Austin Food Sciences

Open Access

Austin Publishing Group

Editorial

Approaches for Disease Resistant Candidate Genes Identification in Plants: Recent Techniques and Trends

Rawat N*
Gulf Coast Research and Education Center, University of Florida, USA

***Corresponding author:** Nidhi Rawat, Gulf Coast Research and Education Center, University of Florida, USA

Received: May 11, 2016; **Accepted:** May 13, 2016; **Published:** May 17, 2016

Editorial

Global food supply for the thriving world population is a big challenge. Annual production of economically important crops such as wheat, rice, maize is adversely affected by several diseases caused by various pests and pathogens resulting almost 20-40% yield loss [1]. Hence, to ensure sustainable food security, we need to engineer long-lasting and broad-spectrum disease resistance in crops. Development and utilization of disease-resistant/tolerant cultivars have been considered as one of the most important tools for mitigating diseases in agronomic, horticultural or forestry crops. One of the major objectives of plant research is to improve our understanding on plant immune system and unravel how this is manipulated by pathogens, in order to engineer crops for durable resistance against pathogens [2]. While the identification of Candidate Genes (CGs) involved in defense response remains a challenge, several strategies have been developed in this subject in the last decade. CGs identification approaches can be broadly classified in three groups.

Genetic Approaches

First instance of disease resistance traditional breeding was reported in wheat [3]. However, at that time, traditional breeding programs were used to identify resistant sources in crops by classical methods and introgressing them into economically important crops by crossing without any knowledge about Resistance (*R*) genes and their respective mechanisms. Major breakthrough in the understanding of plant disease resistance as 'gene-for-gene' concept was designed by the work of Flor [4] which defined that an Avirulence (*Avr*) gene in the pathogen and an *R* gene in the host plant are required to mount an immune response in plant. Currently, transgenic approaches are being used to reduce the estimated time on classical breeding by several years to create disease resistant crops [5]. For the successful deployment of disease resistance genes, we need to have handful of resistance gene candidates to create disease resistance cultivars. CGs involved in defense responses can be broadly classified into two groups. The first group is for the genes those are involved in the initial recognition of pest or pathogen, the resistance (*R*) genes and the other group is for the genes those are involved in Defense Response (DR) triggered by the recognition event [6]. CG approach has proven extremely significant for studying the genetic architecture of complex traits including disease resistance. Genetic linkage analysis (mapping) has been commonly used in plants to screen CGs and subsequently clone these genes for controlling disease resistance/tolerance traits. Genetic maps of markers for genes can assist in choosing positional CGs. This is accomplished by evaluating the closeness of linkage with the trait loci. The CG approach has been used to characterize disease resistance loci. Numerous disease candidate genes involved in pathogen recognition and defense response have been isolated in various crops such as in rice against gall midge insect infestation [7-10]. Linkage analysis involves creating bi-parental populations, genotyping and phenotyping of segregating progeny in the populations, and testing if sequence variations in the CGs co-segregate or co-localize with the loci controlling the disease resistance/tolerance trait in the populations [11]. Association mapping has been widely used in human and animal genetics, and it has emerged as a new approach to screen candidate genes in plants [11,12]. This approach uses existing natural populations and germplasm accessions, collections of cultivars, breeding lines from breeding programs, and statistical analysis to reveal association between sequence variations in candidate genes and phenotypic variations in the association study panels.

Both the approaches have their own advantages and disadvantages [11,13]. Segregating populations need to be constructed for linkage analysis which requires longer time and more resources, but it offers better experimental control of plant materials to be genotyped and phenotyped. On the other hand, association study takes advantage of linkage disequilibrium and historical and evolutionary recombination events, therefore improving the mapping resolution [13]. Since it examines a more diverse pool of germplasm beyond the parents used in linkage analysis, association study may discover additional genes or alleles for the trait of interest. However, the study panels selected for association study are "uncontrolled" populations among the panel members that may cause spurious associations [12]. These two approaches are being widely used complementarily in plants, with one approach to discover candidate genes, and the other to validate the discovery [11].

Genomics Tools

Identifying sequence variations in CGs and determining the genotype of progeny in segregating populations are the most essential steps in linkage or association analysis; yet, for many decades, these steps were the most limiting factors. Thanks to the development of massively parallel target capture/enrichment and next-generation DNA sequencing tools, large-scale sequence variation identification and genotyping are no longer a technical obstacle [14]. Genomic approaches for identification of expressed genes such as Expressed Sequence Tag (EST) using Suppression Subtractive Hybridization (SSH) [15], Serial Analysis of Gene Expression (SAGE) [16] and Massively Parallel Signature Sequencing (MPSS) [17] have been widely used in genome-wide gene expression studies in various organisms.

Citation: Rawat N. Approaches for Disease Resistant Candidate Genes Identification in Plants: Recent Techniques and Trends. Austin Food Sci. 2016; 1(2): 1010.

SAGE and MPSS are two powerful tools for deep transcriptome analysis and have been developed to evaluate the expression patterns of thousands of genes in a quantitative manner without prior sequence information [16,17]. However, complicated cloning procedures involved in the SAGE and MPSS library construction have inhibited a wide use of these two methods in plant species [18]. Among these techniques, SSH has high subtraction efficiency and harbors an equalized representation of differentially expressed sequences which can separate effectively both high and low copy expressed genes mainly because of normalization. EST sequencing was the first method used for rapid identification of expressed genes. In the last several years, many defense related genes have been isolated using Real-time PCR, Affymetrix microarray chips, SSH and cDNA library differential screening methods [19-22]. Among available next-generation sequencing systems, Illumina HiSeq system is currently the most widely used platform in many genome sequencing and re-sequencing projects for its high accuracy, unrivalled output, and superior cost-effectiveness. HiSeq 2000 can generate up to 55 Gb per day from a 2x100 bp run, which enables genome sequencing and genotyping-by-sequencing completed in an unprecedented fashion. The availability of this system has greatly accelerated the discovery of new CGs and allelic variations among susceptible and resistant plant genotypes [23-26].

Next-generation sequencing is used in different ways to address various questions in plant genomics. Whole Exome Sequencing (WES) by high-throughput sequencing of target-enriched genomic DNA (exome-seq) has become common in basic and translational research. Exome sequencing is a capture based method developed to identify variants in the coding region of genes. Exome sequencing has the advantage that oligonucleotides probes (baits) are hybridized to genomic DNA to capture the RNA coding regions and provides better coverage for SNP calling. It is also more economical compared to Whole Genome Sequencing (WGS) allowing analysis of more individuals and populations. However, WGS is more comprehensive and structural variants could be detected in non-coding regions such as promoters regions of disease related CGs. Currently three major exome enrichment platforms are available; NimbleGen's SeqCap (http://www.nimblegen.com/seqcap/), Illumina's TruSeq Exome Enrichment (http://www.illumina.com/techniques/sequencing/) and Agilent Technologies' SureSelect (http://www.genomics.agilent.com/). The first is an array-based, while the other two are solution-based hybridization capture systems. The technologies diverge in their choice of target such as DNA for Nimblegen and Illumina, and RNA for Agilent, bait lengths, bait density and molecule used for capture. These enrichment methods of sequencing have been successfully used to capture and enrich tens of thousands of genes [27]. Nevertheless, SureSelect system needs smaller amounts of input DNA and is easier to multiplex samples and automate sample preparation, which seems to have better specificity and reproducibility [28].

Bioinformatics Approach

In modern world, *in-silico* methods are being used for investigation, establishment, and classification of disease resistant CGs. With the increase in accessible data from numerous plant genomes databases and functional genomics information, methods of identifying CGs are rapidly evolving. Genome databases are now an integral part in the process of candidate disease genes selection. Merging the positional information of CGs from linkage analysis along with the functional characteristics is the usual approach by which candidate disease genes are selected [8,22]. Among various *in-silico* methods available for CGs identification, one of the most commonly used strategy is meta-analysis where converging output of experimental data from different labs with different experimental conditions are mined using same statistical algorithms. Meta-analysis of microarrays and RNA-Seq has been used extensively in animal systems to define robust, regulated probe sets [29]. Recently, meta-analyses have also been used to identify differentially expressed probe sets in plants [30-32]. The other popular ways for CGs identification are HMM search, gene ontology, comparative genomics, molecular evolution, machine learning and cluster analysis [33]. *In-silico* ways from prioritizing the CGs are done by identification of gene structure variation, orthologs, protein-DNA interactions, and co-expression gene analysis and protein-protein regulatory network. These techniques place the potential disease-causing proteins in a functional context, comparative to other known disease related genes, and thus, systematic investigation of such complexes might unmask new candidate genes.

In conclusion, a manageable list of CGs is essential for direct evaluation of these genes with resistance/susceptibility for a given crop disease. The combination of above mentioned CG identification methods such as association mapping, linkage analysis, next-generation sequencing and *in-silico* analysis yield CGs those are required as the first phase for functional analysis of the genotype-phenotype relationship.

References

1. Savary S, Ficke A, Aubertot JN, Hollier C. Crop losses due to diseases and their implications for global food production losses and food security. Food Security. 2012; 4: 519-537.
2. Dangl JL, Horvath DM, Staskawicz BJ. Pivoting the plant immune system from dissection to deployment. Science. 2013; 341: 746-751.
3. Biffen RH. Mendel's laws of inheritance and wheat breeding. J Agr Sci. 1905; 1: 4-8.
4. Flor HH. Host-parasite interaction in flax rust-its genetics and other implications. Phytopathology. 1955; 45: 680-685.
5. Haverkort AJ, Boonekamp PM, Hutten R, Jacobsen E, Lotz LA, Kessel GJ, et al. Societal costs of late blight in potato and prospects of durable resistance through cisgenic modification. Potato Research. 2008; 51: 47-57.
6. Dixon RA, Harrison MJ. Activation, structure, and organization of genes involved in microbial defense in plants. Adv Genet. 1990; 28: 165-234.
7. Sama VS, Rawat N, Sundaram RM, Himabindu K, Naik BS, Viraktamath BC, et al. A putative candidate for the recessive gall midge resistance gene gm3 in rice identified and validated. Theor Appl Genet. 2014; 127: 113-124.
8. Yasala AK, Rawat N, Sama VA, Himabindu K, Sundaram RM, Bentur JS. *In silico* analysis for gene content in rice genomic regions mapped for the gall midge resistance genes. Plant Omics. 2012; 5: 405.
9. Hammond-Kosack KE, Jones JD. Resistance gene-dependent plant defense responses. Plant Cell. 1996; 8: 1773-1791.
10. Bentur JS, Rawat N, Divya D, Sinha DK, Agarrwal R, Atray I, et al. Rice-gall midge interactions: Battle for survival. J Insect Physiol. 2016; 84: 40-49.
11. Pflieger S, Lefebvre V, Causse M. The candidate gene approach in plant genetics: a review. Molecular breeding. 2001; 7: 275-291.
12. Khan MA, Korban SS. Association mapping in forest trees and fruit crops. J Exp Bot. 2012; 63: 4045-4060.

13. Zhu C, Gore M, Buckler ES, Yu J. Status and prospects of association mapping in plants. The Plant Genome. 2008; 1: 5-20.

14. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A robust, simple Genotyping-By-Sequencing (GBS) approach for high diversity species. PLoS One. 2011; 6: 19379.

15. Rounsley SD, Glodek A, Sutton G, Adams MD, Somerville CR, Venter JC, et al. The construction of Arabidopsis expressed sequence tag assemblies. A new resource to facilitate gene identification. Plant Physiol. 1996; 112: 1177-1183.

16. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. Science. 1995; 270: 484-487.

17. Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, et al. Gene expression analysis by Massively Parallel Signature Sequencing (MPSS) on micro bead arrays. Nature Biotechnol. 2000; 18: 630-634.

18. Gowda M, Venu RC, Raghupathy MB, Nobuta K, Li H, Wing R, et al. Deep and comparative analysis of the mycelium and appressorium transcriptomes of *Magnaporthe grisea* using MPSS, RL-SAGE, and oligoarray methods. BMC Genomics. 2006; 7: 310.

19. Lu G, Jantasuriyarat C, Zhou B, Wang GL. Isolation and characterization of novel defense response genes involved in compatible and incompatible interactions between rice and *Magnaporthe grisea*. Theor Appl Genet. 2004; 108: 525-534.

20. Rawat N, Himabindu K, Neeraja CN, Nair S, Bentur JS. Suppressive subtraction hybridization reveals that rice gall midge attack elicits plant-pathogen-like responses in rice. Plant Physiol Biochem. 2013; 63: 122-130.

21. Rawat N, Chiruvuri Naga N, Raman Meenakshi S, Nair S, Bentur JS. A novel mechanism of gall midge resistance in the rice variety Kavya revealed by microarray analysis. Funct Integr Genomics. 2012; 12: 249-264.

22. Rawat N, Neeraja CN, Nair S, Bentur JS. Differential gene expression in gall midge susceptible rice genotypes revealed by Suppressive Subtraction Hybridization (SSH) cDNA libraries and microarray analysis. Rice. 2012; 5: 1-5.

23. Jupe F, Witek K, Verweij W, Sliwka J, Pritchard L, Etherington GJ, et al. Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. The Plant J. 2013; 76: 530-544.

24. Quesada T, Gopal V, Cumbie WP, Eckert AJ, Wegrzyn JL, Neale DB, et al. Association mapping of quantitative disease resistance in a natural population of loblolly pine (*Pinus taeda* L.). Genetics. 2010; 186: 677-686.

25. Rawat N. Mining genome and transcriptome sequences of *Poncirus trifoliata* for NB-LRR genes and their association with HLB disease resistance. In Plant and Animal Genome XXIV Conference. Plant and Animal Genome.

26. Rawat N. Genome-wide identification, characterization and expression analyses of constitutive disease resistance gene and miraculin gene families in citrus. In Plant and Animal Genome XXIV Conference. Plant and Animal Genome.

27. Montenegro G, Powell E, Huang J, Speziani F, Edwards YJ, Beecham G, et al. Exome sequencing allows for rapid gene identification in a Charcot-Marie-Tooth family. Ann. Neurol. 2011; 69: 464-470.

28. Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, et al. Target-enrichment strategies for next-generation sequencing. Nat Methods. 2010; 7: 111-118.

29. Nazri A, Lio P. Investigating meta-approaches for reconstructing gene networks in a mammalian cellular context. PLoS One. 2012; 7: 28713.

30. Shaik R, Ramakrishna W. Machine learning approaches distinguish multiple stress conditions using stress-responsive genes and identify candidate genes for broad resistance in rice. Plant Physiol. 2014; 164: 481-495.

31. Rawat N, Kiran SP, Du D, Gmitter FG, Deng Z. Comprehensive meta-analysis, co-expression, and miRNA nested network analysis identifies gene candidates in citrus against Huanglongbing disease. BMC Plant biol. 2015; 15: 1.

32. Du D, Rawat N, Deng Z, Gmitter FG Jr. Construction of citrus gene coexpression networks from microarray data using random matrix theory. Hortic Res. 2015; 2: 15026.

33. Zhu M, Zhao S. Candidate gene identification approach: progress and challenges. Int J Biol Sci. 2007; 3: 420-427.

Austin Food Sci - Volume 1 Issue 2 - 2016
Submit your Manuscript | www.austinpublishinggroup.com

Citation: Rawat N. Approaches for Disease Resistant Candidate Genes Identification in Plants: Recent Techniques and Trends. Austin Food Sci. 2016; 1(2): 1010.