Analysis of Multiple PCR Bands in Drimia Species Obtained by R-Gene Specific Degenerate Primers

H. M. Hanumantha Rao and L. Rajanna*

Department of Botany, Bangalore University, Jnana Bharathi campus, Bengaluru 560 056.

Abstract: Many studies confirmed that degenerate primers LM638 and LM637 amplify NBS-LRR Type R gene encoding sequences in both monocots and dicots. In addition, these primers also amplify nonspecific multiple bands along with specific amplification. The current study aims to explore the structure, role and reason for nonspecific bands in the amplification. The genomic DNA-degenerate PCR method was used to isolate the NBS-LRR type R-genes in Drimia species (synonym Urginea). Agarose gel electrophoresis was carried out for separation and visualization of amplified bands. The kit-based gel extraction method was followed to purify the PCR bands. Purified PCR bands were subjected to TA cloning. The plasmids which contain desired PCR bands were sequenced using the Sanger sequencing method. Multiple sequence alignment tool CLUSTALW was used to align sequences and MISA-Web tool was used to identify simple sequence repeats (SSR). The results of the study revealed that multiple sequence analysis of PCR bands of 600bp/500bp showed random deletions, presence of indels and frameshift mutations in the alignment. The sequence alignment also showed monomorphic sites between two sequences. Both 600 and 500bp were devoid of simple sequence repeats and the ORF finder confirmed the presence of irregular stop codons in the 600 bp sequences. The multiple sequence alignment tool was used in disclosing the structure and ancestor relationship between aligned Drimia sequences. The current study would help to explore pseudo genes in the genome of Drimia and to differentiate the orthologous and paralogous relationships in the sequences.

Key words: Conserved domain markers (CDMs), Degenerate PCR, NBS-LRR Type R-genes, Functional marker, Sanger sequencing, MSA, CLUSTALW

*Corresponding Author
L. Rajanna, Department of Botany, Bangalore University, Jnana Bharathi campus, Bengaluru 560 056.

Received On 1 March, 2023
Revised On 1 June, 2023
Accepted On 22 June, 2023
Published On 1 November, 2023

Funding: This research did not receive any specific grant from any funding agencies in the public, commercial or not for profit sectors.


This article is under the CC BY-NC-ND Licence (https://creativecommons.org/licenses/by-nc-nd/4.0/)

1. INTRODUCTION

Plants faced many environmental challenges which include both biotic and abiotic factors. In evolution plants have a well adaptive mechanism to withstand environmental stresses. The adaptive mechanisms which can be regulated by signaling cascades and gene transcription networks. The production of reactive oxygen species, plant hormones, R- gene mediated resistance and physical barriers are important mechanisms involved in plants to adopt environmental stress. Similarly, Nucleotide binding site-Leucine rich repeats (NBS-LRRs) type R gene and proteins acts as adaptable guards in plants to defend against various diseases caused by diverse pathogens. The different research groups concluded that Plant NBS-LRR proteins act through a network of signaling pathways and induce a series of plant defense responses. Drimia (synonym Urgines) species is a wild medicinal plant and is also called as wild onions or sea onions. Segregation of this genus becomes a more complex process because of its morphological variations. Hence, many studies have been conducted to explore its molecular phylogeny by using Random amplified polymorphic DNA (RAPD), Sequence related amplified polymorphism (SRAP), Internal transcribed spacer (ITS) and matK gene sequence data. As a result, presence of distinct subclades and intraspecies relationship was confirmed. Apart from these molecular phylogenetic studies, other phytochemical studies have confirmed the presence of antifungal protein which has a broad range of antifungal activity against plant pathogens. As supportive evidence there are no documents which confirm plant pathogens which are affecting Drimia species. However, one study has identified this genus as host for *alfexivirus* which can infect onion and garlic crops. Based on the current knowledge and literature survey, biotic stress is very less in Drimia species and the plant must be evolved with better disease resistance mechanisms against fungal and bacterial pathogens. Many important plant disease resistance genes (R-genes) come under NBS-LRR class genes and its proteins. Hence, our previous study was aimed to explore the diversity of disease resistance mechanisms in Drimia. The Nucleotide binding sites – lucine rich repeats R genes and its sequence diversity was analyzed by using Conserved domain specific Degenerate PCR method. As a result of these comparative studies, it has confirmed that the Drimia species NBS-LRR coding nucleotides comprise components to fight against both fungal and bacterial pathogens. Moreover, the Kinase-2 Motif of Drimia NBS encoding nucleotides contains extra Leucine amino acid which is absent in other characterized R genes. Overall phylogenetic analysis has confirmed that isolated NBS encoding sequences are highly diverse and grouped in separate clades. These studies give further hints on how to use these genetic resources for conducting functional studies in model plants. In a previous study, the results of agarose gel electrophoresis showed multiple banding patterns in both *Drimia* species Viz. *Drimia wightii* and *Drimia indica*. However, these studies presented only those sequences which had homology with NBS-LRR specific protein as per the study objective. But the structure, role and relationship of nonspecific DNA bands were not explored in detail. The multiple banding pattern is a major problem in highly degenerate PCR methods. Visually, these multiple banding patterns are confusing with heterozygosity in the genome and bias the results unless these bands are further analyzed by sequencing. Further investigations are required to confirm the relationship between multiple bands in *Drimia* species. The primers (LM638 and LM637) used in the present study were used in many crop plants and the nonspecific PCR bands along with specific band was common in all studies however most of the studies ignored the nonspecific PCR band and targets only NBS—LRR specific band. Hence, during the present investigation, the study intended to explore the importance and cause of nonspecific bands in the PCR amplification. Therefore, the current study aims to isolate, analyze and confirm the relationship between multiple DNA bands in *Drimia* species which are generated by NBS-LRR conserved DNA degenerate PCR method. The basic and advanced bioinformatic web tools are highlighted in this study as an important tool to explore the different aspects of multiple bands. This technique is helpful for identifying pseudogenes, orthologous and paralogous relationships and also to check the heterozygosity in *Drimia* species.

2. MATERIALS AND METHODS

The Sample collection, Genomic DNA Polymerase chain reaction and gel electrophoresis were performed as per standard protocol used in previous study.

2.1. Sample collection

Two species of *Drimia* such as *D.indica* and *D.wightii* were collected from Tataguni estate, Agara village, Bengaluru Rural District and Kolar dry land forest, Kolar respectively by undertaking frequent field trips and were maintained in the Botanical garden, Department of Botany, Bangalore University, Bengaluru. The collected plants were identified with the help of Gamble flora and authenticated by Dr. K Gopala Krishna Bhat, Taxonomy Research Centre, Department of Botany, Poornapraga College Udipi, Karnataka. The voucher specimen (Voucher No. Bot. 122) was deposited in the herbarium of the Department of Botany, Bangalore University, Bangalore.

2.2. Genomic DNA extraction and isolation and quantification

The leaves of *D.wightii* and *D.indica* were collected between March-April, frozen immediately in liquid nitrogen and stored in -70 °C deep freeze till further use. The GeneJET Plant genomic DNA purification mini kit (Thermo Scientific, USA. Catalog No: K0791) was used for extracting genomic DNA, and protocol was followed. Initially 100 mg of plant tissue was ground in the presence of liquid nitrogen using autoclaved mortar and pestle. Samples are lysed in supplied Lysis Buffers in the presence of RNase A. Proteins and polysaccharides are removed by Precipitation Solution. The lysate is then mixed with the Plant gDNA Binding Solution, ethanol and loaded on the purification column where the DNA binds to the silica membrane. Impurities are effectively removed by washing the column with the prepared wash buffers. Genomic DNA is then eluted under low ionic strength conditions with the Elution Buffer (Pub. No. MAN0016131). The eluted DNA was directly quantified using Nanodrop Spectrophotometer (Thermo scientific, USA) and integrity was checked by resolving 7µl eluted DNA sample in 0.8% agarose gel and electrophoresis was carried out at 100V for 30 minutes. The concentrated DNA samples were stored at -20 °C in different aliquots with respective labelling to minimize the freeze thaw effect while standardizing the PCR.

2.3. Polymerase chain reaction (PCR) components and program

Genei laboratories Taq DNA polymerase kit with compatible buffer (Bangalore GeNei catalog No: 0601600051730) was used for PCR amplification of NBS-LRR nucleotide sequences in genomic DNA. The degenerate primers LM638 and LM637 used in this study were designed and validated in many crop species. The reaction mix contains 2µl of 10x Taq buffer, 3 µl of 2.5mm dNTP mix, 1.5
µl of 10 picomol forward primer and 1.5 µl of 10 picomol reverse primer, 0.4 µl of Taq polymerase, 9.6 µl of PCR grade Nuclease free Water and 2µl of (50-100ng) of genomic DNA was added separately to respective tubes to make the final volume of 20 µl. The reaction was performed in C1000 Touch Thermal cycler, Bio-Rad, USA. The primers concentrations and annealing temperature were standardized. i.e. 95°C for 10 minutes to initial denaturation, 95 °C for 1 minute to secondary denaturation, 48 °C for 1-minute to annealing, and 72 °C for 1.5-minute to extension with 35 cycles and 72°C for 15 minutes to final extension.

2.4. Assessment of Genetic variation

The genetic variation of D.wightti and D.indica was evaluated through the polymerase chain reaction using Conserved DNA markers (CDMs) which targets disease resistance genes in the plant genome. The two highly degenerate primers (forward and reverse) which anchor the NBS domain were designed by using known disease resistance genes 18. The PCR amplified product was separated by using 1% agarose gel electrophoresis; the gel picture was documented. The PCR was repeated with gradient annealing temperature method to check the stability of amplification. The different sized amplified DNA bands were excised from the agarose using GeneJET Gel Extraction Kit, Thermo scientific, USA (catalog No: K0691). Each PCR fragment was cloned into the plasmid using the pGEM-T Easy vector system, Promega, Germany (catalog No: A1360) and transformed into competent Escherichia coli strain JM109, transformed cells were screened using x-gal containing selection media plates. The transformed colonies (white colonies) were selected and screened for the presence of desired insert using colony PCR. The plasmid of an insert containing colonies was isolated using a GeneJET plasmid isolation kit, Thermo scientific, USA (catalog No: K0502). DNA sequencing was conducted at Eurofins Scientific India Pvt Ltd, Bangalore, India. The primary amplification was conducted using vector specific T7 RNA polymerase promoter Primer, sequencing cycle was performed using big dye terminator v3.1 kit (di-deoxy chain termination method), and capillary electrophoresis was conducted by Using 3730 xL genetic analyzer applied Bio-systems, USA. The DNA sequence quality was checked by using chromos software and vector contamination was removed by using NCBI VecScreen Tool. Further FASTA format sequence was extracted from ABI files by using BioEdit software. DNA Polymorphism in the selected sequence was analyzed by DnaSP version 6 software. The occurrence of microsatellite tandem repeats was checked by using Microsatellite identification tool (MISA-Web) 19, 20. The sequence similarity and percent homology were analyzed by multiple sequence alignment tool CLUSTALW, and NCBI-ORF finder was used to search probable ORFs in the nucleotide sequences.

3. STATISTICAL ANALYSIS

The current study was a comparative study which involved pairwise alignment of two different DNA sequences. Most of the work was carried out using Bioinformatics tools and computer programs.

4. RESULTS

The DNA extraction method used in this study was efficient to isolate DNA from Drimia species. Nanodrop quantification of purified DNA was 160ng/µl and 200ng/ µl in Drimia wightti and Drimia indica respectively. The genomic DNA agarose gel electrophoresis has shown intact DNA bands without smearing which confirmed the DNA was good for PCR amplification. As expected, the genomic DNA amplification using degenerate PCR has confirmed the presence of multiple bands in gel electrophoresis i.e. 500bp, 600bp and 1000bp. 100 base pair difference between 500bp and 600bp was exactly appeared as SSR marker derived heterozygote pattern (fig-1). The agarose gel extraction and purification method followed in this study yielded approximately 20-30ng/ µl. The cloning of PCR purified bands was successful in producing four good DNA sequences. These sequences were matched with previous studies. The two FASTA format sequences of Drimia wightii (600bp/500bp) and two sequences of Drimia indica (600bp/500bp) were used to construct multiple sequence alignment. The results of sequence alignment showed highest homology between two 600bp of different species. Similarly, a good homology was observed in 500bp sequences between these species. However, the diversity was very high between 500bp and 600bp due to the presence of SNP, Indels and frameshift mutations. The Overall alignment excluding primer binding sites 230 monomorphic sites, 242 polymorphic sites and 116bp nucleotide sequence deletion were seen (Fig-2). MISA microsatellite -Web search tool revealed the absence of SSRs in the sequences.
In figure 1, the loading dye was added to the PCR product to see the migration during electrophoresis. After 60% of movement of the loading dye in agarose gel, the picture was documented using the gel documentation system several times to obtain a better resolution. In gel picture lane 1 is *Drimia wightii* PCR product, lane 2 is *Drimia indica* PCR product and lane M is 10kb ladder (Known DNA bands from 100bp to 10000bp), the lane B is no template control or blank to confirm the primer quality and other DNA contamination. The arrows in the picture showing bands with 600bp and 500bp appear like SSR marker derived heterozygous patterns. But one more band at 1000 bp in the amplified product was nonspecific and not included in the study.

In Table 1, showing homology information on selected degenerate primer amplified Bands. *Drimia wightii* and *Drimia indica*, derived 600bp and 500bp FAST sequence were searched for homology against nucleotide databases using the BLASTX program. In the search results the 600bp sequences did not show any similarity with *NBS-LRR* protein sequences. Conversely, the search results of 500bp showed homology between *NBS-LRR* Type R genes in the database and kinase 2 motif of *NBS domain* confirmed by protein motif analysis which were elaborated in our previous studies 16, 14.

Table 1: Homology information on selected degenerate primer amplified Bands.

<table>
<thead>
<tr>
<th>Plant species</th>
<th>600bp PCR Band</th>
<th>500bp PCR Band</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Drimia wightii</em></td>
<td>No similarity with NBS-LRR Type R genes</td>
<td>NBS-LRR Type R genes specific</td>
</tr>
<tr>
<td><em>Drimia indica</em></td>
<td>No similarity with NBS-LRR Type R genes</td>
<td>NBS-LRR Type R genes specific</td>
</tr>
</tbody>
</table>

In Table 1, showing homology information on selected degenerate primer amplified Bands. *Drimia wightii* and *Drimia indica*, derived 600bp and 500bp FAST sequence were searched for homology against nucleotide databases using the BLASTX program. In the search results the 600bp sequences did not show any similarity with *NBS-LRR* protein sequences. Conversely, the search results of 500bp showed homology between *NBS-LRR* Type R genes in the database and kinase 2 motif of *NBS domain* confirmed by protein motif analysis which were elaborated in our previous studies 16, 14.

*Fig 2: Showing multiple sequence alignment results output of CLUSTALW.*
In Figure 2, showing multiple sequence alignment results in the output of CLUSTALW. The FASTA format derived from sequence chromatogram files was used for analysis. The sequence length of 600 bp was 593bp in Drimia wightii and 614bp in Drimia indica, but there is not difference in 500 bp sequence in both Drimia species. In the alignment Drimia wightii sequence labelled as “DW” and Drimia indica sequence labelled as DI. The asterisk * present in the alignment showing matching nucleotides in the alignment. Gaps are showing with dash — symbol.

5. DISCUSSION

A PCR primer sequence is called degenerate if some of its positions have several possible bases 21. In polymerase chain reaction application, it is possible to design primers which can amplify similar DNA fragments across different organisms. Primers can be designed within a conserved region of the target sequences in such a way that the annealing properties of the primers remain relatively unchanged across multiple templates 22. In plant genetics this degenerate PCR method allowed to isolate R- genes from crop plants and their wild relatives. Many important R- genes in plants are NBS-LRR Type R genes which have specific functions in defence mechanisms against pathogens and insects. The degenerate primers LM638 and LM637 used in this study was designed by Kanazin et al (1996) from P-loop and the GLPL motifs of the NBS region of known R-genes from crop plants 16. Laterby using these primers many commercial important plants were screened for R-gene diversity 23, 24, 25. The studies using these primers have confirmed 500bp is a specific amplicon which showed NBS domain specific features in these sequences 23, 24. Interestingly one study used single degenerate primers to amplify the grape plant genomic DNA. The PCR mix with only LM638, only LM637 and both LM638/LM637 primers included and performed PCR and gel electrophoresis. They observed LM638 alone did not amplify any DNA bands but LM637 alone amplified non-specific 600bp band and both primers in the reaction amplified the 500bp which is specific to NBS domain sequences 24. But they considered 600bp is unknown and nonspecific. As per this information it is confirmed that LM637 primers have some role in amplification of the 600bp band in the plant genome. In current study on Drimia species the degenerate primers in PCR mix produced both 600bp and 500bp sequences strongly (Fig-1). Similarly, LM638/LM637 primer combination in Chilli genomic DNA also amplified two strong DNA bands with different size and further no information provided for non-specific band 24. Current study confirmed the previous results that were carried out in various plant species. However, we intended to investigate further on non-specific amplified bands by conducting comparative bioinformatics. In our previous work, NBS-specific degenerate PCR showed the presence of more than two bands in Drimia species (1000bp/600bp/500bp) 14, 16. Both 600bp and 500bp bands were in the expected amplification region. However, only 500bp showed similarities with NBS-LRR Disease resistance family proteins as expected (Table 1). But 100 bp difference between two PCR bands showing exactly like locus specific heterozygous type (Fig 1). The heterozygote allitic patterns are common in Simple sequence repeat sequences (SSR) marker studies across the different species and genotypes. Therefore, further investigation of these bands (600bp/500bp) were selected for pairwise sequence analysis. The other PCR band at 1000bp was considered a nonspecific band because it was not at the expected size and removed from the study. The sequence alignment is given in Fig 2. Two sequences of Drimia wightii (600bp/500bp) and two sequences of Drimia indica (600bp/500bp) were used to construct multiple sequence alignment. The multiple sequence analysis showed that there is no good homology between 600bp and 500bp and the diversity was due to presence of SNPs, indels, frameshift mutations, and good conservation was only observed in the position of forward primer (Fig 2). In addition to this there was a random deletion of 116 bp in the sequence alignment which created a functional 500bp band which shares similarities with disease resistance family proteins. Interestingly, the alignment also showed 230 monomorphic sites which confirms these two sequences (600bp/500bp) have also shared some ancestral relationships. The sequences were also used to search probable simple sequence repeats using the MISA-Web tool; the results showed that there are no sequence repeats in the sequences. The overall results confirmed that these two bands (600bp/500bp) are not derived from a single locus like SSR marker. However, these PCR amplified bands alignment showed paralog relationship. Paralogs are genes in the same organism which evolved by gene duplication during evolution 27, 28, 29. Paralogs after duplication, paralogous proteins experience weaker evolutionary pressure and their specificity diverges leading to the emerging of new specificities and functions 30. Specifically, NBS-LRR genes confer defense against the pathogen infection by multiple duplications and diversification 31, 32. Therefore, the Duplication and divergent phenomena are most comparable to current study 33. However, the 600bp sequence ORF finder results showed that these sequences contain premature stop codons which further clues probable pseudogene structure. Pseudogenes are also a paralog gene generated from ancestral functional genes during evolution by gene duplication or retro transposition. Which contains critical defects in their sequences, such as lacking a promoter, having a premature stop codon or frameshift mutations. Generally, pseudogenes are functionless, but recent evidence demonstrates that some of them have potential roles in gene regulation 34. This study is limited to small sequence data which is derived from PCR bands, and the sequence information and data is specific to Drimia species; the same conclusion will not be the same for other plants. However, this study throws light on nonspecific bands in degenerate primer PCR. Species of Drimia are an important medicinal plant which have cardio tonic properties 35. In addition, genetic diversity and molecular studies are vital in discriminating sub sub-populations in Drimia 11. The multiple PCR bands in the species leads to misperception and hinders the genotyping 36. The polyploid nature of plant genomes made it difficult for the analysis of variation within genes 35. In general, compared to wild plants, crop plants are more prone to pests and disease because a large number of plants with uniform genetic background helps pathogens to spread rapidly 36. Wild relatives of crop plants may have developed different mechanisms to defend against biotic challenges and well adapted to climatic changes 37, 38. In general, most of the commercially important crop plants were explored for their diversity in disease resistance mechanisms. But other wild plants like the Drimia species were ignored. To develop a novel approach in disease resistance mechanism more genetic resources from different sources are required. Development of disease resistance crop plants by using biotechnology application is the future of plant breeding programs 40.

6. CONCLUSION

Multiple sequence alignment technique is an important tool in exploring the different aspects of the nucleotide sequences. The absence of simple sequence repeats and presence of indels, frameshift mutations, missing sequences were shown in the sequence alignment. This further confirms that heterozygous allitic patterns in the PCR bands (600bp/500bp) are not due to presence of SSR or ortholog genes. Gene duplication and rearrangements created a paralog in the genome. Further, the presence of immature stop codon in the nonspecific sequence is confirmed by the presence of pseudogene families created from ancestor functional genes by gene duplication in two Drimia species.
I am thankful to my Guide Professor L. Rajanna for supporting and completing this study by providing timely suggestions and guidance.

Conflict of interest declared none


