mahmood Ashraf[1], Shafia Arshad[2], Sohail Razak[3], Tayyaba Afsar[3], Ali Almajwal[3], Huma Shafique[3]

[1]School of Biochemistry and Biotechnology, University of Punjab, Lahore 52254, Pakistan. naeem.sbb@pu.edu.pk

[2] The Islamia University of Bahawalpur. E-mail: shafia.arshad@iub.edu.pk

[3]Department of Community Health Sciences, College of Applied Medical Sciences, King Saud University, Riyadh, Saudi Arabia. E-mail: smarazi@ksu.edu.sa, tayyaba_sona@yahoo.com, aalmajwal@ksu.edu.sa, humashafique@gmail.com

Corresponding author*: Fatima Muccee

E-mail: fatima.sbb@pu.edu.pk

Abstract: *Oryza sativa* (rice) being the contributor of more than 2 million tons of staple food plays assorted role in agrarian economy of Pakistan. Variety of pathogens and use of insecticides result in serious loses of its yield. To develop immunity at genetic level in *O. sativa* crops is the crucial need of the time. A diverse disease resistant gene restricts the polycyclic epidemics development induced by pathogens in the plants. Nucleotide

binding site-leucine rich repeat proteins (NBS-LRR) being associated with regulation of plant proteins targeted by pathogens and detection of diverse pathogens are considered as adaptable guards for plants. Eight genes from *O. sativa* encoding eight NBS-LRR protein variants have been targeted for their diversity analysis in present study. For characterization of NBS-LRR disease resistant protein in *O. sativa*, eight nucleotide sequences were retrieved from Rice Genome Annotation Project (RGAP) database and translated via Expasy translation tool. The eight protein variants were then assessed for their physicochemical properties, sub-cellular localizations, 2D and 3D configurations using ProtParam, CELLO2 and SOPMA tools and SWISS MODEL server. These eight forms of disease resistant protein showed great diversity in their attributes. Highest number of amino acids and molecular weights were observed in case of NBS-LRR2 and LRR4. Isoelectric point (I), instability index, aliphatic and GRAVY showed only slight variation. Extinction coefficient showed considerable variation with highest and lowest values for LBS-LRR2 and LRR1, respectively. Diversity was also observed in sub-cellular localization of these NBS-LRR variants i. e. LRR1, 2, 4 and 6 (nuclear + cytoplasmic), LRR2, 3 and 7 (nuclear only) and LRR5, 6 and 8 (cytoplasmic only). Secondary structure was highly variable in terms of extended strand (ranging between 9.51% and 12.75%) and beta turn (ranging between 1.88% and 5.65%) while only small difference was observed in case of alpha helix and random coil. The 3D configurations of all the variants of protein documented in present study showed great variation and high level of complexity. This characterization revealed significant

heterogeneity among different forms of disease resistant NBS-LRR proteins. Selective breeding of rice plants containing the diverse forms of the gene, addressed in present study will help to remove homogeneity in rice crops. This might result in reduced epidemics, yield stability and resistance durability.

Keywords: Oryza sativa, divergence, variants, isoforms, NBS-LRR, encoding, resistance

Introduction

Agri food sector has been facing challenges due to increase in human population since last few decades. Growing food needs are mostly compensated by cereals production. *Oryza sativa* (rice) is one of the most cultivated cereals and major agronomic species which feeds 50% of the world population. Its production is 25% of world cereals [1, 2]. Asian countries including Indonesia, Pakistan, China and India produce 92% of the world rice which is a food source of approximately 4.6 billion consumers [3, 4]. In addition to Asian countries, rice is also a staple food in U.S. In year 2020, rice crops harvested on an area of 2.987 million acres in U.S valued at 2495 million dollars [5, 6]. In 2021-2022, milled rice production in U.S was estimated to be 511.7 million metric tons [7]. Nutritionally, rice comprises of 75-80% starch, 12% water, 7% proteins with 4% lysine minerals like calcium, phosphorus, magnesium, copper, manganese, iron and zinc [8-10].

Globally 30% of rice production is lost due to pathogens [11]. The *Lissorhoptrus oryzophilus* (rice water weevil), *Diatraea saccharalis* (sugarcane

borer), *Chilo plejadellus* (rice stalk borer), *Nilaparvata lugens* (brown planthopper), *Pseudaletia unipuncta* (armyworm), *Oebalus pugnax* (stink bug), *Cofana* sp. (white leafhopper) and *Magnaporthe oryzae* (rice blast fungus) have been reported to considerably affect the rice yield thus playing a devastating role in economy [12]. Conventionally, several methods have been employed to minimize pests build up like field draining to stop aquatic larvae, planting trap crop for plant borers, weeding, land preparation, mixed cropping, fertilizer management, tillage, synchronous and asynchronous planting, time of planting and use of pesticides etc. [13-15]. These methods suffer from multiple drawbacks like accidental poisoning, environmental pollution, and expensive equipment for insecticide application[16-19].

These above mentioned pathogens with diverse modus operandi have made rice crop security a serious concern. Managing these pathogens via proteome engineering is the most sustainable solution. Manipulating resistant proteins and cloning various isoforms in the same plant through breeding and genetic engineering techniques is an emerging biocontrol strategy [20]. To engineer proteins, genes associated with pathogens resistance pathways in plant should be targeted. Resistant genes play crucial role in pathogen resistance mechanism of plants known as effector-triggered immunity (ETI). One such type of potent genes is Nucleotide binding sites-leucine rich repeat (NBS-LRR). In ETI pathway, resistant genes encode transmembrane receptors which bind and form complex with specific avirulence (Avr) proteins of pathogens. This complementary binding initiate the conformational change in LRR and amino terminal

domains of NBS-LRR protein [21, 22]. This change stimulates NBS domain of protein to exchange ADP for ATP thereby activating mechanism that stops pathogen spreading by cell death [23].

NBS-LRR are the disease resistant genes. Majority of the isoforms constituting this group are uncharacterized. These are constitutively expressed in plants. NBS-LRR have two major groups i.e. one which encodes coiled coil motif at amino terminal (CC-NBS-LRR) and other comprises of N-terminal domain homologous to toll interleukin 1 receptor domain (TIR) of mammals [23].

In present study we have selected eight, Avr effector interacting, variants of NBS-LRR resistant protein which have been reported to be associated with resistance development against *Magnaporthe oryzae* in two varieties of rice cultivars i.e. BR2655 and HR12 [24]. Objectives of present project were the characterization of these variants at the level of 2D and 3D configuration and physicochemical properties which might be helpful in isolating these resistance imparting genes from different rice plants and engineering them into same plant thus producing rice crops with potential pathogens resistance.

Rapid change of environmental conditions and excessive use of pesticides, pathogens are developing into multiple races. To combat these pathogens and to attain environmental sustainability, immunity development at the level of genes is crucial.

2. Methodology

2.1 Retrieving the sequences of eight variants of NBS-LRR gene from database

Sequences of eight variants of NBS-LRR gene in *O. sativa* were retrieved from Rice Genome Annotation Project (Osa1) Release 7 Annotation (rice.uga.edu/amalyses_search_locus.shtml, accessed on July 2023). [25].

2.2 ExPASy translation tool

The coding nucleotide sequences (CDS) of NBS-LRR genes were translated into proteins by using ExPASy translation tool available at (https://web.expasy.org/translate/, accessed on July 2023). [26] .

2.3 ProtParam tool

Physicochemical properties of NBS-LRR protein variants were assessed using ExPASy ProtParam tool available at (https://web.expasy.org/protparam/, accessed on July 2023) [27].

2.4 CELLO: Subcellular localization predictor

To predict the sub-cellular localization of NBS-LRR protein variants, CELLO v.2.5: Subcellular localization predictor available at (cello.life.nctu.edu.tw, accessed on July 2023) was employed [28].

2.5 SOPMA tool

For assessment of secondary structure of NBS-LRR variants, SOPMA secondary structure prediction method online server available at (npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html,

accessed on July 2023) was used [29]. Using this tool following attributes of 2D configuration were analyzed i.e. alpha helix, extended strand, beta turn and random coil.

2.6 SWISS-MODEL

Protein structure homology modelling server, SWISS-MODEL available at (https://swissmodel.expasy.org, accessed on July 2023) was consulted for analysis of 3D configuration of NBS-LRR protein variants [30].

2.7 Phylogeny of NBS-LRR protein variants

To analyze the evolutionary relationship of NBS-LRR protein variants of *O. sativa* documented in present study, among themselves and also with the same proteins of some other plants like *Solanum lycopersicum* (tomato), Arabidopsis thaliana (thale cress) and *Glycine max* (soybean), protein sequences were retrieved from NCBI (National Center for Biotechnology Information) GenBank database available at (https://www.ncbi.nlm.nih.gov/nucleotide/, accessed on July 2023). Clustal Omega Multiple Sequence Alignment software was used for multiple sequence alignment of variants documented in present study [31]. Ungapped aligned sequences were then subjected to MEGA11 software and phylogenetic tree was constructed [32]. Neighbour joining tree was constructed using the bootstrap value of 100 as reliability index.

3. Results

Nucleotide sequences retrieved from RGAP for NBS-LRR gene variants and their protein sequences are shown in Supplementary data Table 1 and

Supplementary data Figure 1. To predict the diversity at different levels among these variants, protein sequences were analyzed further. The results of analyses are discussed in detail.

3.1 Prediction of sub-cellular localization of variants

CELLO predictor showed diversity in the sub-cellular localization of present study protein forms. NBS-LRR1, LRR2, LRR4, LRR6 and LRR8 were found to be localized in cytoplasm and nucleus with reliability scores of 1.388 & 1.138, 1.377 & 2.432, 1.290 & 1.618, 1.676 & 1.248 and 2.705 & 1.325, respectively. These values are showing that LRR2 is more localized in nucleus and LRR8 in cytoplasm. In all other cases, there is equal probability of localization in cytoplasm and nucleus. NBS-LRR3 and LRR7 are found in nucleus with score of 3.145 and 3.671, respectively. NBS-LRR5 variant is found to be localized in cytoplasm with reliability score of 3.145 (Figure 1).

3.2 Prediction of physicochemical properties of variants

Differences among these protein variants at the level of physicochemical properties was analyzed using ProtParam tool (Table 1). Analysis revealed all the variants comprised of different number of amino acids and hence different molecular weights. NBS-LRR1 and LRR3 were the shortest in length with 2574 and 2709 amino acids, respectively while NBS-LRR2 and LRR4 were comprised of the highest number of amino acids i.e. 4479 and 4443, respectively. Not much deviation was observed in case of pI which ranged between 4.77 and 4.90. Highest extinction coefficient was observed in case of NBSS-LRR2 (57875 $M^{-1}cm^{-1}$) and LRR4 (57750 $M^{-1}cm^{-1}$). The

lowest value was observed for LRR1 i.e. 32625 $M^{-1}cm^{-1}$. While LRR3, LRR5, LRR6, LRR7 and LRR8 were showing intermediate values i.e. 34250, 36375, 40250, 50875 and 36250, respectively.

3.3 Prediction of 2D structure of protein variants

NBS-LRR protein variants were found to exhibit diversity in 2D structure in terms of extended strand and beta turn (Table 2 and Figure 2). While in case of alpha helix and random coil, the deviation was less. Extended strand ranged between 9.51 and 12.75. LRR4, LRR5 and LRR8 were found to exhibit 9.93%, 9.97% and 9.51% extended strand content. LRR3 showed highest content with 12.75%. LRR1 exhibited the second highest i.e. 11.20%. Beta turn ranged between 1.88% and 5.65%, the lowest being observed in case of LRR2 and the highest for LRR7. No significant variation was observed in case of alpha helix and random coil content which ranged between 50.68% - 56.82% and 30.49% - 36.39%, respectively.

3.4 Prediction of 3D configuration of protein variants

Considerable diversity was observed in 3D configuration of NBS-LRR protein variants (Figure 3). LRR7 exhibited the simplest level of folding as compared to others followed by LRR1, LRR2 and LRR4. While the remaining variants were found to have complex level of folding.

3.5 Phylogeny analysis

Neighbour joining tree constructed via MEGA11 revealed closest relation of NBS-LRR1 with O. sativa as both originated from same branch point via bootstrap value of 100. Remaining seven isoforms were unrelated

because they did not directly share a clade with O. sativa (Figure 4). These both shared clade with Arabidopsis thaliana only with bootstrap value of 27. LRR2 and LRR6 shared clade but with bootstrap value of only 46. These two also shared clade with LRR6. LRR5 and LRR8 were found phylogenetically closely related due to originating from same branch point with bootstrap value of 100. These two also showed closeness to LRR3 and LRR4 but the reliability index was very small. As per this phylogenetic tree, LRR3, LRR4, LRR5 and LRR8 were more evolutionary close to Glycine max than O. sativa. This analysis showed considerable divergence among the NBS-LRR protein isoforms documented in present study.

Discussion

Several studies have reported the divergence among NBS-LRR gene in different varieties of rice plants [33-38]. Most of these studies focused genome wide association analysis to target NBS-LRR profiling [39-41]. In a study, different genes from NBS-LRR group i.e. qBFR4 and qLBL5 from rice plant have been characterized through *in-silico* technique of Blast2GO [42]. Another investigation retrieved the coding sequences of NBS-LRR group genes from GenBank and compared at the level of number of nucleotides, amino acids and protein configuration [43].

As per the literature, NBS-LRR proteins are cytoplasmic in localization [44]. This is almost consistent with our study as seven out of eight proteins have been found to be localized in cytoplasm, although in some cases the proteins were also predicted to be found in nucleus in

addition to cytoplasm. This might be contributed by significant diversity among these proteins.

The pI of eight isoforms were approximately same ranging from 4.77 to 4.90 showing their acidity. This is consistent with literature which reports acidic I of six out of eight variants. In case of NBS-LRR3 and LRR8, our findings were not consistent with previous one because these two isoforms showed basic values [24]. The molecular weights were found to be 213.008 kDa, 369.87806 kDa, 223.861 kDa, 364.19279 kDa, 253.28646 kDa, 267.58843 kDa, 308.94675kDa and 252.92825 kDa. These values are considerably different from the one reported in literature for the same eight isoforms, the minimum and maximum being recorded to be 97.8 kDa and 168.3 kDa, respectively.

The values for aliphatic index were found in the range of 27.23 to 32.05 which are totally inconsistent with the values range 88.09 to 103.58 reported in previous work [24]. Aliphatic index is the measure of the proportion of protein comprising of aliphatic amino acids and is directly related with thermal stability of protein [45]. Eight variants analyzed in present work does not seem to be thermally stable while the value reported previously showed high thermal stability of the isoforms. This is the first study reporting instability index, extinction coefficient and GRAVY for these eight isoforms of NBS-LRR protein as no one has performed analysis of these parameters. Stability of protein in test tube is indicated by its instability index. The value above 40 shows stable nature pf protein [46]. So, in present study, all the variants except NBS-LRR5 and LRR8 were stable.

The variation with respect to 2D configuration among eight variants was prominent in cases of extended strand and beta turn. Our value of number of amino acids participating in alpha helix formation i.e. 50.68 to 56.80% was not consistent with earlier reported literature in which 25-46% residues have been found to be associated with alpha helix in these eight variants. Earlier work reports involvement of 9-14% residues with beta turn while in present study these values were lower than the reported one. As it was found to be 1.88 to 5.65%. Similarly, the amino acids participating in random coil formation in present study were predicted to range from 30.49 to 36.39%. This finding is not in consistence with earlier reported literature where this value ranged between 48-60% [24].

Our finding of highly folded and complex 3D configuration of NBS-LRR protein variants is consistent with earlier reported findings where different isoforms of this protein exhibited complication structures [24, 43].

Our phylogenetic analysis has revealed that the isoforms of NBS-LRR protein evolved considerably and showed divergence when aligned with *O. sativa*. This finding is not consistent with earlier study which reports conserved nature of different isoforms of the protein addressed in present study [47].

The marked diversity observed in present study is in accordance with previous literature as a study has reported high diversity in nucleotides

and copy number of NBS-LRR gene loci in fourteen wild rice populations and twenty cultivars [48].

Conclusion

Pyramiding of these isoforms of NBS-LRR gene into a single rice plant genome through artificial breeding and genetic engineering strategies might induce broad-spectrum pathogen resistance. Phenotypic characteristics cannot help in identification of plants with potential resistant genes. The characteristics of variants analyzed in present study might be used as markers for selection of these genes which can then be isolated and introduced into a plant simultaneously thus developing strong resistance in rice crops. Coupling of these resistant plants with chemical or cultural pathogenic control strategies might considerably enhance the crops yield and may be more sustainable solution.

Statements and Declarations:

Funding: The authors extend their appreciation to the Researchers Supporting project number (RSP2023R502), King Saud University, Riyadh Saudi Arabia for funding this project.

Acknowledgements: The authors extend their appreciation to the Researchers Supporting project number (RSP2023R502), King Saud University, Riyadh Saudi Arabia for funding this project.

Author contributions: F.M designed methodology, N.M.A perceived the idea, S.A retrieved data from database, S.R performed analysis, T.A write up of manuscript, A.A visualization, H.S formal analysis

Informed consent: N/A

Competing interests: The authors have no competing interests.

Data availability statement: The sequences of gene variants of NBS-LRR documented in present study are available at Rice Genome Annotation Project (Osa1) Release 7 Annotation (rice.uga.edu/amalyses_search_locus.shtml).

Submission declaration and verification: The work is not published previously and it is not under consideration for publication elsewhere.

References

1. FAO. Crop prospects and food situation-Quarterly global report No. 1, March 2020. *Rome* 2020.
2. Vinci G, Ruggieri R, Ruggeri M, et al. Rice production chain: environmental and social impact assessment—a review. *Agriculture* 2023; 13: 340.
3. Gnanamanickam SS and Gnanamanickam SS. Rice and its importance to human life. *Biological control of rice diseases* 2009: 1-11.
4. Muthayya S, Sugimoto JD, Montgomery S, et al. An overview of global rice production, supply, trade, and consumption. *Annals of the new york Academy of Sciences* 2014; 1324: 7-14.

5. NASS U. Crop Values 2017 Summary. *US Department of Agriculture National Agricultural Statistics Service* 2018.

6. Godara N. *Evaluation of factors that contribute to injury to quizalofop-resistant rice from quizalofop applied postemergence*. University of Arkansas, 2022.

7. Bagnall DK, Shanahan JF, Flanders A, et al. Soil health considerations for global food security. *Agronomy Journal* 2021; 113: 4581-4589.

8. Verma DK and Srivastav PP. Bioactive compounds of rice (*Oryza sativa* L.): Review on paradigm and its potential benefit in human health. *Trends in Food Science & Technology* 2020; 97: 355-365.

9. Guan C, Long X, Long Z, et al. Legumes flour: a review of the nutritional properties, physiological functions and application in extruded rice products. *International Journal of Food Science & Technology* 2023; 58: 300-314.

10. Beaulieu JC, Boue SM and Goufo P. Health-promoting germinated rice and value-added foods: a comprehensive and systematic review of germination effects on brown rice. *Critical Reviews in Food Science and Nutrition* 2022: 1-34.

11. Savary S, Willocquet L, Pethybridge SJ, et al. The global burden of pathogens and pests on major food crops. *Nature ecology & evolution* 2019; 3: 430-439.

12. Oyediran IO, Ndongidila A and Heinrichs EA. Strepsipteran parasitism of white leafhoppers, *Cofana* spp.(Hemiptera: *Cicadellidae*) in

lowland rice in Cote d'Ivoire. *International Journal of Pest Management* 2000; 46: 141-147.

13. Fahad S, Nie L, Hussain S, et al. Rice pest management and biological control. *Sustainable agriculture reviews: cereals* 2015: 85-106.

14. Huang S, Wang L, Liu L, et al. Nonchemical pest control in China rice: a review. *Agronomy for sustainable development* 2014; 34: 275-291.

15. Heong KL, Teng PS and Moody K. Managing rice pests with less chemicals. *GeoJournal* 1995; 35: 337-349.

16. Abdullah AR, Bajet CM, Matin MA, et al. Ecotoxicology of pesticides in the tropical paddy field ecosystem. *Environmental Toxicology and Chemistry: An International Journal* 1997; 16: 59-70.

17. Khoury WEl and Makkouk K. Integrated plant disease management in developing countries. *Journal of Plant Pathology* 2010: S35-S42.

18. Damalas CA and Koutroubas SD. Botanical pesticides for eco-friendly pest management: Drawbacks and limitations. *Pesticides in Crop Production: Physiological and Biochemical Action* 2020: 181-193.

19. Carrão DB, Perovani IS, de Albuquerque NCP, et al. Enantioseparation of pesticides: A critical review. *TrAC Trends in Analytical Chemistry* 2020; 122: 115719.

20. Pandit MA, Kumar J, Gulati S, et al. Major biological control strategies for plant pathogens. *Pathogens* 2022; 11: 273.

21. Chang M, Chen H, Liu F, et al. PTI and ETI: convergent pathways with diverse elicitors. *Trends in Plant Science* 2022; 27: 113-115.

22. Yuan M, Ngou BPM, Ding P, et al. PTI-ETI crosstalk: an integrative view of plant immunity. *Current opinion in plant biology* 2021; 62: 102030.

23. DeYoung BJ and Innes RW. Plant NBS-LRR proteins in pathogen sensing and host defense. *Nature immunology* 2006; 7: 1243-1249.

24. Chandrakanth R, Sunil L, Sadashivaiah L, et al. In silico modelling and characterization of eight blast resistance proteins in resistant and susceptible rice cultivars. *Journal of Genetic Engineering and Biotechnology* 2020; 18: 1-15.

25. Yuan Q, Ouyang S, Wang A, et al. The institute for genomic research Osa1 rice genome annotation database. *Plant physiology* 2005; 138: 18-26.

26. Artimo P, Jonnalagedda M, Arnold K, et al. ExPASy: SIB bioinformatics resource portal. *Nucleic acids research* 2012; 40: W597-W603.

27. Garg VK, Avashthi H, Tiwari A, et al. MFPPI–multi FASTA ProtParam interface. *Bioinformation* 2016; 12: 74.

28. Salvatore M, Shu N and Elofsson A. The SubCons webserver: a user friendly web interface for state-of-the-art subcellular localization prediction. *Protein Science* 2018; 27: 195-201.

29. Hundertmark M, Popova AV, Rausch S, et al. Influence of drying on the secondary structure of intrinsically disordered and globular proteins. *Biochemical and biophysical research communications* 2012; 417: 122-128.

30. Bienert S, Waterhouse A, De Beer TAP, et al. The SWISS-MODEL Repository—new features and functionality. *Nucleic acids research* 2017; 45: D313-D319.

31. Sievers F and Higgins DG. Clustal omega. *Current protocols in bioinformatics* 2014; 48: 3.13. 11-13.13. 16.

32. Tamura K, Stecher G and Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. *Molecular biology and evolution* 2021; 38: 3022-3027.

33. Zhou T, Wang Y, Chen JQ, et al. Genome-wide identification of NBS genes in japonica rice reveals significant expansion of divergent non-TIR NBS-LRR genes. *Molecular Genetics and Genomics* 2004; 271: 402-415.

34. Quinton-Tulloch MJ and Steele KA. SNPs, indels, and microsatellites within and near to rice nbs-lrr resistance gene candidates. *Agronomy* 2021; 11: 2297.

35. Sallaud C, Lorieux M, Roumen E, et al. Identification of five new blast resistance genes in the highly blast-resistant rice variety IR64 using a QTL mapping strategy. *Theoretical and Applied Genetics* 2003; 106: 794-803.

36. Ma J, Lei C, Xu X, et al. Pi64, encoding a novel CC-NBS-LRR protein, confers resistance to leaf and neck blast in rice. *Molecular Plant-Microbe Interactions* 2015; 28: 558-568.

37. Wang GL and Zhou B. Functional and evolutionary analysis of the Pi2/9 locus in rice. In: *Advances in Genetics, Genomics and Control of Rice Blast Disease* 2009, pp.127-135. Springer.

38. Kumar P, Pathania S, Katoch P, et al. Genetic and physical mapping of blast resistance gene Pi-42 (t) on the short arm of rice chromosome 12. *Molecular breeding* 2010; 25: 217-228.

39. Yu H, Shahid MQ, Li R, et al. Genome-wide analysis of genetic variations and the detection of rich variants of NBS-LRR encoding genes

in common wild rice lines. *Plant Molecular Biology Reporter* 2018; 36: 618-630.

40. Qian LH, Wang Y, Chen M, et al. Genome-wide identification and evolutionary analysis of NBS-LRR genes from *Secale cereale*. *Frontiers in Genetics* 2021; 12: 771814.

41. Singh S, Chand S, Singh NK, et al. Genome-wide distribution, organisation and functional characterization of disease resistance and defence response genes across rice species. *Plos one* 2015; 10: e0125964.

42. Kumar IS, Amran NA and Nadarajah K. In silico analysis of qbfr4 and qlbl5 in conferring quantitative resistance against rice blast. *Journal of Pure & Applied Microbiology* 2018; 12.

43. Terensan S, Fernando HNS, Silva JN, et al. In silico molecular and morphological analysis of rice blast resistant gene Pi-ta in Sri Lankan rice germplasm. *Journal of Genetic Engineering and Biotechnology* 2021; 19: 1-12.

44. Belkhadir Y, Subramaniam R and Dangl JL. Plant disease resistance protein signaling: NBS–LRR proteins and their partners. *Current opinion in plant biology* 2004; 7: 391-399.

45. Karshikoff A, Nilsson L and Ladenstein R. Rigidity versus flexibility: the dilemma of understanding protein thermal stability. *The FEBS journal* 2015; 282: 3899-3917.

46. Gamage DG, Gunaratne A, Periyannan GR, et al. Applicability of instability index for in vitro protein stability prediction. *Protein and peptide letters* 2019; 26: 339-347.

47. Mizuno H, Katagiri S, Kanamori H, et al. Evolutionary dynamics and impacts of chromosome regions carrying R-gene clusters in rice. *Scientific reports* 2020; 10: 872.

48. Yang S, Gu T, Pan C, et al. Genetic variation of NBS-LRR class resistance genes in rice lines. *Theoretical and applied genetics* 2008; 116: 165-177.

Figure Legends:

Figure 1: Prediction of sub-cellular localization of eight variants of NBS-LRR protein documented in present study using Cello2 tool

C: cytoplasm, N: nucleus, PM: plasma membrane, E: extracellular, M: mitochondria, P: peroxisomal, CP: chloroplast, G: golgi bodies, ER: endoplasmic reticulum, L: lysosome, CK: cytoskeleton, V: vacuole

Figure 2: Prediction of secondary structure of eight variants of NBS-LRR protein documented in present study using SOPMA tool

Figure 3: Prediction of 3D configuration of eight variants of NBS-LRR protein documented in present study using Swiss-Model

Figure 4: Phylogenetic tree construction to predict evolutionary relationship of eight variants of NBS-LRR protein documented in present study using MEGA11 software

Neighbour-joining tree was constructed with bootstrap value of 100

Table 1: Physicochemical attributes of eight variants of NBS-LRR protein in *O. sativa* predicted using ProtParam tool

#	Protein ID	No. of amino acids	Molecular weight	pI	Extinction coefficient ($M^{-1}cm^{-1}$)	Instability index	Aliphatic index	GRAVY
1	NBS-LRR1	2574	213008.80	4.90	32625	42.33	27.23	0.703
2	NBS-LRR2	4479	369878.06	4.77	57875	46.43	30.39	0.789
3	NBS-LRR3	2709	223861.65	4.89	34250	41.21	26.73	0.691
4	NBS-LRR4	4443	364192.79	4.77	57750	45.73	28.56	0.754
5	NBS-LRR5	3102	253286.46	4.87	36375	38.58	29.56	0.714
6	NBS-LRR6	3243	267588.43	4.85	40250	41.66	28.68	0.724
7	NBS-LRR7	3825	308946.75	4.80	50875	44.83	32.05	0.859
8	NBS-LRR8	3096	252928.25	4.87	36250	39.33	30.20	0.728

Table 2: Secondary structure of eight variants of NBS-LRR protein in *O. sativa* predicted using SOPMA tool

#	Protein ID	Alpha helix (%)	Extended strand (%)	Beta turn (%)	Random coil (%)
1	NBS-LRR1	54.73	11.20	3.27	30.81
2	NBS-LRR2	51.54	10.19	1.88	36.39
3	NBS-LRR3	53.99	12.75	2.33	30.93
4	NBS-LRR4	50.68	9.93	4.53	34.86
5	NBS-LRR5	56.82	9.97	2.71	30.49
6	NBS-LRR6	53.98	10.83	3.06	32.13
7	NBS-LRR7	51.96	10.91	5.65	31.48
8	NBS-LRR8	56.80	9.51	2.82	30.87

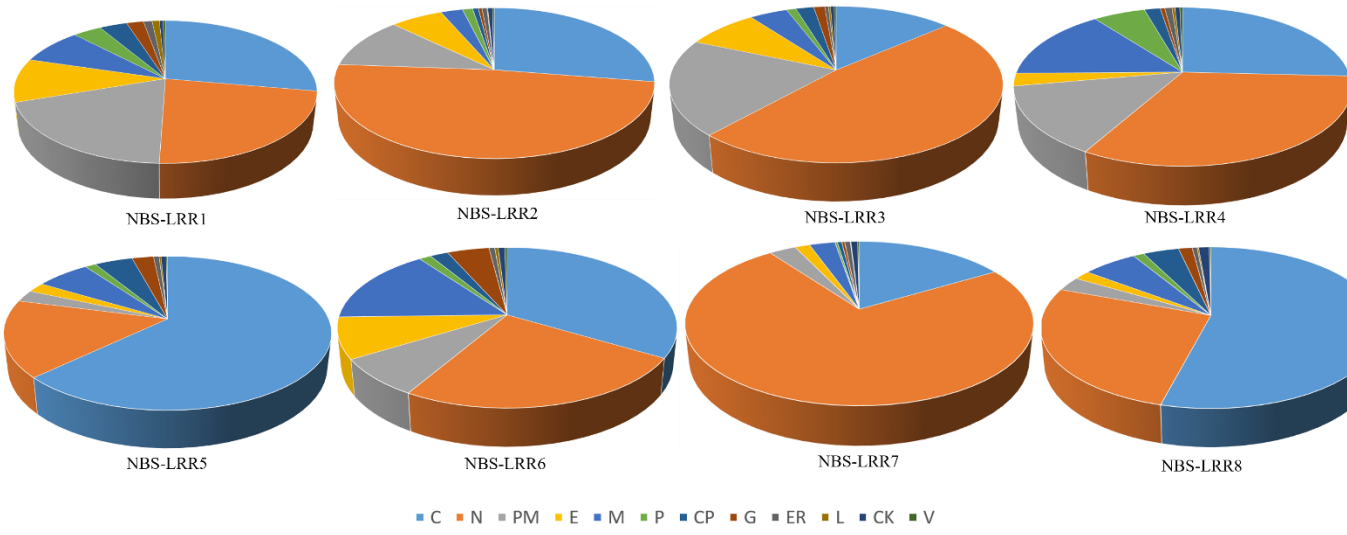
NBS-LRR1
NBS-LRR2
NBS-LRR3
NBS-LRR4
NBS-LRR5
NBS-LRR6
NBS-LRR7
NBS-LRR8
C N PM E M P CP G ER L CK V

Fig 1.

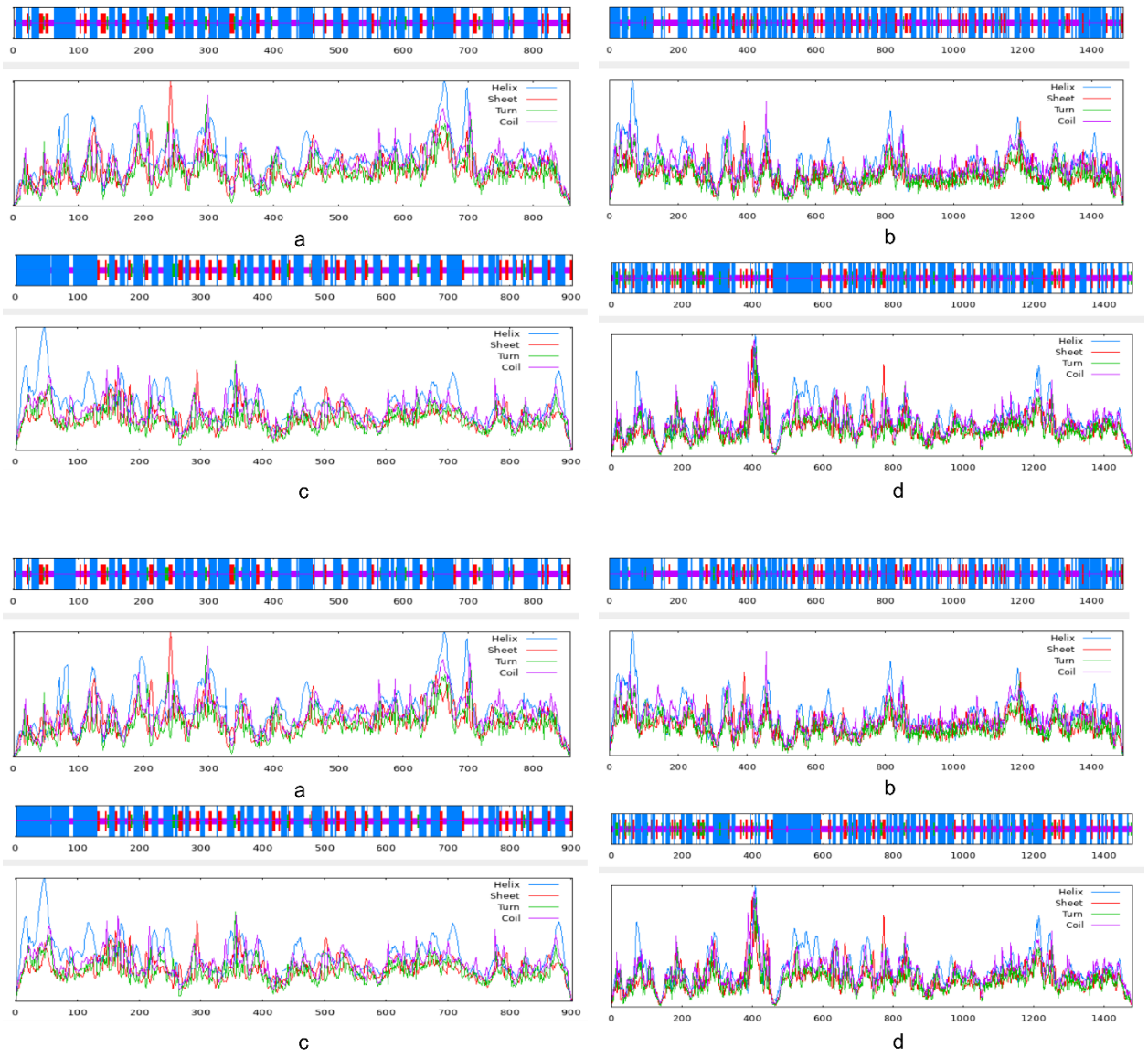

Fig 2

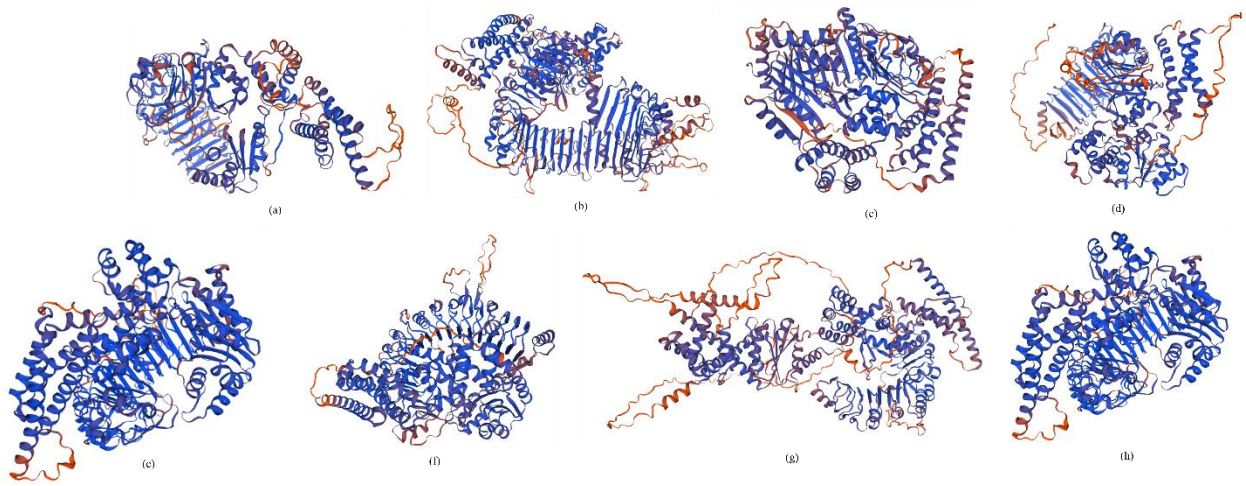

Fig 3

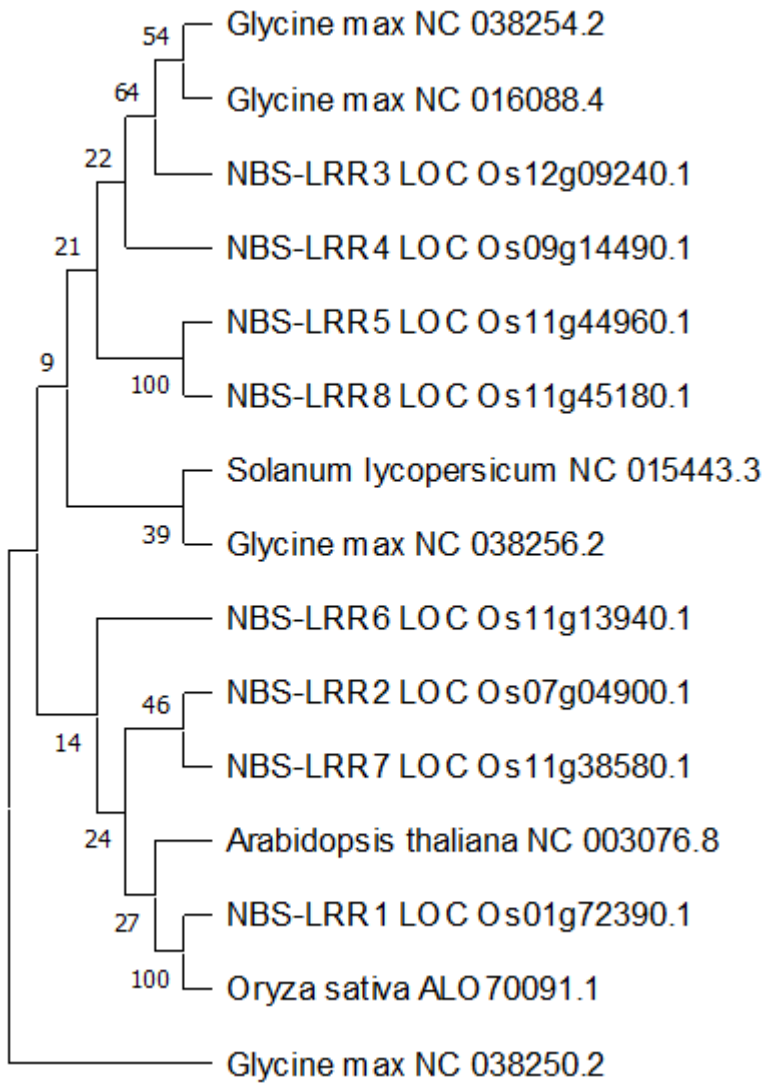
54
Glycine max NC 038254.2
64
Glycine max NC 016088.4
22
NBS-LRR3 LOC Os12g09240.1
21
NBS-LRR4 LOC Os09g14490.1
NBS-LRR5 LOC Os11g44960.1
9
100
NBS-LRR8 LOC Os11g45180.1
Solanum lycopersicum NC 015443.3
39
Glycine max NC 038256.2
NBS-LRR6 LOC Os11g13940.1
46
NBS-LRR2 LOC Os07g04900.1
14
NBS-LRR7 LOC Os11g38580.1
24
Arabidopsis thaliana NC 003076.8
27
NBS-LRR1 LOC Os01g72390.1
100
Oryza sativa ALO70091.1
Glycine max NC 038250.2

Fig 4