RGAs Approach in Identification of Disease Resistance Genes and their Deployment in Crops Improvement

Pavan Kumar¹, Sameer Chandra², V.S. Chauhan¹ and Sangeeta Srivastava³ *

¹Department of Biotechnology, Bundelkhand University, Jhansi 284128, India.
²Department of Plant Science (Faculty of Applied Sciences), MJP Rohilkhand University, Bareilly-243006, India.
³Division of Crop Improvement, ICAR- Indian Institute of Sugarcane Research, Raebareli Road, Lucknow 226002, India.
Corresponding author email: sangeeta_iisr@yahoo.co.in

Abstract

Plant pathogens are chief consequent to effect crop productivity. Globally up to 40% crop lost per annum due to plant disease. More amount of food will be required in coming time as increasing the world's population. Disease resistance genes contribute resistance to plants against pathogens and may be economic and eco-friendly biological approach to improve crop productivity by developing disease resistant crop in compare to other methods. Several research studies underway for attaining durable and broad spectrum resistance using deployment of resistance genes include R genes and QDR genes. Generally, resistance gene analogue approach with bioinformatics tool is used to identify disease resistance gene and thousands of candidate disease resistance genes or resistance gene analogues (RGAs) identified in different plants.

Keywords: Plant pathogen, R-genes, Durable and Broad spectrum resistance, RGAs, QDR genes

INTRODUCTION

Phytopathogens are a major constraint for plants because they affect crop productivity and quality. About 20-40% of all crops are lost globally due to pests and diseases every year [1]. The world population is expected to rise from 7.4 billion people to 8.4 billion in 2030 and 9.5 billion in 2050 [2]. The greatest challenge in the 21st century
would be to increase the crop productivity and quality, and to reduce the loss of crops. A big question arises that how much we are ready and conscious to resolve the upcoming problems. The conventional practice used to control crop diseases commonly fall into following categories: chemical, cultural practices and biological control methods. The chemicals being used for this purpose are not having sustainable approach since they adversely affect the human and whole environment by bioaccumulation and magnification process. Practices such as mix cropping, crop rotation, manipulation of planting time, plant density and spacing etc. play a major role in prevention of crop diseases [3]. However, in many cases impractical, inadequate or financially nonviable and sometimes the use of cultural practices may affect the wild life. Biological control methods include biological control agent (e.g. bacteria, fungi etc.) and host resistance commonly used to control crop diseases. However, biological control is failure in management of destructive diseases. The host resistance provided by resistance (R) genes is an economic and eco-friendly biotechnological approach to control crop diseases and improve crop productivity by developing disease resistant crop. The development of disease resistance through conventional plant breeding has been a cost effective, eco-friendly and successful method to develop of disease resistant varieties [4]. On the other hand breeding process take more than ten years to release a commercial variety. Resistance gene analogues (RGAs) approach has proved to be a land mark step in identification of disease resistance gene in different crops. A large number of resistance genes have been identified to date and have been successfully cloned in a few crops. RGAs have not only been used as genetic markers for genetic mapping of resistance genes but also opens new avenues in studying of genetic organization and evolutionary features of various classes of resistance genes among different plant species [5]. Identification of resistance genes still remains a difficult assignment due to the lack of information about genomic structure of several valuable crops.

(i) Phytoresistance genes: a step towards future crop improvement

Plants disease resistance can be divided into two categories: qualitative disease resistance and quantitative disease resistance (QDR). Commonly two groups of genes take part in conferring resistance against disease, include receptor genes and second are attributed by act in response to a pathogen attack via altering levels of expression or posttranslational modifying their encoding proteins. The receptor genes comprise R genes, host pathogen recognition receptor (HPRR) genes and defense responsive genes. HPRR genes and defense responsive genes actively provides durable and broad spectrum resistance against race nonspecific pathogen and qualitative resistance mediated by R genes confers complete resistance to race specific pathogen [6], [7], [8]. The gene mechanisms underlying QDR mostly unknown and differ from classical gene for gene concept [8]). Multiple genes that can interact with the environment as well as with each other [9] governs QDR, widely used in breeding programme because of stable and durable resistance due to racial unspecificity [10]. Plants have a complex immune system that bears testimony to co-evolution of plant and pathogen. The initial perception between host and pathogen carried out by two receptors: pattern recognition receptors (PRRs) and R-gene receptors. This eventually leads to induction
RGAs Approach in Identification of Disease Resistance Genes and their...

of signal transduction pathway and defense responses against pathogen. PRRs are generally receptor like kinases (RLKs) [11], attached to cell membrane and recognize pathogen or microbe associated molecular pattern (PAMPs or MAMPs) such as flagellin or chitin, lipopolysaccharides, murein etc. The resistance proposed by PRRs against pathogen or microbe is referred as PAMP triggered immunity (PTI or MTI) or basal resistance. Another perception of pathogen virulence molecule (effector) is recognized by an intracellular receptor called R protein and triggers defense against pathogen. This kind of resistance is commonly known as effectors triggered immunity (ETI) or R-gene mediated resistance. Several research findings suggest that the concept of plant immunity including QDR needs to be more explored in order to reveal the mystery of resistance. The R-genes have three key characteristics: pathogen detection, induction of signaling pathway and adaptation for quick evolution of pathogenic specificity [12]. However, race specific resistance provided by R genes; that operate in gene for gene fashion. The plant disease resistance (R) gene encodes large number of pathogen receptors which have the capability to recognize a pathogen attack and assist a counter attack across the pathogen [13].

(ii) Categorization of phytoresistance gene

There are eight classes of resistance genes (Table 1) which have been classified on the basis of their conserved protein structure [14].

(I). Resistance genes characterized by cytoplasmic coiled coil (CC) domain and NBS (nucleotide binding site) domain fused with LRR domain such as the RPM1 and RPS2 gene of Arabidopsis gene and I2 resistance gene of tomato.

(II). Genes is represented by mammalian toll interleukin-1 receptor (TIR) domain and NBS domain fused to LRR domain. The tobacco N gene and flax L6 gene and RPP5 gene belong to this class [15]

(III). Genes contains an extracellular leucine rich repeat (eLRR) domain and a trans membrane (TM) domain for example the gene Cf-2, Cf-4, Cf-9 of tomato. The eLRR indirectly plays a major role in activation of defense protein and pathogen recognition [16], [17].

(IV). Resistance gene is characterized by an eLLR, transmembrane and serine/threonine protein kinase domain, such as Xa21 rice gene and the Fls2 gene of Arabidopsis thaliana [18], [19].

(V). The fifth class of resistance genes contain a transmembrane (TM) protein domain, fused to the coiled-coil (CC) domain, such as RPW8 gene of Arabidopsis.

(VI). Gene contains eLRR and transmembrane domain fused to PEST (Pro-Glu-Ser-Thr) domain. PEST domain is responsible for protein degradation and short protein motifs or RME (receptor mediated endocytosis). The tomato genes Ve1 and Ve2 belong to this category.

(VII). Genes contain a TIR domain, NBS domain and LRR domain associated with NLS (nuclear localization signal) and WRKY domain. The WRKY domain is a 60-
amino acid region which is defined by consensus amino acid residues WRKYGQK at its N-terminal end, associated with novel zinc finger like motif. The \textit{RRS 1} gene is an example of this class.

(VIII). Genes is attributed by one of following: \textit{Hm 1} gene in maize [20], encodes detoxifying enzymes (Toxin reductase). The tomato \textit{Pto} gene; the \textit{At1} gene of melon [21], [22] contain a protein kinase (PK) domain. The \textit{Rpg1} gene of barley encodes for a receptor like kinase protein that contains two protein kinase (kinase-kinase) domains [23].

\textbf{Figure 1: Different classes of resistance gene.} 1. CC (Coiled coil) -NBS (nucleotide binding site) -LRR 2. TIR (Toll interleukin-1 receptor) -NBS-LRR 3. Extra cellular leucine rich repeat (eLRR) -Trans membrane (TM) 4. eLRR-TM -Protein kinase (PK) 5. TM -CC 6. eLRR -TM - PEST 7. TIR-NBS-LRR-NLS (nuclear localization signal)-WRKY 8. Protein Kinase and Kinase- Kinase.

\textbf{Table 1:} R-gene structure and different classes of Plant resistance gene

<table>
<thead>
<tr>
<th>Class</th>
<th>Structure</th>
<th>R-gene</th>
<th>Plant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CC-NBS-LRR</td>
<td>RPM1, RPS2, I2</td>
<td>\textit{Arabidopsis}, Tomato</td>
</tr>
<tr>
<td>2</td>
<td>TIR-NBS-LRR</td>
<td>N, L6, RPP5</td>
<td>Tobacco, flax</td>
</tr>
<tr>
<td>3</td>
<td>eLRR-TM</td>
<td>Cf-2, Cf-4, Cf-9</td>
<td>Tomato</td>
</tr>
<tr>
<td>4</td>
<td>eLRR-TM-PK</td>
<td>Xa 21, Fls2</td>
<td>Rice, \textit{Arabidopsis}</td>
</tr>
<tr>
<td>5</td>
<td>TM-CC</td>
<td>RPW-8</td>
<td>\textit{Arabidopsis}</td>
</tr>
<tr>
<td>6</td>
<td>eLRR-TM-PEST</td>
<td>Ve-1, Ve-2</td>
<td>Tomato</td>
</tr>
<tr>
<td>7</td>
<td>TIR-NBS-LRR-NLS-WRKY</td>
<td>RRS-1</td>
<td>\textit{Arabidopsis}</td>
</tr>
<tr>
<td>8</td>
<td>Toxin reductase</td>
<td>\textit{Hm-1}</td>
<td>Maize</td>
</tr>
<tr>
<td></td>
<td>Kinase</td>
<td>\textit{Pto, At-1}</td>
<td>Tomato</td>
</tr>
<tr>
<td></td>
<td>Kinase Kinase</td>
<td>\textit{Rpg 1}</td>
<td>Melon</td>
</tr>
</tbody>
</table>
The NBS-LRR class of resistance genes grouped into Toll NBS-LRR (TNL) and CC-NBS-LRR (CNL) groups [24], [25]. Both TNL and CNL are commonly referred as NBS-LRR class of resistance genes. The CNL groups of genes are found most frequently in monocots as well as dicots whereas TNLs have been identified mostly in dicots [26]. The NBS-LRR class of resistance genes is extensively used to isolate and identify candidate resistance genes from different valuable plants. The NBS-LRR protein consists of two domains, the NBS and LRR domain [27], [28]. The NBS domain of NBS-LRR class of resistance gene consist of around 300 amino acid and four conserved motifs viz. P loop (kinase 1a), kinase 2, kinase 3a and GLPLAL motif [24]. The C-terminal LRR domain consists of many leucine rich sequences and helpful in recognition of pathogen ligand. Fascinatingly, a single amino acid residue of the extremely conserved Kinase 2 motif (LLVLDDVW/D) can be used to determine the two subfamilies (TNL and CNL) with 95% accuracy. TIR is signified by an aspartic acid (D) residue; whereas, tryptophan (W) is replaced with the D in the case of Coiled coiled or Non-TIR [29].

(iii) Quantitative Disease Resistance (QDR)

Recently, QDR genes turn out to be significant part of research due to their wide range of resistance against pathogen like bacteria, virus, fungi and nematodes [6]. Number of QDR genes have been identified which provide durable and broad spectrum resistance against pathogen. The ABC transporter *Lr34* gene of wheat confers QDR to both ascomycete and basidiomycete [30]. Soybean gene *Rgh1* and *Rgh4* provide QDR against nematodes [31], [32]. Rice *Pi21* gene provide resistance to pathogen causing blast disease [33]. maize associated wall kinase *ZnWak* gene provides quantitative resistance against fungus *Sporisorium reilianum* which causes head smut disease in maize [34]. The *Htn1* gene of maize a wall associated receptor like kinase confers resistance to *Exserohilum turcicum* causing northern corn leaf blight [35]. Wheat *Yr36* gene confers resistance against fungal pathogen of wheat stripe rust [36]. The *RKS1* gene of *Arabidopsis* provides resistance to pathogen *Xanthomonas campestris* [37]. The Cloning of these genes may help in opening of the new avenue for enhancement of durability, understanding of the mechanism of resistance provide by these genes and their cross connection to R gene [6].

(iv) Resistance Gene Analogue (RGAs) newly develop tool to identify disease resistance gene

Isolation of resistance genes has traditionally involved map-based cloning and transposon tagging, both of which are very laborious and expensive approaches [38]. Generally, two common methods are used to identify candidate disease resistance gene: the PCR-based method and data mining. In PCR-based methods, the degenerate primers derived from the conserved domains of known R genes are used to identify resistance gene analogues (RGAs) or candidate disease resistance genes from a number of plant species [39]. Data mining has been successfully used for
identification of RGAs from sugarcane, wheat, maize and bean [40], [41],[27], [32]. In case of identification of RGAs through data mining of EST sequences, identified RGAs belong to expressed genes, whereas the RGAs identified from genome sequences may be unexpressed pseudogenes. In plant kingdom, most of the cloned and functionally identified resistance genes to date are classified in different classes on the basis of their predicted protein structure. Using data mining and PCR, several RGAs have been isolated by designing of degenerate primers against conserved motifs present in NBS-LRR class of gene in different crops worldwide. In China, Que et al. [42] reported 11 RGAs from smut resistant sugarcane variety-NCo376 [46]. Fifteen RGAs were identified from a red rot resistant variety of sugarcane in Pakistan [43]. Bertioli et al., [44] identified 78 RGAs in Peanut and its wild relatives with degenerate primers based on NBS domain. Gao et al., [39] identified 100 RGAs in Tobacco based on NBS and PK domain of resistance gene. In coconut, 243 RGAs were identified [45]. McIntyre et al., [46] isolated 54 RGAs from sugarcane species from which 10 were ESTs based and remaining 44 were based on conserved motifs of RGAs . Liu et al., [47] identified 385 expressed peanut RGAs from peanut ESTs GenBank data base. Srivastava et al., [48] identified 28 putative RGAs from sugarcane based on NBS-LRR and kinase domain motif.

(v) Bioinformatics and Putative disease resistance gene

Various bioinformatics methods such as multiple sequence alignment, BLAST search, Phylogenetic analysis and domain and motif analysis play important role in identification of RGAs and prediction of their function. In the identification and characterization of putative disease resistance genes, some common bioinformatics steps along with degenerate PCR method are involved in RGAs isolation (Figure 2). Initially the collection of resistance gene candidate and protein database from nearest members is required by using NCBI data base and second step is multiple sequence alignment (MSA) to identify conserved motifs in resistance gene homologues. Such identified conserved motifs are helpful in designing of degenerate primers. The degenerate primers are derived from the conserved domain of other cloned resistance genes also.

Generally, there is not a single universal technique to design degenerate primers and RGA identification in various plants. The designing of degenerate primer is a critical thing in identification of RGAs. This followed by PCR based amplification of DNA and sequencing of desirable amplicon. Finally, BLAST is done for sequence similarity and grouping of identified RGAs into different classes according to observed conserved motifs or domains or combination of both. The designing of degenerate primers is generally carried out by bioinformatics tools involving multiple sequence alignment of protein or nucleotide sequence to characterize the conserved motifs or regions in them and using these motifs for primer designing by software such as CODEHOP [49]. This case of primer designing requires high similarities among conserved regions of proteins in global multiple alignments. To reduce this problem, new methods have been employed for designing of degenerate primer which
involve local multiple sequence alignment based on Multiple Expectation – Maximization for motif Elicitation algorithm (MEME) is done to find conserved regions [50]. Protein family database such as Pfam, Inter ProScan and MARCOIL are used to determine the functional protein domain, such as CC or TIR domain in NBS-LRR class of resistance genes. A program COILS, is used to detect coiled coil (CC) domain in protein families [51]. Another aspect required to take care in identification of resistance genes, is that, identified resistance genes should not possess pseudogenes. To overcome this problem, use of cDNA in place of genomic DNA is useful because they have expressed sequence.

Figure 2: The steps involved in identification of RGAs
(vi) Deployment of phytoresistance gene in crop improvement

The genes involved in Qualitative resistance and quantitative resistance are significant in providing resistance against pathogens, widely used in crop improvement. The R gene-mediated resistance or effector triggered immunity (ETI) in crop varieties involves direct or indirect recognition of complementary pathogen effector protein which halting the disease. The resistance provided by R-genes is a major strategy to control the diseases. The use of host resistance generally involves introduction of single or multiple genes into susceptible cultivars via backcrossing or more recently genetic engineering approaches. Through the GM (genetically modified) plant approach, a minimum of 10-15 years could be saved as compared to the conventional breeding approaches. However, R-genes are vulnerable to evolutionary potential of pathogens as a result of which resistance provided by single R- gene is not durable as observed in many cases, e.g. resistance mediated by Rlm gene was overcome within five years of its deployment in Brassica [52]. Several recent studies show that the management strategies of resistance source could be useful for achieving durability of race specific resistance genes. With regards to durability, the selection of R-genes and previous knowledge about genetic background of susceptible cultivar into which R-gene introgression applied is very important [53], [54]. Gene stacking or pyramiding or combination of genes in a single genotype in place of single R gene is expected to overcome the evolutionary effect of pathogen virulence thereby increasing durability. Marker assisted gene pyramiding has been used successfully for powdery mildew resistance in wheat [55] whereas bacterial blight resistance in rice by Huang et al., 2004 [56]. Recently, Fukuoka et al., [57] reported that the pyramiding of QTL alleles enhances blast disease resistance in rice [57]. The sequential or alternate use of different R genes in rotation may probably reduce the emergence of new virulent strain and may also contribute to durability [58]. Gene stacking and rotation of R-genes lowers the evolutionary risk posed by the pathogen in comparison to single R gene deployment in monoculture. Genetic engineering can be used to attain durability and broad spectrum resistance through introgression of gene stacking of R-genes into susceptible cultivars. Depending on the crop, stacking via genetic engineering may commonly be more practical as compared to other breeding techniques [59]. For instance, by using this approach durability and broad spectrum of resistance have been reported in potato [60], [61], [62]. From the point of view of durability and contribution of QDR genes in combination with R- gene for broad spectrum resistance may be significant as supported by research findings. An R-gene pair (RPS4/RPS1) in Arabidopsis forms a complex in vivo and provides broad spectrum resistance against bacterial and fungal pathogen [63].This kind of R-gene cooperation may be deployed in different crops to enhance durability and broad spectrum resistance. It has been noticed that the transfer of R-genes between unrelated species shows good success rate in attainment of durable disease resistance [64]. Further studies are under way for how to achieve durability of resistance using R genes.
CONCLUSION AND FUTURE PROSPECTIVE

Phytopathogens are a serious threat to plant and crop productivity. A sustainable and eco-friendly biological strategy is needed to overcome this problem. Identification of disease resistance gene using RGAs opens new avenues to develop molecular marker linked with these resistance genes in different crops. Despite the identification of a number of candidate resistance gene in various plants has remained a great challenge due to clustering of resistance genes and occurrence of polyploidy in many plants. The use of resistance genes could become a unique approach to develop resistant cultivars and control various plant diseases. The development of a disease resistant variety through the insertion of cis resistance gene and trans resistance gene will be required in future taking concern to the ethical policies. Another aspect is durability and broad spectrum resistance in resistant cultivar which is very important for developing a disease resistant variety. Several research approaches are in progress to achieve such durability and broad spectrum disease resistance. Hopefully, the coming time will be of genetically modified disease resistant crops to face such upcoming challenge ahead.

REFERENCE


RGAs Approach in Identification of Disease Resistance Genes and their

rust resistance,” Genome., 48(3), 391-400.


7773.


