International Journal of Botany Studies
www.botanyjournals.com
ISSN: 2455-541X
Received: 04-01-2022, Accepted: 19-01-2022, Published: 04-02-2022
Volume 7, Issue 2, 2022, Page No. 168-178

Isolation and comparative analysis of NBS-LRR disease resistance gene analogs in medicinal plant *Drimia wightii. L*

Hanumantha Rao H M[1], L Rajannna[2], S Prabhakara[3]

[1] Department of Botany, Jnana Bharathi Campus, Bangalore University, Bangalore, Karnataka, India

[2] Professor, Department of Botany, Jnana Bharathi Campus, Bangalore University, Bangalore, Karnataka, India

[3] Senior Research Scientist, Central Research Lab, Raja Rajeswari Medical College and Hospital, Bangalore, Karnataka, India

Abstract
The NBS-LRR Type R- gene diversity is well studied in crop plants but there are very little data available in non-domesticated wild medicinal plants. This study explored the diversity of NBS-LRR Type R genes in the wild medicinal plant *Drimia wightii.* L and compared with characterized NBS Type R genes to understand the organization of R genes in wild plant genome. The NBS encoding sequences were isolated using degenerate PCR and TA cloning methodology. Sequence homology was conducted by using the NCBI-BLASTx tool; polymorphism was analyzed by DnaSP version 6 software. MEGAx software and Translate tool – ExPASy was used for obtaining protein sequences and motif analysis, protein alignment was edited and represented using JALVIEW. Further, DNA Multiple sequence alignment of NBS type RGAs were constructed by using the T-COFFEE tool and represented by BOXSHADE program. The methodology followed in this study helped to isolate seven NBS encoding sequences from *D. wightii. L* and nucleotide sequence was submitted to GenBank for unique accession numbers. The multiple sequence analysis showed only a few point mutations contributed to diversity. In addition, results disclosed the presence of parsimony information site containing sequences. The comparative analysis revealed that last residues of Kinase-2 motif are conserved and consists amino acid Tryptophan (W) which confirmed that all *D.wightii. L* sequence isolated is Non-TIR-NBS-LRR Type (CNLs). The results of the present study revealed that universal degenerate primers designed for highly conserved regions of the NBS domain (LM638, LM637) are efficient to amplify the NBS encoding genes in the wild medicinal plant *Drimia wightii*. This study further confirms the absence of TNLs in monocots genomes which were confirmed previously in cereal crop plant genomes. The importance of parsimony information sites in diversity analysis and identifying common ancestors was highlighted.

Keywords: NBS-LRR Type R- gene, *Drimia wightii.* L, DNA Multiple Sequencing analysis, Protein Motif analysis. Degenerate PCR

Introduction
Molecular characterization is an important concept which discloses the genetic diversity and functionality of living organisms [1]. In plants, the molecular breeding programs and plant genetic resource conservation strategies require a better DNA Marker system to disclose the genetic diversity [2]. Recently, the functional markers gain importance in plant diversity studies to disclose the diversity of functionally characterized genes in plants [3]. The Nucleotide Binding Site- Leucine Rich Repeats (NBS-LRR) encoding sequences are functionally linked to disease-resistant mechanisms, and the isolation and characterization of NBS encoding sequence studies assisted in the screening of resistant and susceptible genotypes in crop plants [4]. Consequently, in many economically important crop plants, NBS-LRR Type R-genes are isolated and characterized using degenerate PCR and TA cloning methods [5]. But this methodology is limited to only economically important crop plants and as per the review and literature, there are no documents of NBS-LRR type Disease resistance genes and protein structures studied in non-domesticated wild medicinal plants like *Drimia* species. *Drimia* plant species also called “wild onions or sea onions” have a considerable role in curing several human diseases [6]. Several studies were conducted to explore its phytochemical constituents and their bioactive characteristics [6, 7, 8]. As a special usage, the extracts of *Drimia* plant species are also used as a Bio-pesticide against several plant pathogens [9]. The leaf and bulb extracts of *Drimia* species have been confirmed for their antimicrobial activity against both plant and human pathogens [10]. The insight of a broad range of antimicrobial activity in *Drimia* plant species shows that this plant may have a well-established defence mechanism that may differ from other crop plants. In addition, *Drimia* genetic and chromosomal variations were also still being researched to distinguish the different populations systematically [8, 11, 12].

Consequently, genetic variability and genome studies are interesting topics in this plant to resolve some basic questions in the classification system. Hence the present study aims to explore genetic resources linked to disease-resistant genes in *Drimia* species. The major objective is to study the diversity of disease-resistant linked genes and proteins in *Drimia wightii* and Comparative analysis of disease-resistant proteins between characterized R-genes and *Drimia wightii*. This is the first study interested in exploring the evolution of NBS-LRR Type R gene sequences in wild non-domesticated medicinal plant *Drimia wightii*.L.

Materials and methods

Drimia plant sample collection

Field trips were undertaken to collect the plant specimens of *Drimia wightii* and its bulbs from the dry land forest of Kolar district Karnataka, India. The plant samples were collected and authenticated by late Dr.Shiva Kameshwari associate professor Department of Botany JB campus, Bangalore University.

Genomic DNA isolation and quantification

Leaf samples of *Drimia wightii* collected and stored immediately at -80^{0}C until the sample was used for extraction. The genomic deoxyribonucleic acid (DNA) from leaf samples was extracted by using column-based thermo scientific GeneJET Plant Genomic DNA purification mini kit. The extracted DNA was quantified using nanodrop spectrophotometer and integrity was checked by resolving 7µl eluted DNA sample in 1% agarose gel and electrophoresis was carried out at 100V for 30 minutes.

Degenerate Oligonucleotide Primed Polymerase Chain Reaction (DOP-PCR)

The universal degenerate primer set was designed for conserved P-loop sequence and GLPL motif within the NBS domain (LM638 5′- GGIGGIGTIGGIAAIACIAC-3′ and LM-637 5′-ARIGCTARIGGIARICC-3′) [13, 14]. PCR was carried out using genei laboratories taq DNA polymerase kit (catalog No: 0601600051730). 2µl of 10x Taq buffer containing 10mM Tris-HCl (Ph 9.0) 1.5mM Mgcl2, 50mM KCl and 0.01% gelatin and 3µl of 2.5mm dNTP mix, 1.5µl of 10 picomol forward primer and 1.5µl of 10 picomol reverse primer, 0.4µl of Taq polymerase, 2µl of (50-100ng) of genomic DNA and 9.6µl of PCR grade Nuclease free Water added to make up final reaction volume 20µl. The polymerase chain reaction program at 95^{0}C for 10 minutes initial denaturation, 95^{0}C for 1 minute secondary denaturation, 48^{0}C for 1-minute annealing, and 72^{0}C for 1.5-minute extension with 35 cycles and 72^{0}C for 15 minutes final extension.

Cloning and sequencing of PCR products

Each PCR fragment was cloned into the plasmid using the pGEM-T Easy vector system (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 gene jet plasmid isolation kit (catalog No: K0502) and 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.

Bioinformatics tools used for sequence analysis

The bioinformatics tools were used as per the methods of *Lei et.al., 2014* [15] and *Palomino et.al., 2006* [16] with slight modification. Initially, the DNA sequence quality was analyzed using chromos software and the FASTA format of sequences was checked for vector contamination using NCBI VecScreen Tool. The specific sequences searched using NCBI-BLASTx tool for sequence homology search against the database. Sequence polymorphism and a number of nonsynonymous substitutions and synonymous sites were calculated using DnaSP version 6 software. Sequences are translated into amino acid sequences using the MEGAx software and Translate tool – ExPASy. The deduced amino acid sequences were then subjected to motif analysis, using the online version CLUSTALW multiple sequence alignment Programme of the European bioinformatics institute and edited with the JALVIEW program. DNA Multiple sequence alignment of NBS type RGAs were constructed by using T-COFFEE and box shade [17, 18] online bioinformatics program.

ORF analysis (Open Reading Frame analysis)

The obtained sequences in this study were analyzed for Open Reading Frames using NCBI Open Reading Frame Finder (https://www.ncbi.nlm.nih.gov/orffinder/) and Translate tool – ExPASy.

Results

Isolation of Drimia wightii NBS Type R- gene analogues

The resistance gene analogs of *Drimia wightii.L* (*DW*) was isolated using universal degenerate primers set LM638 and LM-637 is designed for conserved P-loop sequence and GLPL motif within the NBS domain. A total of 3 amplified multiple bands are obtained at 500 bp, 600bp, and 1000bp from genomic DNA (Figure 1). Based on the literature, this primer set is designed to amplify the 500-540 bp of DNA fragment from the

genomic DNA of crop plants. PCR products of both 500 and 600 were extracted and purified from agarose gel for TA cloning. The purified PCR fragments were ligated to pGEM-T Easy Vector by following kit protocol. Based on the literature the 1000bp PCR product is considered Nonspecific and not taken for TA cloning. In total, 26 positive clones were obtained after the blue-white screening with antibiotic selection for both 500 and 600bp. These 26 positive clones were further taken for colony PCR confirmation and segregated as eleven 500bp clones (viz., *DW1*, *DW2*, *DW5*, *DW6*, *DW7*, *DW9*, *DW10*, *DW11*, *DW12*, *DW15*, and *DW16*), a six 600bp clones (*DW3*, *DW4*, *DW8*, *DW13*, *DW14*, and *DW17*) after removing false-positive colonies form the group. From the remaining, non-specific partial sequences were also removed and the identical sequences were merged. Finally, a total of 7 sequences such as *DW2*, *DW5*, *DW6*, *DW7*, *DW10*, *DW11*, and *DW12* were obtained for 500bp, and only three sequences namely *DW13*, *DW14*, and *DW17* for 600bp. The sequencing results showed that the sequence lengths of 500bp group sequences are 480 to 500bp and sequence lengths of 600bp were obtained between 599 bp and 613 bp.

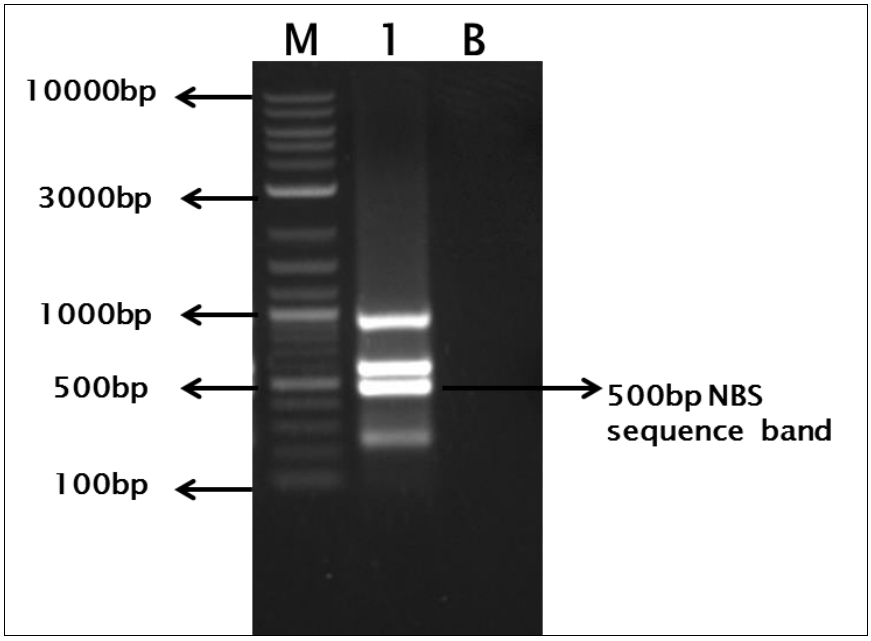

Fig 1: PCR amplification of NBS RGAs in *Drimia wightii* plant, M- 10kb ladder, Lane 1- *Drimia wightii* (Kolar) amplified NBS RGAs, Lane 2- B-Blank No template Control.

Homology analysis of Drimia wightii NBS sequences

The obtained 10 RGA clones were selected for homology study. The homology searches against Gene bank database using BLASTX algorithm revealed the 600bp sequences are similar to some uncharacterized protein of several plant species such as uncharacterized protein *Asparagus officinalis*-XP_020262776.1, *Musa acuminata subsp.malaccensis*-XP_017702161.1 (wild *banana* species), *Phoenix dactylifera*-XP_017702161.1 (date palm), *Quercus lobata*-XP_030923529 (*valley oak or roble*). But none of these sequences showed similarity to cloned NBS-LRR Type R genes in the database. In addition, the translated DNA sequence confirms that presence of premature and irregular stop codons in 600bp sequences. Hence these clones are removed from the group for analysis. Further, the BLASTX analysis of all 500bp sequences shown significant similarities with disease resistant proteins of several plant species such as *SUMM2*-like disease resistance proteins of *Elaeisguineensis*-XP_029122853.1 (*oil palm tree*), *Vitis riparia-XP_034679810.1* (*Wild Grape Vine*), putative disease resistance proteins of *Asparagus officinalis*-XP_020270986.1, NBS-LRR-like resistance protein of *Oryza sativa*-AL070094.1, Disease resistance protein *RPS5 Vitis vinifera*-RVW24676.1 (*common grape vine*) and disease resistance protein *RPS2*-like *Oryza brachyantha*-006651225.1. The Blast results of *DW* 500 bp sequences showed 45 to 58 percent identity with other characterized R-genes.

Gene polymorphism and homologous analysis of NBS Type R-gene analogues in Drimia wightii

Multiple sequence alignment of 500bp NBS type RGAs of *Drimia wightii* revealed that there was not much difference between in the seven sequences (GenBank accession numbers: OK562412, OK562413, OK562414, OK562415, OK562416, OK562417 and OK562418). However, single nucleotide polymorphism was observed in the sequences. The polymorphic sites were analyzed using DnaSP version 6 software. In these seven sequences, from region 1 to 500 bp identified 487 monomorphic sites without gaps or missing data. Further in the sequence alignment 13 polymorphic sites or mutations sites were identified. In that, 12 singleton variable sites and 1 parsimony informative sites at position 209th in sequences *DW 5* and *DW7*. The DNA polymorphism analysis showed that all obtained sequences are Haplotypes with diversity (Hd) is 1 and Nucleotide diversity (Pi) is 0.00781. The conserved region analysis showed that sequence conservation (C) value is 0.974 and all the sequences were conserving at position starts from 343 to 474 (Figure 2).

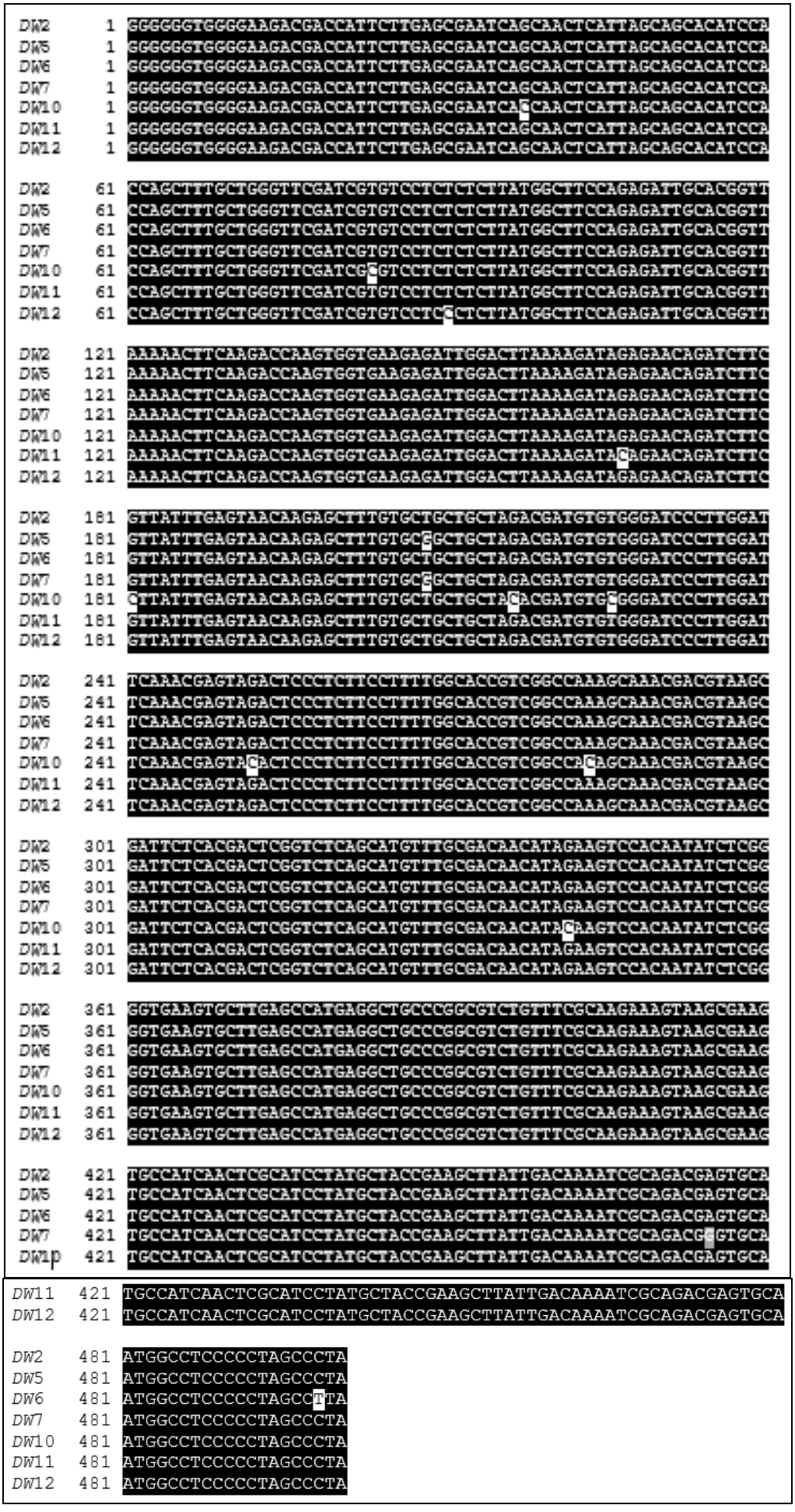

Fig 2: Alignment of 500bp NBS type –R genes in *Drimia wightii*.

The Clustlw online tool was used to measure the percent identity between 7 NBS sequences. The extent of homology between these sequences is from 98% to 99%. The evolutionary history of 7 sequences of *Drimia wightii* was inferred using the Neighbor-Joining method [19]. The bootstrap consensus tree inferred from 500 replicates [20] is taken to represent the evolutionary history of the taxa analyzed [20]. Branches corresponding to partitions reproduced less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) is shown next to the branches [20]. The evolutionary distances were computed using the Maximum Composite Likelihood method [21] and are in the units of the number of base substitutions per site. All ambiguous positions were removed for each sequence pair (pairwise deletion option). Evolutionary analyses were conducted in MEGA X [22]. The tree is divided into two groups in Group1 (*DW7*, *DW5*, *DW11*, *DW12*) and Group 2 (*DW6*, *DW10*, *DW2*) with a common ancestor (Figure 3).

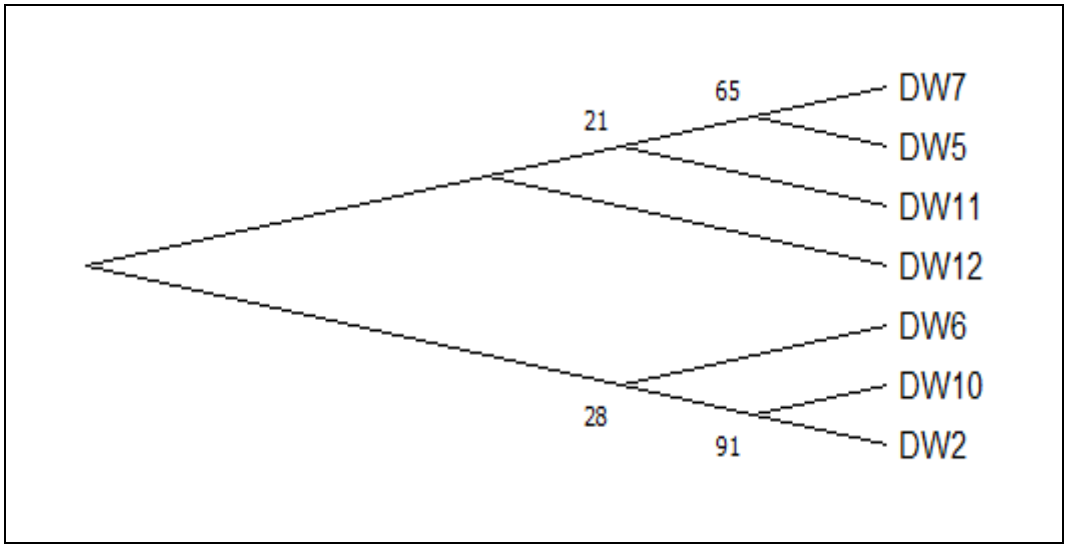

Fig 3: Phylogenetic tree of NBS resistance RGAs in *Drimia wightii*.

Open Reading Frame analysis

All obtained DW sequences were checked for Open Reading Frame using NCBI ORF finder tool. To obtain complete ORF information of the sequences the NCBI ORF finder parameters were set to detect minimum ORF length of 30 nucleotide. Initially results was collected by choosing "ATG as only starting codon" and later for the same sequence collected results by choosing "ATG and other alternate Codons" as starting codon. ORF results are given in table 1 and table 2 respectively. Expasy Translation tool was used to confirm the protein sequences. ORF numbers and positions are similar to all 500bp sequences due to fewer variations in sequences. The results also confirmed that presence of small nested ORFs in the main-ORFs. Compared to ORFS obtained with ATG as only starting codon; the ATG with other alternate codons option showed better results with two main-ORFs with 68 and 67 amino acid strand that covers most of the motifs in the NBS domain. The protein product of OFR containing Kinase-2 motif (ORF1-start from 97 to ends at 303) of NBS domain was selected for smart blast search against protein database. The results matched to CC-NBS-LRR class Disease resistance gene of *Arabidopsis thaliana*.

Table 1: Showing common ORF found in all sequences of 500bp sequences when selected only ATG as start codon, the important motifs are highlighted i.c Kinase-2 motif in ORF 1- Frame 1 of forward strand.

Orfs	Frame	Strand	Start	Stop	Protein Sequence
ORF 1	1	Forward strand	97	303	>lcl|ORF1 Masrdctveklqdqvvkrldlkdreqifsylsnksfvllddvwdpldlk Rvdslfllapsakandvs (68 amino acids)
ORF 2	1	Forward strand	379	414	>lcl|ORF2 Mrlpgvcfark (Nested -ORF-11 amino acids)
ORF3	1	Forward strand	421	495	>lcl|ORF3 Mpstrilcyrslltksqtsamasp(24 amino acids)
ORF4	2	Forward strand	221	253	>lcl|ORF4 Mcgipwisne (Nested -ORF-10 amino acids)
ORF5	3	Forward strand	441	>500	>lcl|ORF5 Mlpklidkiadecnglplal (20 amino acids)

Table 2: Showing common ORF found in all sequences of 500bp sequences when selected ATG and alternate codon as start codon, the important motifs are highlighted i.c Kinase-2 motif in ORF 1- Frame 1 of forward strand.

Orfs	Frame	Strand	Start	Stop	Protein SEQUENCE
ORF 1	1	Forward strand	97	303	>lcl|ORF1 Masrdctveklqdqvvkrldlkdreqifsylsnksfvllddvwdpldlk Rvdslfllapsakandvs (Main-ORF-68 amino acids)

ORF 2	1	Forward strand	379	414	>lcl|ORF2 MRLPGVCFARK (Nested ORF -11 amino acids)
ORF3	1	Forward strand	421	495	>lcl|ORF3 Mpstrilcyrslltksqtsamasp (Nested ORF-24amino acids)
ORF4	2	Forward strand	68	145	>lcl|ORF4 Mlgsivssllwlpeiarlknfktkw (25 amino acids)
ORF5	2	Forward strand	221	253	>lcl|ORF5 Mcgipwisne (Nested ORF-10 amino acids)
ORF6	2	Forward strand	329	382	>lcl|ORF6 Mrqhrspqylggevlep(Nested ORF -17 amino acids)
ORF7	3	Forward strand	300	>500	>lcl|ORF7 Milttrsqhvcdnievhnisevkclsheaarrlfrkkvsedainshpmlp KLIDKIADECNGLPLAL (Main-ORF-67 amino acids)

Conserved motifs analysis in sequences

The amino acid sequences of the 7 NBS-type –RGAs sequence were deduced and the length of the protein sequence consists of 160 to 166 amino acids. The BLAST analysis detected the sequence homology with typical NBS domain internal conserved motif Kinase-2 (VLDDVW). The presence of an internal kinase-2 motif region [LDDVW] in all the *DW* 500bp confirms that these clones resemble the NBS type of R genes. However, in *DW* RGAs the variants in the kinase-2 motif were observed. Initially the kinase motif in all DW sequences mutated with amino acid L instead of V in the motif [LLDDVW]. In the DW10 clone, the amino acid mutated to H instead of D, and R instead of W was also observed [LLHDVR]. The P-loop motif [GGVGKTT] is conserved in most of the NBS-LRR type R-gene sequences in many species [16]. This P-loop Motif was well conserved in all isolated DW RGAs clones and no other variants in the motif were observed. Multiple sequence alignments of deduced amino-acid sequences of the 7 DW RGA clones with known qualified NBS-LRR Type R-genes such as RPS2, XA1, N, M, and RPP4 were carried out using CLUSTALW and edited with Jalveiw program (Figure 4). The phylogenetic analysis of DW RGAs amino acid sequence with an already characterized R gene was inferred using the Neighbor-Joining method (Figure 5) [19]. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) is shown next to the branches [20]. The evolutionary distances were computed using the Poisson correction method [21] and are in the units of the number of amino acid substitutions per site. This analysis involved 13 amino acid sequences. Evolutionary analyses were conducted in MEGA X [22].

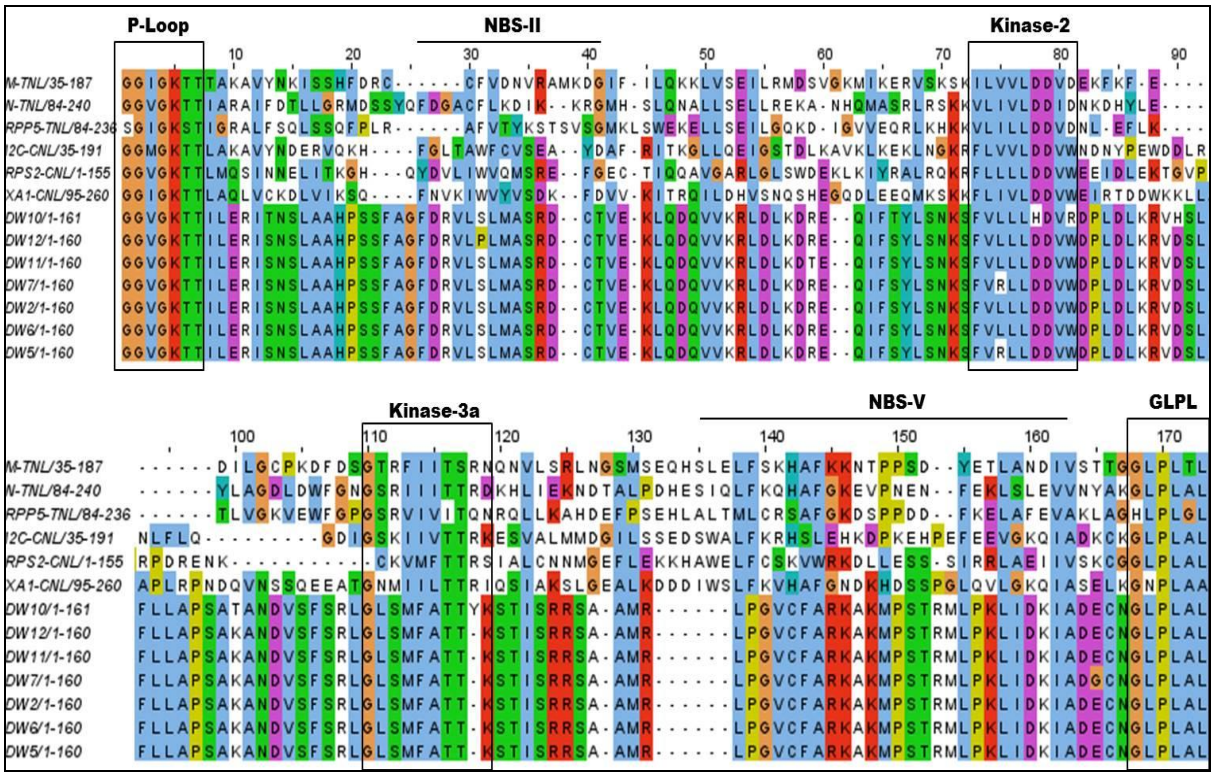

Fig 4: Alignment of deduced amino acid sequences of 7 RGAs of *Drimia wightii* and the nucleotide Binding site domain (NBS-Domain) of 6 known R-genes of *Tobacco* N, *Tomato* I2 and M, *Arabidopsis* RPP5, *Rice* Xa1 and RPS2, RPS5 and RPP5 of *Arabidopsis*, Xa1 of *rice*. Identical amino acids and CLUSTAL2 and JALVIEW programs used for multiple sequence alignment, conserved sequences are blocked with the same color. The conserved motifs are shaded and labeled as described in the Pan Q *et.al, 2000* [23] and Palomino C *et.al. 2006* [16].

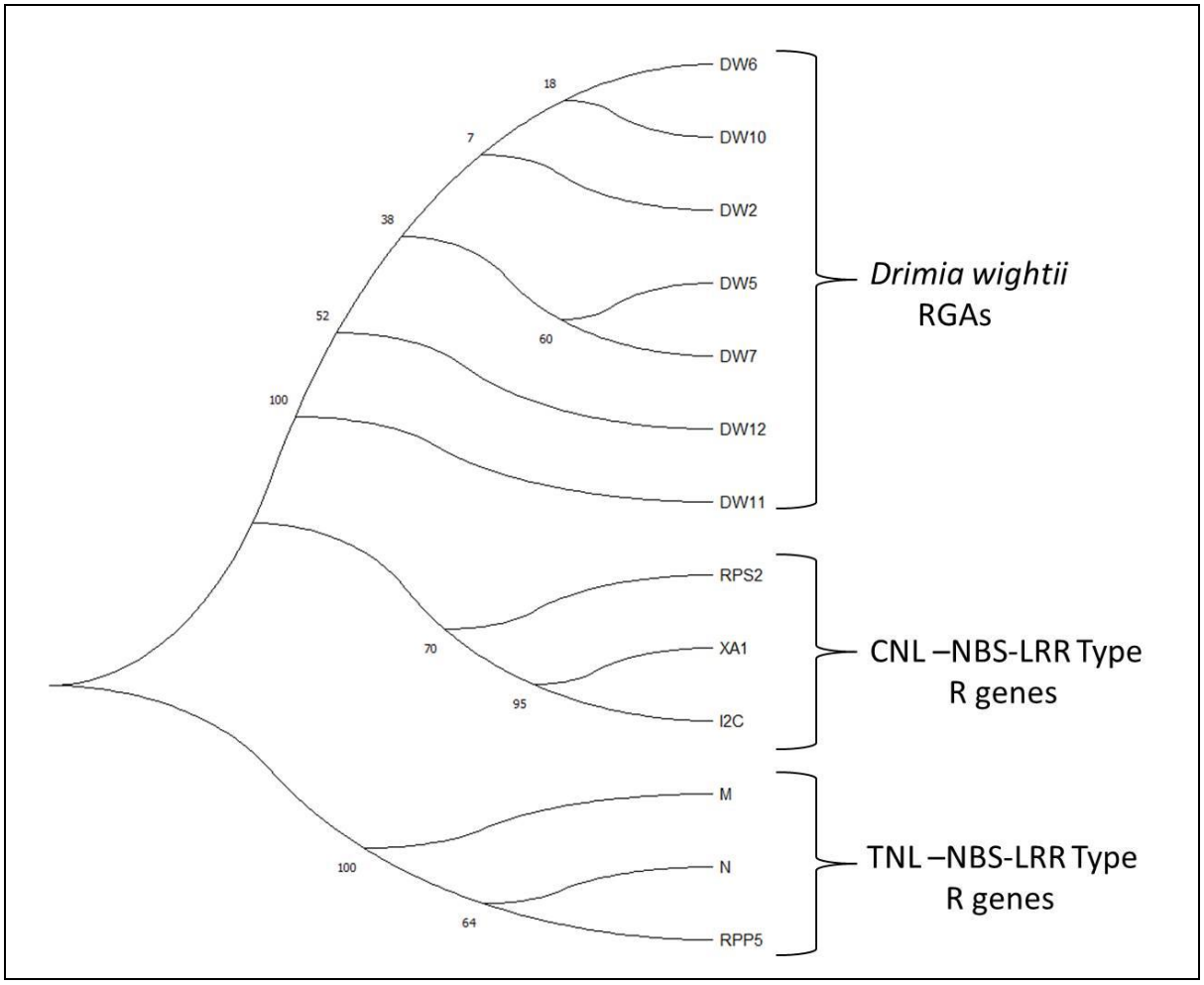

Fig 5: Phylogenetic relationship of amino acid sequences of 7 RGAs isolated form *Drimia wightii* and six characterized NBS-LRR Type R-genes from *Arabidopsis*(RPS2, RPS5), *Rice*(Xa1), *tomato*(I2C, M) and *tobacco*(N) plants.

Discussion

The defence mechanism in plants is a complex process, which involves different signaling pathways and mechanisms [24]. Generally, plants' defence mechanisms are classified into active and passive. The passive defence is the first line of defence, which is involved in physical barriers like cuticles and thorns. Active defence is classified into two types, basal plant defence- which restricts the invasion of a virulent pathogen or insect. The other one is involved in specific recognition of the invading pest or pathogen by plant resistance (R) genes [25]. In addition, based on the genetics of R gene-mediated defence mechanisms, crop Plants are grouped into resistance and susceptible genotypes. During the present investigation, the disease resistance (R) genes in Drimia species were studied to see the diversity of R genes in different plants.

The NBS-LRR nucleotide-binding sites leucine-rich repeats are major family genes that are involved in disease resistance mechanisms against diverse pathogens in many crop plants [26]. But these R-genes sequences (NBS-LRR) are not explored in medicinally important plants especially in *Drimia* species. Therefore, necessary to explore the disease-resistance mechanism in different plants to understand its role and find out novel genes for disease resistance mechanisms. Disease resistance mechanism and organization of the NBS-LRR sequences are fundamental resources for molecular characterization of R- genes. Structurally NBS-LRR proteins have 4 different domains' i.e. Variable amino-terminal domain, the NBS domain, the LRR region, and Variable Carboxy-Terminal Domain [26-27]. Specifically, the NBS domain contains highly conserved subdomains such as the P loop and hydrophobic putative membrane-spanning GLPL domain. By using PCR with degenerate primers specific to these conserved regions (P loop and GLPL domains) of the NBS domain many NBS-LRR type R genes are isolated and cloned to characterize their structure and function in many different crop plants [27].

In addition, commercially important medicine valued plants like *turmeric* NBS-LRR type R genes are also characterized by using PCR-based methods [28]. Similarly, in the present study, PCR techniques with degenerate primers were used to identify putative disease resistance genes in *Drimia wightii*. The degenerate Primers used in this study have already been used in different crop plants for the identification of resistant gene analogs [29, 13]. The primers LM-638 and LM-637 primers also known as universal primers set to isolate NBS-LRR Type RGAs (Resistance Gene Analogues) designed to the highly conserved motifs of the NBS domain [29]. LM638 is a forwarded primer designed to GGV/IGKTT (P-loop) and primer LM-637 is a reverse primer designed to

GLPLAL (hydrophobic domain) of characterized NBS-LRR Type R-genes from N, L6, and RPS2 genes [16]. The homology search against the database using the BLASTx tool confirmed that 500 bp sequences matched significantly with putative disease resistance proteins of *Phoenix dactylifera* (Date Palm), disease-resistant protein -SUMM2 LIKE of *Elaeis guineensis* (oil *palm* or *African oil palm*), and other monocot species. Similarly, 600bp sequences also matched with some uncharacterized proteins of *Musa acuminata subsp. malaccensis* (wild banana species), *Phoenix dactylifera* (date palm), *Quercus lobata* (valley oak or roble) but not to any characterized Disease resistance genes. Remarkably, the 500bp of *Drimia wightii* protein motif sequences highly matched with characterized NBS-LRR Type R-genes of *Rice*, *Tomato*, and *Wheat* and *Arabidopsis* plants.
Drimia plant species is known for phenotypic and genotypic complexities, and variations in accessions within species lead to the appearance of a wide range of intermediate forms, which led to confusion taxonomically [12]. In the classification system, *Urginea* is separated from *Urginea* and included in *Drimia* [31, 32]. Further, the Plastids DNA Molecular studies by P*fosser and Speta in 1999* placed *Drimia* in the Tribe *Urginea*, Order *Scillodeae*, and Family *Asparagaceae* [11]. With reference to this, an important observation was recorded in the homology search, the sequences containing parsimony informative sites at 209th position (*DW5* and *DW7*) sequences have significant similarity with 54% identity to putative disease resistance protein At1g63350 - *Asparagus Officinalis-XP_020270986*; the sequence DW10 also shown match to uncharacterized protein A4U43_C08F29190 of *Asparagus Officinalis*-ONK61370.1. Currently, in the classification system, the *Drimia* species come under clade: *monocots*, order: *Asparagales*, and family: *Asparagaceae*. This confirms that the sequence *DW5* and *DW7* shares common ancestry history with family *Asparagaceae* plants. In the case of remaining sequences, they have identified with other monocot plants like *Phoenix dactylifera* (Date Palm) and *Musa acuminata subsp.malaccensis* (wild banana species) however all the sequences showed 38-45% identity with *Asparagus Officinalis* protein sequences.
Gene polymorphism and homologous analysis of NBS-LRR Type R-gene analogs of *Drimia wightii* showed good homology. Overall, the diversity between the sequences is only due to the presence of point mutations. This further infers that the NBS-LRR Type R genes in the Drimia wightii are evolving slowly with point mutations and not randomly evolving with recombination events. Thus the translated DNA sequence of 600 bp showed the presence of irregular Stop codons and premature Stop codons in the protein sequence. This confirms that irregular recombination events in the genome created possible pseudogene families in the *Drimia wightii* genome.
The conserved motifs analysis of *Drimia wightii* translated DNA sequence results show that all *Drimia wightii* sequences show good homology containing P-LOOP and GLPLAL motifs. The comparative analysis of *Drimia wightii* RGAs with characterized NBS-LRR Type-R genes shows that the motif Kinase-2 (VLDDVW) is well conserved. However, in the *DW* sequence Kinase-2 motif is slightly changed compared to the typical 'VLDDVW/D' pattern to 'VLLLDDVW' due to the insertion of two extra Leucine amino acids. And in the *DW10* amino acid sequence, the replacement of D (aspartic acid) with H (Histidine) and replacement of W (Tryptophan) with R (Arginine) was also observed (VLLLHDVR). In addition, the *DW5* and *DW7* sequences which contained parsimony informative sites in DNA showed the replacement of L (Leucine) with R (Arginine) amino acid however the variation in this position with characterized amino acid motifs are common. The synonymous and non-synonymous mutation in the DNA sequences was analyzed between every 2 sequences using DnaSP version 6 software. The comparison between parsimony information sequences (*DW*7 and *DW*5) with other sequences revealed that mutation G/T at 209th position is nonsynonymous which replaces the amino acid L to R in the Kinase-2 motif. In *DW*7 one more mutation at 475th position leads amino acid replacement E with G. Interestingly, the sequence DW10 has a slightly different Kinase-2 motif with replacing typical LDDVW/D amino acid combination to LHDVR combination this indicates some non-synonymous mutation in the sequences changes in Kinase motif.
The last residue of Kinase-2 motif (VLDDVW/D) of NBS domain has been used to classify RGAs into the TIR-NBS-LRR family or Non-TIR-NBS-LRR family Resistant gene analogs (RGAs) in plants. The conservation of tryptophan (W) and aspartic acid (D) are distinct characteristics of Non-TIR-NBS-LRR and TIR-NBS-LRR proteins. For example, the Kinase-2 motif with ending residue D is classified as TIR-NBS-LRR proteins (VLDDVD) and the Kinase-2 motif with ending residue W are classified as Non-TIR-NBS-LRR proteins (VLDDVW). The Alignment of deduced amino acid sequences of 7 RGAs of *Drimia wightii* with known NBS-LRR Type R-genes shows this characteristic conservation of amino acids which can be used for classifying the *Drimia wightii* RGAs. In this study, the six Known NBS-LRR-type R-genes were used. Those are N genes from *tobacco*, M gene from *Tomato* and RPP5 from *Arabidopsis* are known as TIR-NBS-LRR Type R-genes (TNLs); and I2C gene from *Tomato*, Xa1 gene from *Rice* and RPS2 of *Arabidopsis* are known non-TIR-NBS-LRR (also called as CC-NBS-LRR). The protein sequence alignment perfectly shows that the last residue of the Kinas-2 motif in the NBS domain shows the presence of amino acid 'D' for all TNLs and 'W' for all CNLs type RGAs (figure 4). Many studies have confirmed that cereals and grass species lost these TNLs and contain only Non-TIR-NBS-LRR (CNLs) [33] and many dicots have both types of NBS-LRR type R-genes in the genome. Interestingly, the *Drimia wightii* RGAs contains "W" as the last residue of Kinase-2 motif and confirmed as CNLs and the phylogenetic analysis of amino acids sequence with six characterized R-genes shows that all *DW* sequences grouped with a single common ancestor with known CNL-Type R genes such as I2C gene from Tomato, Xa1 gene from Rice and RPS (Figure 5). However, the DW sequence is slightly different in kinase motifs and NBS domain; Hence, *Drimia wightii* RGAs grouped distant from known CNLs of crop plants.

Further, The DW10 amino acid sequence with the kinase-2 motif 'VLLLHDVR' was searched in the Protein database using the "Protein Blast tool".

Search parameters were adjusted to search for a short input sequence. The results show that the 'VLLLHDVR' amino acid sequence matched 100% with the unnamed protein product of *Triticum turgidum subsp. Durum* (Accession Number: VAI91578.1) and hypothetical protein Tci_416672 *Tanacetum cinerariifolium* (Accession Number: GEY44698.1) but not to any functionally characterize protein. As well, with one amino acid difference (VLLLH/DDVR) this motif shows 88% identity with predicted disease resistance protein RPS2-like of *Oryza brachyantha* (Accession Number: XP_006652512.1) and disease resistance protein RPS2-like of *Phoenix dactylifera* plant (Accession Number: XP_006652512.1). The presence of 'R' (arginine) amino acid as the last residue in the Kinase-2 motif in the NBS domain was detected also in NBS-LRR Type R-genes of other monocot species. This further confirms that the variant (VLLLDDVW/R) in the kinase motif also exists in some predicted plant R-genes.

Overall, the study shows that there is less diversity in NBS-LRR encoding genes in selected *Drimia wightii* plant. However, The *Drimia wightii* NBS domain DNA sequences are not matched to any currently available NBS-LRR nucleotide databases. The only translated DNA sequence showed similarity with protein motif between available NBS-LRR disease resistance proteins. The most common reason homologs are missed in DNA sequences, rather than translated DNA sequences, is because Protein similarity searches are much more sensitive than DNA: DNA searches. DNA: DNA alignments rarely detect homology after more than 200–400 million years of divergence; but protein: protein alignments routinely detect homology in sequences that last shared a common ancestor more than 2.5 billion years ago [34]. Compared to characterized disease-resistant NBS-LRR Type R-gene sequences, the translated DNA sequences of *drimia wightii* are significantly diverse and grouped separately in the phylogenetic tree (Figure 5). It is documented that A small number of *R* genes can provide defence against diverse pathogens if a limited number of effector targets are present [35]. The function of NBS-LRR Type R-genes is completely not known, and Resistance genes in plant evolve by divergent selection process [35]. In addition to this, previously the smaller ORFs nested within main-ORFs generally are not interpreted features for coding sequences, but recent survey on alternative ORFs and small ORFs proving the examples for importance of nested small ORFs in regulation of main-ORF translation initiation in eukaryotes [36]. In this study the ORF analysis of all 500 base pairs of DW sequences displayed nested small ORF Overlapping the main-ORF (Table 1 and Table 2). Further, the protein product of these small nested ORFs was successfully matched to functional proteins motifs in the protein database (data not shown). This may give hints that these nested small ORFs might be involved in translating diverse proteins during different biotic stress condition. But, to confirm this further functional studies are required.

As limitations, *Drimia wightii* is a wild non-cultivated medicinal plant that was collected from natural habitat. Hence, specific phenotype and genotype study against different biotic stress were not conducted and documented. Furthermore, complete set of NBS-LRR genes are required to study the organizations of NBS-LRR genes in genome but this can be achieved only in whole genome sequenced plants. As a future perspective the obtained genomic information can be used to design the functional studies to understand the role of NBS- LRR type R-gene importance in different plants systems.

Conclusion

NBS-LRR Type R-genes and proteins were explored in a non-domesticated wild medicinal plant *Drimia wightii*. The diversity between predicted NBS-LRR types R-genes in the *Drimia wightii* suggests that R-genes in the genome are slowly evolving with point mutations but the complete protein sequence of *DW* group separately in the phylogeny. The protein-coding analysis also showed the presence of pseudogene in the genome due to random recombination. This study further supports that the primer combination used in the current investigations is efficient in amplifying NBS-LRR encoding genes in the wild medicinal plant. The loss of TIR-NBS-LRR Type encoding genes was first confirmed in cereal genomes based on the amino acid residues of the Kinase-2 motif of the NBS domain. This study further confirms the absence of TIR-NBS-LRR encoding genes in the wild monocot plant genome other than the cereal species.

Reference

1. Laurentin Hernán. "Data analysis for molecular characterization of plant genetic resources." Genetic Resources and Crop Evolution,2009:56(2):277-292.
2. Staub Jack E, Felix C Serquen, Manju Gupta. "Genetic markers, map construction, and their application in plant breeding." HortScience,1996:31(5):729-741.
3. Anderson, Dane. "Functional Markers in Zea mays L.", 2021.
4. Panwar Preety, Anand Kumar Jha, Pandey PK, Arun K. Gupta *et al*. "Functional markers based molecular characterization and cloning of resistance gene analogs encoding NBS-LRR disease resistance proteins in finger millet (Eleusine coracana)." Molecular biology reports,2011:38(5):3427-3436.
5. Kanazin Vladimir, Laura Fredrick Marek, Randy C. Shoemaker. "Resistance gene analogs are conserved and clustered in soybean." Proceedings of the National Academy of Sciences,1996:93(21):11746-11750.
6. Pandey Dhananjay, Gupta AK. "Bioactive Compound in Urginea indica (Kunth.) from Bastar and its Spectral Analysis by HPLC, UV-Vis, FT-IR, NMR, and ESI-MS." SN Comprehensive Clinical Medicine,2019:1(4):241-254.

7. Mahato Dipa, Abhishek Prasad Sahu, Sharma HP. "Phytochemical and antioxidant evaluation of Urginea indica Kunth.", 2018.
8. Lakshman Bijul A, Paramasivam G. "Biosystematics studies on medicinal plant Urginea indica Kunth. liliaceae-A review." International Journal of Pharmacy & Life Sciences,2012:3:1.
9. Metin Mert, Betül Bürün. "Effects of the high doses of Urginea maritima (L.) baker extract on chromosomes." Caryologia,2010:63(4):367-375.
10. Pandey Dhananjay, Ashwini Kumar Gupta. "Antimicrobial activity and phytochemical analysis of Urginea indica from Bastar district of Chhattisgarh." Int J Pharm Sci Rev Res,2014:26:273-81.
11. Pfosser Martin, Franz Speta. "Phylogenetics of Hyacinthaceae based on plastid DNA sequences." Annals of the Missouri Botanical Garden, 1999, 852-875.
12. Geetha HL, Nijagunaiah R. "Karyotypic variations in different accessions of Urginea indica (Roxb) Kunth. and Urginea wightii Lakshmin. Hyacinthaceae." Cytologia,2016:81(3):329-334.
13. Kanazin Vladimir, Laura Fredrick Marek, Randy C. Shoemaker. "Resistance gene analogs are conserved and clustered in soybean." Proceedings of the National Academy of Sciences,1996:93(21):11746-11750.
14. Miller Robert NG, David J Bertioli, Franc C Baurens, Candice MR Santos, Paulo C Alves *et al*. "Analysis of non-TIR NBS-LRR resistance gene analogs in Musa acuminata Colla: Isolation, RFLP marker development, and physical mapping." BMC Plant Biology,2008:8(1):1-15.
15. Lei Xintao, Quansheng Yao, Xuerong Xu, Yang Liu. "Isolation and characterization of NBS–LRR resistance gene analogues from mango." Biotechnology & Biotechnological Equipment,2014:28(3):417-424.
16. Palomino C, Satovic Z, Cubero JI, Torres AM. "Identification and characterization of NBS–LRR class resistance gene analogs in faba bean (Vicia faba L.) and chickpea (Cicer arietinum L.)." Genome,2006:49(10):1227-1237.
17. Notredame Cédric, Desmond G, Higgins, Jaap Heringa. "T-Coffee: A novel method for fast and accurate multiple sequence alignment." Journal of molecular biology,2000:302(1):205-217.
18. Hofmann K, Baron MD. Boxshade server, 2000.
19. Saitou Naruya, Masatoshi Nei. "The neighbor-joining method: a new method for reconstructing phylogenetic trees." Molecular biology and evolution,1987:4(4):406-425.
20. Felsenstein Joseph. "Phylogenies from restriction sites: a maximum-likelihood approach." Evolution,1992:46(1):159-173.
21. Tamura Koichiro, Masatoshi Nei, Sudhir Kumar. "Prospects for inferring very large phylogenies by using the neighbor-joining method." Proceedings of the National Academy of Sciences 101,2004:30:11030-11035.
22. Kumar Sudhir, Glen Stecher, Michael Li, Christina Knyaz, Koichiro Tamura. "MEGA X: molecular evolutionary genetics analysis across computing platforms." Molecular biology and evolution,2018:35(6):1547.
23. Pan Qilin, Jonathan Wendel, Robert Fluhr. "Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes." Journal of molecular evolution,2000:50(3):203-213.
24. Brown John Frederick, Ogle HJ. Plant pathogens and plant diseases. Published by Rockvale Publications for the Division of Botany, University of New England, 1997.
25. Kaloshian Isgouhi. "Gene-for-gene disease resistance: bridging insect pest and pathogen defense." Journal of chemical ecology,2004:30(12):2419-2438.
26. McHale Leah, Xiaoping Tan, Patrice Koehl, Richard W. Michelmore. "Plant NBS-LRR proteins: adaptable guards." Genome biology,2006:7(4):1-11.
27. Panwar Preety, Anand Kumar Jha, Pandey PK, Arun K Gupta, Anil Kumar. "Functional markers based molecular characterization and cloning of resistance gene analogs encoding NBS-LRR disease resistance proteins in finger millet (Eleusine coracana)." Molecular biology reports,2011:38(5):3427-3436.
28. Joshi RK, Mohanty S, Subudhi E, Nayak S. "Isolation and characterization of NBS-LRR- resistance gene candidates in turmeric (Curcuma lonqa cv. surama)." Genetics and Molecular Research,2010:9(3):1796-1806.
29. Yeo Brandon Pei Hui, Lian Chee Foong, Sheh May Tam, Vivian Lee, Siaw San Hwang. "Gene isolation using degenerate primers targeting protein motif: a laboratory exercise." Biochemistry and Molecular Biology Education,2018:46(1):47-53.
30. Reddy IN, Bheema Lingeswara, Srinivas Reddy D, Lakshmi Narasu M, Sivaramakrishnan S. "Characterization of disease resistance gene homologues isolated from finger millet (Eleusine coracana L. Gaertn)." Molecular Breeding,2011:27(3):315-328.
31. Jessop JP "studies in the bulbous liliaceae in South africa: Vii. The taxonomy of drimia and certain allied genera.", 1977.
32. Stedje Brita. "A revision of the genus Drimia (Hyacinthaceae) in East Africa." Nordic Journal of Botany,1987:7(6):655-666.
33. Nepal Madhav P, Ethan J Andersen, Surendra Neupane, Benjamin V. Benson. "Comparative genomics of non-TNL disease resistance genes from six plant species." Genes,2017:8(10):249.
34. Pearson William R. "An introduction to sequence similarity ("homology") searching." Current protocols in bioinformatics,2013:42(1):3-1.

35. Meyers Blake C, Alexander Kozik, Alyssa Griego, Hanhui Kuang, Richard W. Michelmore. "Genome-wide analysis of NBS-LRR–encoding genes in Arabidopsis." The Plant Cell,2003:15(4):809-834.
36. Orr Mona Wu, Yuanhui Mao, Gisela Storz, Shu-Bing Qian. "Alternative ORFs and small ORFs: shedding light on the dark proteome." Nucleic Acids Research,2020:48(3):1029-1042.