Original Research Article

https://doi.org/10.20546/ijcmas.2019.805.036

Conceptual Advances in Pathogen Effector Related Studies

Ankita¹* and Vijay Kamal Meena²

¹Department of Plant Pathology, Dr. Yashwant Singh Parmar University of Horticulture and Forestry, Nauni, Solan, Himachal Pradesh, India
²Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India

*Corresponding author

A B S T R A C T

Plants are always under the radar of the pathogen, with their different means of infecting them. And conventional methods of managing plant diseases are showing transient effectiveness, as they never last long. Introduction of next generation sequencing, has provided with whole lot of information regarding host and pathogen. It analyses a large number of pathogen isolates and host varieties rapidly at lower cost. In this host pathogen dynamics, effector molecules play a very crucial role. Genomic approaches like RNA analysis, microarrays have unveiled the role of effectors in plant host responses and have advanced our knowledge in terms of pathogen’s potential virulence factor, pathogenicity island, information related to R genes and its cloning, effector diversity of pathogen and global monitoring of pathogen. Through this paper different aspects of genomics as well as role of effectors in plant immunity are reviewed.

Keywords
Pathogen, Genomic approaches, R genes, Potential virulence factor

Article Info
Accepted: 04 April 2019
Available Online: 10 May 2019

Introduction

There is a constant threat to all the plants present in our ecosystem from the diseases of fungal, bacterial, viral and other sources. Directly and indirectly they are affecting every aspect of our life, be it food sources, timber industry, medicinal sources and etc. Different methods are employed in order to manage these diseases. Right from the chemical method, physical methods, biological methods, resistance breeding and many more but each method has its own effectivity and limitations. Nowadays many chemicals which were once effective have now lost its effectiveness, leading to the development of resistant strain and deterioration of the environment. The other way around, in order to avoid the use of chemicals, resistance breeding is a good option. It is cost-effective and ecologically sound as well. But the conventional methods of breeding are not very effective because they are somewhat negligent towards the pathogen, so with the inclusion of genomics approach, different generation of sequencing material and its decoding has helped in increased understanding and decrypting code that revolves around the host, the pathogen and their interaction giving us a new insight.
on R genes and its utilization in breeding programme.

Entering into 21st century, many geneticists began to change their focus from genes to genomes. The change in their focus from genes is made possible by large-scale DNA sequencing that has revealed the complete genomic sequence of many organisms. The genome sequencing was accompanied by advanced computational methods for recognizing genes and for comparing genomes. Soon methods were devised for identifying which genes in the genome were transcribed in particular tissue types, at specific times in development, or at different stages of the cell cycle. Going beyond transcription, methods were also devised for identifying which proteins in the cell interact physically with which other proteins. New terms were coined to emphasize the changes in perspective. The field of genomics deals with DNA sequence, organization, function and evolution of genomes.

Identifying the pathogen is the core of the whole process, from there we derive information’s related to the pathogenicity of the microorganism. Genomic study of the pathogen by following different genome sequencing methods will unleash significant amount of information related to the identification of potential virulence factors, pathogenicity islands, and many other genes that could be related to therapeutic targets of plants. Also the selective forces driving evolution of new pathotypes allowing the prediction of pathotypes that may act as a threat to particular genotypes within specific geographic regions and provide an early-warning system for crop vulnerabilities (Derevnina and Michelmore, 2015).

Genome sequencing and new genomics studies, such as pathogenomics has revealed a new area of study of microorganisms and discovered unexpected aspects of pathogen biology (Pallen and Wren, 2007). Sequencing of isolates from different geographical areas will reveal the prevalence and distribution of known virulence genes as well as uncover new virulence genes. High-throughput genomics offers the possibility for analyzing a large number of pathogen isolates and host varieties rapidly and at low cost. It detects genes being expressed and therefore the determinants of the interaction thus, non-expressed genes present in the genome do not obscure genotype-phenotype correlations.

History of genomics

Genomics is a quite new discipline. Although the DNA was first isolated as early as 1869 (Dahm et al., 2004), it took more than one century for the first genomes to be sequenced (Anderson et al., 1981). The history of modern genomics begins in the 1970s, however, it was possible only due to important discoveries in the field by a small group of scientists in the period after the Second World War. It would therefore be unfair to start history of genomics without mentioning the breakthroughs two decades earlier.

Modern genomics would not be possible without the technological advances in the 1950s such as creating isotopes and radiolabel biological molecules (Chargaff et al., 1950), while the greatest breakthrough was most certainly the description of the structure of the DNA helix that was made by James D. Watson and Francis H. C. Crick in 1953 (Watson and Crick 1953). This allowed determination of the DNA replication, gene expression, protein synthesis, etc. Just as important were the later technological advances and advances in methodology, in the first place the automated DNA sequencing and polymerase chain reaction (PCR), a technique which was developed in the early
1980s and allowed DNA amplification from extremely small amounts of material (Mullis et al., 1990). Automated DNA sequencing, on the other hand, dramatically increased the pace of genome sequencing. Both automated DNA sequencing and PCR played the key role in the Human Genome Project which, however, was completed only in 2007 (Hood et al., 2013).

The first genome was sequenced by Fred Sanger who sequenced the genomes of a virus and mitochondrion in the early 1970s. He and his team also created techniques for sequencing, data storage, genome mapping and other techniques and methodologies which still play an important role in genomics. For his contribution to genomics, he was awarded the Nobel Prize in Chemistry in 1980 (for the second time) (Anderson et al., 1981). Genome sequencing of different organisms still remains an important part of genomics, however, the knowledge of full genomes paved the way to the so-called functional genomics which is focused on the study of DNA at the gene level with an aim to understand the relationship between the organism’s genome and phenotype.

Walter Fiers is another scientist who played an important role in modern genomics. In 1972, he and his research group from the Laboratory of Molecular Biology of the University of Ghent in Belgium became the first to sequence a gene. They sequenced the gene of Bacteriophage MS2, an RNA virus which infects the E. coli bacterium. Thus Bacteriophage MS2 became the first organism to be completely sequenced (Fiers et al., 1976).

In 1995, Hamilton O. Smith and his team from The Institute for Genomic Research became the first to sequence a genome of a free living organism – that of *Haemophilus influenza* (Fleischmann et al., 1995). Since then, genomes of various organisms including human (mostly sequenced in 2001, completed in 2007) have been sequenced at a great pace. Today, we have complete sequences for more than 2,700 viruses, over 1,000 bacteria and archaea and 36 eukaryotes, with the new ones being sequenced virtually on a daily basis.

**Classification of genomics**

Geneticists have used the term genome for over seven decades to refer to one complete copy of the genetic information or one complete set of chromosomes (monoploid or haploid) of an organism. In contrast, genomics is a relatively new term. The word genomics appears to have been coined by Thomas Roderick in 1986 to refer to the genetics sub discipline of mapping, sequencing, and analyzing the functions of entire genomes, and to serve as the name of a new Journal-Genomics-dedicated to the communication of new information in this sub discipline (Snusted and Simmons, 2016).

As more detailed maps and sequences of genomes became available, the genomics sub discipline was divided into structural genomics- the study of the genome structure, functional genomics- the study of the genome function; and comparative genomics- the study of genome evolution. Functional genomics includes analyses of the transcriptome, the complete set of RNAs transcribed from a genome, and the proteome, the complete set of proteins encoded by a genome. Indeed, functional genomics has spawned an entirely new discipline, proteomics, which has as its goal the determination of the structures and functions of all the proteins in an organism (Snusted and Simmons, 2016). Whereas structural genomics is quite advanced with the complete nucleotide sequences available for many organisms, functional genomics is presently in an explosive growth phase. New array
hybridization and gene-chip technologies allow researchers to monitor the expression of entire genomes- all the genes in an organism- at various stages of growth and development or in response to environmental changes. These powerful new tools promise to provide a wealth of information about genes and how they interact with each other and with the environment (Snusted and Simmons, 2016).

Recent advancement in genomics that also serves in host pathogen interaction

Micro-array

Now a day’s development of new technologies allow scientist to produce microarrays and gene chips, which contain many hybridization probes on a single membrane. Microarrays can be classified into different categories like- Custom/spotted/two-color microarrays (cDNAs, BACs), High-density oligonucleotide arrays (GeneChip, Affymetrix), Long oligonucleotide microarrays (Agilent, Illumina, Nimblegen) (Bumgarner, 2013).

Microarray have many broad applications like (i) Gene expression analysis- In this application, RNA is extracted from the cells of interest and either, labeled directly, converted to a labeled cDNA or converted to a T7 RNA promoter tailed cDNA which is further converted to cRNA through the Eberwine amplification process (Van Gelder et al., 1990), (ii) Transcription factor binding analysis- Microarrays have also been used in combination with chromatin immunoprecipitation to determine the binding sites of transcription factors (Horek and Snyder, 2002). (iii) Genotyping- Microarrays have been widely used as single-nucleotide-polymorphism (SNP) genotyping platforms. Several alternative approaches have been used to detect SNP’s but the most commonly used are allele discrimination by hybridization as used by Affymetrix (Wanf et al., 1998).

Next-generation sequencing (NGS)

Next generation sequencing (NGS), massively parallel or deep sequencing are related terms that describe a DNA sequencing technology which has revolutionized genomic research (Behjati et al., 2013). All NGS platforms perform sequencing of millions of small fragments of DNA in parallel. Bioinformatics analyses are used to piece together these fragments by mapping the individual reads to the human reference genome (Behjati et al., 2013). Each of the three billion bases in the human genome is sequenced multiple times, providing high depth to deliver accurate data and an insight into unexpected DNA variation. NGS can be used to sequence entire genomes or constrained to specific areas of interest, including all 22 000 coding genes (a whole exome) or small numbers of individual genes (Behjati et al., 2013). Solexa (Illumina), 454 (Roche), SOLiD (Applied Biosystems) are the methods widely used in next generation sequencing.

Interaction between Host and Pathogens

During the evolution of host and pathogen over a period of time there is a ongoing war of existence between them where each tries to have a upper hand. And it is a never ending process because alternative to coexistence is elimination of either of them (Chaubé and Pundhir, 2008). The molecular genetic researches on host plant interaction have established that plants being sessile in nature has a well-established innate immune system to act against the pathogens. There is an appropriate response of plants emerged from the perception of a pathogen’s signature molecule called as elicitors that are recognized by the plant cell surface receptors. Elicitors are the signature molecules that are originated from the pathogen as well as from plant’s surface during the pathogen attack (Bonas, 2002). The innate immune system of plants helps to identify some alien molecule.
It detects the PAMP, characteristic of potentially harmful microorganism. Even endogenously plant cell component also produces general elicitors by the action of enzymes produced by the microbes. Among elicitors there are two types – General elicitors and specific elicitors.

General elicitors are pathogen derived molecules of diverse form and triggers the basal defense responses independently of the kind of host. They are indispensable in the normal functioning of microbes and thus provides a fitness penalty for the pathogen in cases where they are recognized by plant immune/surveillance system (Nurenberger et al., 2004).

As for example - General elicitors like Glucans, produced by Phytophthora and Pythium, derived from oomycete cell wall, induce phytoalexins. Chitin oligomers are produced by higher fungi, from chitin of fungal cell wall, induce phytoalexins and lignifications. Pectin oligomers, produced by fungi and bacteria, from degraded cell wall, inhibit proteins and defense genes. Harpins, by several gram-negative bacteria, part of type III secretion, cause HR and defense gene response.

Race-specific elicitors i.e. avr gene products, Avr proteins, by fungi and bacteria, in some cases promoting virulence, HR, and PCD. Elicitins, by Phytophthora and Pythium, scavengers of sterol, induce HR in tobacco. Enzymes, e.g., endoxygenase, by Trichoderma viride, fungal enzymes, induce defense genes and HR. Viral proteins, e.g., viral coat proteins, by TMV, structural component, HR in tobacco, tomato Protein (Agrios, 2005)

**Effectors**

But still pathogens are able to cause disease, because they have developed next level of mechanism to overcome PAMP triggered basal defense, where they interfere with the disease signaling by manipulating or suppressing the whole process by delivering effector proteins into the plant cells.

If the effector protein meets a matching R gene in the plant, it becomes a specific elicitor and the plant defense system is activated by the R protein (Rashid et al., 2010). Effectors are proteins expressed by plant pathogens to aid infection of specific plant species by altering different plant processes. Continuous evolution of effectors is posing a greater threat to the plants, even with the introduction of R genes. In such a situation the pathogen effector protein act as virulence factor that suppresses the second wave of plant host defense. This defense system is cultivar specific, gene-for-gene disease resistance system determined by pathogen-encoded effector proteins and the corresponding plant derived R proteins (Bonas et al., 2002). This interaction induces effector triggered immunity (ETI) in plants that results in the rapid, localized and necrotic programmed cell death termed as hypersensitive response that in turn limits the disease progression.

The act of resistance and susceptibility is a never ending process, so pathogens can also secrete effectors that suppress ETI and progression of disease continues. They can be used as molecular probes to study the structural changes that occur during plant infection at subcellular level (Bozkurt et al., 2012). Effector genes are under dynamic even opposite selection pressure depending on the host plant genotype that the pathogen is interacting with. All the relevant information related to effectors can be traced easily and more precisely with the advent of new genomics approach.

**Classes of effectors**

Effector proteins based on their target site in the host plant classified into two types-
Apoplastic and Cytoplasmic effectors. Apoplastic effectors are released into the plant outer cellular space, whereas cytoplasmic effectors are translocated inside the plant cell, where they target different subcellular organelles (Kamoun et al., 2006). Both classes of effectors are modular proteins with cleavable amino-terminal secretion signals. Cytoplasmic effectors carry an additional domain after the signal peptide that mediates translocation inside host cells and is defined by conserved motifs, such as the RXLR amino acid sequence (Kamoun et al., 2006). Some of the examples of effectors molecule presented in table 1.

**Effector gene expression**

Effectors are protein molecule that is encoded by the genes present in the pathogen. The main questions with the effectors are Why did they get expressed? How they get expressed? At which stage they get expressed?

The gene expression profiling in plant pathogenic microorganisms is dynamic and shows uneven expression of the gene with respect to the kind of cell and stage of the host, demonstrating an epigenetic expression. Only studying the expression profile of such genes wouldn’t be of much help, therefore high throughput gene expression analysis by c-DNA library, micro array and RNA-sequencing technology has helped in identifying the candidate virulence factors, that may be in the form of either DNA or RNA or protein (Raffaele et al., 2010). Such methodology will also help in investigating the role of different gene family that may have some role in the effector gene evolution. Effector genes are expressed to produce virulence factor in order to suppress basal defense responses. Such gene expression is very stage specific. Expression studies related to this is done by RNA analysis. As for example, in Phytophthora path system Cytoplasmic effector like RXLR will have its peak expression during the bio trophic stage of its life cycle, evident by higher level of RNA which reduces to the basal levels at the necrotrophic phase of infection (Haas et al., 2009).

In Phytophthora capsici CRN gene family having two distinct classes of CRN effectors, shows distinct expression pattern during the infection on tomato (Stam et al., 2013). Even several apoplastic effector genes encoding Cysteine rich secreted protein and protease inhibitors shows similar expression profile (Haas et al., 2009).

**Host plant responses to effectors**

Plants respond in multiple way to the pathogen effectors which is itself dependent on the genotype of the plant and the invading pathogen. Response of the host to the pathogen comes in the form of oxidative burst, callose deposition, expression of pathogenesis related (PR-) genes or Hypersensitive response (HR). Effectors can be used to give us an idea regarding the kind of cellular processes that are affected, by linking the effector sequences to plant phenotype and the processes. It also helps us to recognize effector, their progression and virulence activities (Oh et al., 2009).

Screening of effectors can be done by either in planta transient expression of effector genes or delivery of candidate effector protein by bacterial Type III secretion system (T3SS) (Fabro et al., 2011). It is performed on heterologous expression system as they will have higher intensity of gene expression. Here effectors can be screened effectively by linking them to different cellular process. As for example Screening of Hyaloperonospora arabidopsis RXLR effector for the suppression of callose deposition, (Fabro et al., 2011) fluorescently tagged effectors in
**Nicotiana benthamiana** and *P. capsici* CRN carboxy terminal domain, targeted the host nucleus (Stam et al., 2013) suggesting that the plant cell nucleus has a crucial role in virulence and immunity. It also helps us to understand the plant cell processes that take place during the host-pathogen interaction. We can assign functions to new immune receptors based on the virulence study of effectors (Oh et al., 2009).

**Table 1** Examples of Effector Molecules

<table>
<thead>
<tr>
<th>Effector</th>
<th>Species</th>
<th>Identification Strategies</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Apoplastic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Enzyme inhibitors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucanase inhibitors GIP1 and GIP2</td>
<td><em>P. sojae</em></td>
<td>Biochemical purification</td>
<td>(Rose et al., 2002)</td>
</tr>
<tr>
<td>Serine protease inhibitors EPI1 and EPI10</td>
<td><em>P. infestans</em></td>
<td>Bioinformatic prediction of secreted and in planta induced proteins coupled with protein domain annotation and followed by biochemical analyses</td>
<td>(Orsomando et al., 2001)</td>
</tr>
<tr>
<td>Cysteine protease inhibitors EPIC1 and EPIC2</td>
<td><em>P. infestans</em></td>
<td>Bioinformatic prediction of secreted and in planta induced proteins coupled with protein domain annotation</td>
<td>(Kamoun et al., 2006)</td>
</tr>
<tr>
<td><strong>B. Small cysteine-rich proteins</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INF1 elicitin</td>
<td><em>P. infestans</em></td>
<td>Biochemical purification</td>
<td>(Ricci et al., 1989)</td>
</tr>
<tr>
<td>PeF-like SCR74 and SCR91</td>
<td><em>P. infestans</em></td>
<td>Bioinformatic prediction of secreted and in planta induced proteins coupled with analyses of polymorphic diversity</td>
<td>(Kamoun et al., 2006)</td>
</tr>
<tr>
<td><strong>C. Nep1-like (NLP) family</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PaNie</td>
<td><em>Pythium aphanidermatum</em></td>
<td>Biochemical purification</td>
<td>(Bos et al., 2003)</td>
</tr>
<tr>
<td>PINPP1</td>
<td><em>P. infestans</em></td>
<td>Bioinformatic prediction of secreted proteins coupled with in planta expression</td>
<td>(Veit et al., 2001)</td>
</tr>
<tr>
<td><strong>2. Cytoplasmic effectors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. RXLR protein family</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATR1&lt;sup&gt;Nd264&lt;/sup&gt;</td>
<td><em>H. parasitica</em></td>
<td>Positional cloning</td>
<td>(Rehmany et al., 2005)</td>
</tr>
<tr>
<td>AVR3a</td>
<td><em>P. infestans</em></td>
<td>Bioinformatic prediction of secreted and polymorphic proteins coupled with association genetics</td>
<td>(Armstrong et al., 2005)</td>
</tr>
<tr>
<td>B. CRN protein family (CRN1 and CRN2)</td>
<td><em>P. infestans</em></td>
<td>Bioinformatic prediction of secreted proteins coupled with in planta expression</td>
<td>(Torto et al., 2003)</td>
</tr>
</tbody>
</table>

Use of effectors in large-scale screens of germplasm has facilitated the discovery of new resistance genes and their classification into discrete recognition specificities, accelerating the cloning of R genes while avoiding any repetition. Effectors can be used to identify R gene homologs in plant species that are more compatible for breeding. E.g. *Solanum stoloniferum* and *S. papita* for *P. infestans* RXLR effector. Monitoring of effector allele diversity in pathogen populations. This can provide valuable information to assess the potential of a given R gene regarding its spectrum and durability, and to design control strategies based on the dynamic distribution of virulence alleles in a
given population, allowing the early detection of races that can overcome the deployed R genes (Pais et al., 2013). Pathogenomics provides the data on host and pathogen needed for deployment of effective resistance genes. Global monitoring of the pathogen at continuous level will enable anticipatory strategies that take into consideration the frequency and relative fitness costs of the targeted virulence effectors. Continual surveillance of pathogen populations will allow the impact of different resistance genes and strategies to be monitored. It will facilitate rapid responses when resistance breaks down and provide enough room for temporal and spatial adjustments in resistance-gene deployment.

In conclusion, genomics has a substantial impact on both basic as well as applied plant biology. And effectors have turned out to be very important molecular probes for a variety of plant cellular processes and can impact on our knowledge of basic plant processes. Use of genomics approach in plant pathology has revealed how pathogenic microorganisms adapt to particular host, can overcome innate immune responses and changes the host range, as well as how evolution of new pathogenic species occur. Conceptual advances in plant pathology with the inclusion of genomics can be used to locate sources of new resistance genes, develop new methods for durable resistance, understanding the dynamics of pathogen in a particular geographical area and its movement in it.

References

Derevnina L and Michelmore RW. (2015). Wheat rusts never sleep but neither do sequencers: Will pathogenomics...
transform the way plant diseases are managed? Genome Biology, 16(1), 44. doi:10.1186/s13059-015-0615-3


Pallen MJ and Wren BW. (2007). Bacterial

**How to cite this article:**

doi: [https://doi.org/10.20546/ijcmas.2019.805.036](https://doi.org/10.20546/ijcmas.2019.805.036)