Proceedings of 22nd Agricultural Research Symposium (2024) 336-340

# Exploration of Blast Resistant Genes and their Allelic Determinant in Sri Lankan Rice Varieties using their Domain Structures for Utilization in Breeding

T. WIJESINGHE[1], S. TERENSAN[2] and N.S. KOTTEARACHCHI[1]

[1]*Department of Biotechnology, Faculty of Agriculture and Plantation Management, Wayamba University of Sri Lanka, Makandura, Gonawila (NWP),60170, Sri Lanka*

[2]*Department of Agricultural Biology, Faculty of Agriculture, University of Jaffna, Ariviyal Nagar, Kilinochchi, Sri Lanka*

## ABSTRACT

**Numerous rice blast resistance (R) genes have been identified, and a comprehensive understanding of their underlying mechanisms was deemed essential. Therefore, three major genes (*Pb1*, *Pi21*, and *Pita*) associated with rice blast (BL) were studied using genome sequences of 32 Sri Lankan rice varieties. The domain structure of each gene, *Pb1*, *Pi21*, and *Pita* of each variety was comprehensively analyzed in comparison to the three resistant varieties, *Modan*, *Owarihatamochi*, and *Tetep*. Results revealed that *Alagusamba* and *Mudaligawee* could be possessed relatively a stronger resistance compared to other varieties, due to the presence of complete pathogen recognition LRR domain which is necessary for triggering innate immunity. The resistant allele of the *Pb1* gene is reported to contain two deletions with 21 bp and 48 bp giving rise to a low amount of proline amino acids, but such lengthy deletions could not be found in any of the Sri Lankan varieties. Few Sri Lankan varieties contained alanine at the 918$^{th}$ amino acid position which is considered as a feature of a resistant allele of the *Pita* gene. The results of the characterization of 32 Sri Lankan rice varieties regarding three blast-resistant genes would be useful for breeding rice for blast resistance**.

**KEYWORDS:** Amino acid polymorphism, Blast, LRR domains, *Magnaporthe oryzae*, Molecular breeding

## INTRODUCTION

According to the Department of Agriculture (DOA) Sri Lanka, rice blast (BL) disease which is caused by the fungal pathogen *Magnaporthe oryzae* is able to cause a significant yield loss of about 60%. The northern province of Sri Lanka, renowned for its substantial role in rice production within the country, experienced a severe outbreak of blast during the *Maha* season of 2016/2017, resulting in significant economic losses during the period spanning from 2013 to 2016 the development and adoption of resistant cultivars (*Terensan et al*., 2022).

Among the cloned and well-characterized rice blast resistance (R) genes, the majority are nucleotide-binding site (NBS) and leucine-rich repeat (LRR) proteins, which are defined by their NBS and LRR domains. Of the three major domains, the coiled-coil (CC) domain is involved in protein-protein interactions or localization within the cell. The NBS domain is associated with signal transduction, and the LRR domain is crucial for pathogen recognition. Specifically, the LRR domain (leucine-rich repeat) plays a pivotal role in identifying pathogen-specific effectors. Notably, LRR domains are highly variable and are subject to diversifying selection (Devi *et al*., 2016).

Therefore, identifying the diversity of major domains in Sri Lankan germplasm is crucial, as it will provide insights into the molecular mechanisms that confer resistance to rice blast disease. This knowledge is essential for developing more effective strategies to enhance resistance in rice varieties. Given the limited studies on domain diversity in Sri Lankan germplasm, this research aimed to address this gap by examining the different mutations in Sri Lankan rice accessions based on the domain structure of resistance genes.

The NRC-16-16 project has revealed the whole genome sequences of At 354 and Bg 352 Sri Lankan rice varieties by next-generation sequencing using Illumina Hiseq 2500 100PE platform and mapped to Nipponbare and R498 genomes (Abhayawickrama *et al*., 2020). This study used these two rice varieties, At 354 and Bg 352 together with 30 Sri Lankan rice accessions reported in the International Rice Research Institute (IRRI), Philippines to find out the rice blast-resistant gene diversity in comparison with wild varieties reported to have R genes as reference varieties so that the information could be used in breeding programmes.

## METHODOLOGY

### *Identification of Genes Related to Rice Blast using Previously Reported QTL Maps*

Previously reported QTL maps were used to gather information on 18 cloned R Genes of blast stress resistance. They are, *Os01g0149500* (*Pit*), Os01g0781200 (*Pi64*), Os01g0781700 (*Pi37*), Os01g0782100 (*Pi35*), Os04g0401000 (*P21)*, Os04g0620950 (*Pi63*), Os06g0286700 (*Pi9*), Os06g0287500 (*Pid4*), Os06g0330100 (*Pid3*), Os06g0494100 (*Pid2*), Os08g0150150 (*Pi36*), Os09g0327600 (*Pi5*), Os09g0328951 (*Pi56*), Os11g0639100 (*Pi54*), Os12g0281300 (*Pita*), Os12g0285100 *(Ptr*), Os11g0598500 (*Pb1*), Os11g0675200 (*Pb3*). Among them, three genes were selected for further analyses based on the highest no of publications. They are, Os12g0281300 (*Pita*), Os11g0598500 (*Pb1*) and Os04g0401000 (*P21*).

### *Retrieving Gene Sequences of 30 Sri Lankan Rice Varieties*

Relevant BAM files of 30 Sri Lankan rice accessions were downloaded from Rice SNP SEEK database (https:snp-seek.irri.org/).

### *Extraction of Gene Sequences*

Sequences of the selected four genes were extracted from 30 rice varieties available in the Rice SNP Seek database and the gene sequences of At 354 and Bg 352 rice varieties were obtained from NRC 16-16 project (Abhayawickrama *et al*., 2020). The gene sequence of the wild variety (*Modan*) was taken from a published journal paper and the gene sequence of *Tetep* (GQ918486.1), gene sequence of *Owarihatamochi* (AB430853.1) was taken from the NCBI database as reference gene (Hayashi *et al*., 2010).

### *Multiple Sequence Alignment*

The chromosomal locations of each gene (Nipponbare) were identified from the RAP-DB database (https://rapdb.dna.affrc.go.jp/) with correct complementary gene arrangements (+/-). A multiple sequence alignment was carried out using the genes screened with variants as follows. The relevant BAM file was opened in UGENE version 40.1, and the sequences of the respective genes of At 354 and Bg 352 were copied. Then, sequences obtained from the SNP SEEK database and sequences of At 354 and Bg 352 were aligned together using MEGA 11 software.

### *Aligning of the cDNA Sequences*

The cDNA sequences of At 354, Bg 352 and 30 rice accessions were derived by removing the intron regions from their respective genomic sequences. Subsequently, multiple sequence alignment was conducted using the constructed cDNA sequences of each gene. Genes that exhibited variants in exons were then selected for further analysis.

### *Constructing Amino Acid Sequences*

The genes of At 354, Bg 352 and 30 rice accessions that exhibited variants in exons were used to construct amino acid sequences. The open reading frames (ORFs) were obtained using the Translate tool on Expasy (https://web.expasy.org/translate/), and the longest ORF was selected to derive the amino acid sequence. Subsequently, multiple sequence alignment was conducted using the amino acid sequences.

### *Identification of Functional Domains*

The location of specific domain regions was identified by using pfam software (http://pfam.xfam.org/).

### *Screening of Mutations in Sri Lankan Rice Varieties and Phylogenetic Analysis*

Mutations were screened among 32 Sri Lankan rice varieties in comparison to reference varieties used for each gene. Phylogenetic tree for the *Pita* gene was created by using nucleotide sequences of 32 rice varieties including the gene sequence of *Tetep* wild rice variety.

## RESULTS AND DISCUSSION

### *Analysis of Pb1 Gene*

Table 1 shows the analysis of *Pb1* gene with respect to the reference gene of the *Modan* variety. The Pb1 gene encodes a protein consisting of 1296 amino acids, encompassing the CC domain (residues 1-44), NBS domain (residues 486-905), and LRR domain (residues 928-1296) (Hayashi *et al*., 2010).

Accordingly, when that gene screening of the Sri Lankan rice accessions, it was observed that none of these accessions contained the full length of the protein sequence with a maximum length of 1290 amino acids. When comparing these accessions with the wild variety *Modan*, *Podiwee* and At 354 exhibited identical amino acids at positions 6, 14, 17, and 41 within the CC domain.

**Table 1. The analysis of *Pb1* and *Pi21* genes in comparison with *Modan* and *Owarihatamochi***

| Accession | | Pb1 | | | | | | Pi21 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | CC domain(aa) 1-44 | | | | NBS (aa) 486 - 905 | LRR (aa) 928 - 1296 | MBD (1-68) | Proline-rich domain (bp) Indels | |
| **Amino Acid Positions** | | **6** | **14** | **17** | **41** | | | **7** | **238-273** | **421-468** |
| *Modan* | | S | V | R | N | P | P | | | |
| *Owarihatamochi* | | | | | | | | L | 21 | 48 |
| *Alagusamba* | (IRIS 313-10610) | S | V | R | H | P# | P# | L | 7 | - |
| *Mudaligawee* | (IRIS 313-11916) | S | V | R | H | P# | P# | S | 9 | - |
| *Balasooriya* | (IRIS 313-8699) | S | V | H | N | NP | NP | L | 14 | 2 |
| *Chandina* | (*IRIS 313-9917)* | S | V | H | N | NP | NP | S | 8 | - |
| *Godawel* | (IRIS 313-10722) | A | V | R | N | NP | NP | L | 1 | 13 |
| *Hondarawala* | (IRIS 313-10020) | S | V | G | N | NP | NP | L | 13 | - |
| *Kahatawee* | (IRIS 313-10662) | S | V | G | N | NP | NP | L | 14 | 2 |
| *Kaluilankayan* | (IRIS 313-8968) | S | E | G | H | NP | NP | L | 4 | 15 |
| *Karuthaseenati* | (IRIS 313-10718) | S | V | G | N | NP | NP | L | 3 | 2 |
| *Kurkarupan* | (IRIS 313-9861) | S | V | R | N | NP | NP | S | 4 | - |
| *Kuruluwee white* | (IRIS 313-8342) | S | V | G | H | NP | NP | L | 7 | 14 |
| *Kuruluthuda* | (IRIS 313-8925) | S | V | G | N | NP | NP | L | 7 | - |
| *Mahatholuwa* | (IRIS 313-10609) | S | V | G | N | NP | NP | S | 7 | - |
| *Modaikaruppan* | (IRIS 313-9862) | S | V | G | N | NP | NP | L | 3 | 2 |
| *Muttu samba* | (IRIS 313-9039) | S | V | G | N | NP | NP | L | 4 | 2 |
| *Pachchaperumal* | (IRIS 313-10476) | S | V | G | N | NP | NP | S | 7 | - |
| *Pannithi* | (IRIS 313-9949) | S | V | H | N | NP | NP | L | 7 | 2 |
| *Periyawellei* | (IRIS 313-10717) | S | V | G | N | NP | NP | L | 9 | - |
| *Podiheenati* | (IRIS 313-9639) | S | V | G | N | NP | NP | L | 4 | - |
| *Podi wee* | (IRIS 313-9831) | S | V | R | N | NP | NP | L | 3 | - |
| *Pokkali* | (IRIS 313-8244) | S | V | G | N | NP | NP | S | 9 | 14 |
| *Puttunellu* | (IRIS 313-9969) | S | V | G | H | NP | NP | S | 9 | - |
| *Raceperumal* | (IRIS 313-9970) | S | V | G | N | NP | NP | S | 9 | - |
| *Ranruwan* | (IRIS 313-1191) | S | V | G | N | NP | NP | L | 12 | - |
| *Samba* | (IRIS 313-10661) | S | V | G | N | NP | NP | L | 9 | 28 |
| *Sigardis* | (IRIS 313-9867) | S | V | H | N | NP | NP | S | 9 | - |
| *Sinnasithirakale I* | (IRIS 313-9091) | S | V | G | N | NP | NP | L | 1 | - |
| *Siththiyankottaisamba* | (IRIS3139936) | S | V | H | N | NP | NP | L | 4 | - |
| *Welleikolomban* | (IRIS 313-10719) | S | V | G | N | NP | NP | S | 9 | - |
| At 354 | | S | V | R | N | NP | NP | L | 3 | - |
| Bg 352 | | S | V | G | N | NP | NP | L | 3 | - |

*P: Domain Present NP: Domain Not Present P#: Domain Present with mutations, MBD: Metal Binding Domai*

In Pb1-mediated blast resistance, the presence of amino acids at 16(I), 23(L), 30(V) and 37(M) positions of the CC domain are crucial for the induction of the Salicylic acid signalling pathway. Because, WRKY-45, a transcription factor binds through the CC Domain in order to start the defence signalling (Inoue *et al*., 2013). Accordingly, all Sri Lankan accessions included those four amino acids without any mutation. In terms of the NBS and LRR domains, only the *Alagusamba* and *Mudaligawee* accessions possessed these two domains. However, these two accessions exhibited numerous mutations in the NBS and LRR domains compared to the wild variety. Consequently, in terms of Pb1-mediated blast resistance, it is suggested that *Alagusamba* and *Mudaligawee* may exhibit a certain level of resistance relative to the other accessions. Nonetheless, further studies, such as functional analyses, are required to substantiate this finding.

***Analysis of Pi21 Gene***

The *Pi21* gene encodes a protein featuring a heavy metal binding domain (1-68) and proline-rich region (Fukuoka *et al*., 2009). Fukuoka and Okuno (2019) have reported that resistant varieties have 21 bp and 48 bp

deletions in proline-rich regions giving evidence for being a blast resistance gene. The susceptible variety does not have deletions in the proline-rich region. However, it is considered that having two deletions specifically on both sides from 238-273 bp and 421-468 bp resulting loss of function of *Pi21* making disease resistant by reducing the proline amino acid content (Fukuoka *et al*., 2009). *Owarihatamochi* is considered as the resistant variety and *Aichi-Asahi* is considered as the susceptible variety (Fukuoka *et al*., 2009). According to the level of deletions in Sri Lankan rice accessions, *Kaluilankayan, Kuruluwee white, Pokkali, Samba, Balasooriya, Godawel* and *Kahatawee* can be considered as resistant varieties.

As *Sinnasithirakalei* included only 1 bp deletion, it could be considered as an accession with a susceptible *Pi21* gene. In addition to that, both Bg 352 and At 354 had similar levels of *Pi21* mediated less resistance. The DOA has reported that At 354 and Bg 352 are moderately susceptible and moderately resistant to blast, respectively, which could probably be due to the presence of 3 bp deletion in the same region.

***Analysis of Pita Gene***

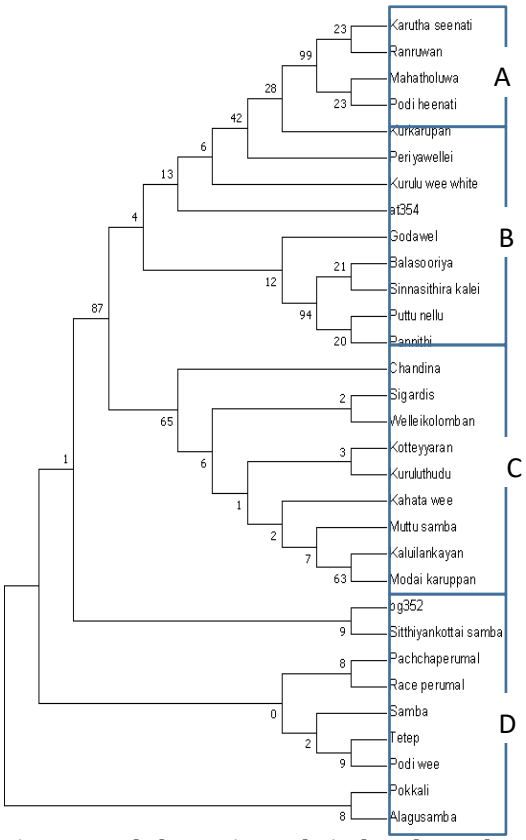


**Figure 1. Phylogenetic analysis based on the nucleotide sequence of Sri Lankan rice accessions in comparison with *Tetep* with bootstrap replications 1000**

From the 32 accessions, only four accessions, *Karuthaseenati, Mahatholuwa, Podiheenati* and *Ranruwan* exhibited mutations at the specified positions with N, C, and F replacing the original amino acids at positions in *Tetep* resistant variety, 711(T), 724(R), and 816(L), respectively and they were included in cluster A (Figure 1). Most accessions in Cluster B comprised varieties that were reported to be susceptible to blast disease (Terensan *et al*., 2021) while Cluster C included varieties that were moderately resistant to rice blast (Terensan *et al*., 2021) (Figure 1).

Notably, all accessions with alanine at the 918$^{th}$ amino acid position which is considered to causing a possible resistance were observed in three different clades mentioned as Cluster D. They are *Siththiyakottaisamba, Raceperumal, Podiwee, Pachchaperumal, Alagusamba, Samba* and Bg 352 variety (Figure 1).

When evaluating the resistance characteristics of Sri Lankan rice accessions based on domain architecture, it was observed that most rice varieties exhibit mutations in the LRR domains. This trend was consistent across all three genes examined. In contrast, the amino acids within the CC domain remained constant, exhibiting similar residues to those found in wild varieties.

## CONCLUSIONS

In Pb1-mediated resistance, based on the LRR domain, *Alagusamba* and *Mudaligawee* could be possessed relatively stronger resistance compared to other varieties, due to the availability of almost complete pathogen recognition domain which is the crucial region for initiating defence response. Conversely, in Pi21-mediated resistance, the varieties *Kuruluwee white, Kaluilankayan, Samba*, and *Pokkali* demonstrated possible higher resistance than the other accessions according to the *in silico* analysis. In Pita-mediated blast resistance, *Samba, Alagusamba, Pokkali, Pachchaperumal* and *Siththiyamkottaisamba* showed possible resistance. Bg 352 was reported as a variety with moderate resistance and it could be proved due to the presence of alanine amino acid at 918$^{th}$ position in the *Pita* gene.

When considering both *Pb1* and *Pita* genes *Alagusamba* could be considered to possess the most resistance trait to blast, but further studies on phenotypic assessment are necessary to confirm the results.

**ACKNOWLEDGEMENTS**

Authors acknowledge the NRC 16-16 project for providing whole genome sequences of At 354 and Bg 352.

**REFERENCES**

Abhayawickrama, B., Gimhani, D., Kottearachchi, N., Herath, V., Liyanage, D. and Senadheera, P. (2020). *In silico* identification of QTL-based polymorphic genes as salt-responsive potential candidates through mapping with two reference genomes in rice. *Plants*, **9**(2),15-20.

Devi, S.J.S.R., Umakanth, B., Vishalakshi, B., Aglawe, S.B., Babu, V.R. and Madhav, M.S. (2016). Insights on diversity of leucine-rich repeat (LRR) domain among major blast resistance genes of Rice. *Journal of Genomes and Exomes*, **5**, 1-7.

Fukuoka, S. and Okuno, K. (2019). Strategies for breeding durable resistance to rice blast using pi21. *Crop Breeding, Genetics and Genomics*, **1**(2).

Fukuoka, S., Saka, N., Koga, H., Ono, K., Shimizu, T., Ebana, K., Hayashi, N., Takahashi, A., Hirochika, H., Okuno, K. and Yano, M. (2009). Loss of function of a proline-containing protein confers durable disease resistance in rice. *Science*, **325**(5943), 998-1001.

Hayashi, N., Inoue, H., Kato, T., Funao, T., Shirota, M., Shimizu, T., Kanamori, H., Yamane, H., Hayano-Saito, Y., Matsumoto, T., Yano, M. and Takatsuji, H. (2010). Durable panicle blast-resistance gene *Pb1* encodes an atypical CC-NBS-LRR protein and was generated by acquiring a promoter through local genome duplication. *Plant Journal*, **64**(3), 498-510.

Inoue, H., Hayashi, N., Matsushita, A., Xinqiong, L., Nakayama, A., Sugano, S., Jiang, C.J. and Takatsuji, H. (2013). Blast resistance of CC-NB-LRR protein *Pb1* is mediated by WRKY45 through protein-protein interaction. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(23), 9577-9582.

Terensan, S., Fernando, H.N.S., Silva, J.N., Perera, S. A. C. N., Kottearachchi, N. S. and Weerasena, O.V.D.S.J. (2021). *In silico* molecular and morphological analysis of rice blast resistant gene *Pi-ta* in Sri Lankan rice germplasm. *Journal of Genetic Engineering and Biotechnology*, **19**(1), 163.

Terensan, S., Fernando, N.H.S., Silva, J.N., Perera, S.A.C.N., Kottearachchi, N.S. and Weerasena, O.V.D.S.J. (2022). Morphological and molecular analysis of fungal species associated with blast and brown spot diseases of *Oryza sativa*. *Plant Disease*, **106**(6), 1617-1625.