Structural Analysis of Resistance (R) Genes in Potato (*Solanum* Species) Genome

14

Soma S. Marla

Abstract

Late blight pathogenesis involves active interplay between effector proteins secreted by the oomycete *Phytophtera infestans* and immune receptors coded by resistance (R) genes in potato plant. This chapter discusses how computational structural studies of *P. infestans* effector and potato immune receptor molecules can assist in understanding this molecular cross-talk. Structural modeling and building three-dimensional structures of effector and resistant proteins enable analysis of their structures, compositions, and variability and can elucidate their functions involved in late blight pathogenesis. Predicted pathogen or enzyme-binding sites and molecular docking partially explain the mechanisms underlying virulence and the possible recognition or avoidance of *P. infestans* effector molecules. Predicted effector functions from the sequenced *P. infestans* genome and a comparison with available *Solanacae* gene databases enable germplasm screening and identification of corresponding new candidate R genes. Some clues obtained from bioinformatics structural investigations may potentially enrich our knowledge of the co-evolution of pathogen and host plant in late blight infection.

14.1 Introduction

Modern agriculture is characterized by monoculture and cultivation of only a few crops in large areas and thus shrinking the biodiversity. Consequently, some pathogens have gained a bigger foothold and rapidly have spread in the ecosystem, resulting in the wide incidence of serious diseases such as late blight caused by oomycete *Phytophtera infestans*. Potato resistance, as in

S. S. Marla (✉)
Genomic Resources Division, ICAR-National Bureau of Plant Genetic Resources, New Delhi, India
e-mail: soma.marla@icar.gov.in

S. K. Chakrabarti et al. (eds.), *The Potato Genome*,
Compendium of Plant Genomes, https://doi.org/10.1007/978-3-319-66135-3_14

other plants, subscribes to gene-for-gene interactions between host resistance (*R*) genes and corresponding pathogen avirulence (*Avr*) genes (Flor 1974). Late blight pathogen *P. infestans* secretes effector proteins that enter and alter the host plant's metabolism, defeating its immune system to advance the infection. The development of potato cultivars possessing durable resistance mandates has detailed the study of various molecular processes underlying host–pathogen interaction. Genomes of *P. infestans* (Haas et al. 2009) and potato (Potato Genome Sequencing Consortium 2011) have been sequenced. Genome data analysis using next-generation sequencing techniques allows identification of putative pathogen effectors, their diversity and their corresponding to candidate R genes. Application of computational techniques such as structure modeling helps in the comparison of effector homologues, the identification of ligand or enzyme-binding sites and may provide functional information about key enzymes involved in various processes such as host cell wall destruction, effector entry into the host cytoplasm or their recognition by NBS-LRR immune receptors and host proteases. The analysis of the structural biology of individual effectors (or matching the host receptors) in isolation or interaction with target molecules may provide valuable insights into understanding of host-pathogen interaction during late blight infection. The structural biology of pathogen effectors and their matching host immunity receptors may also possibly explain the molecular basis for effector recognition by NBS-LRR protein complexes, thus predicting various virulence functions. A major resource for in silico investigations is the availability of deposited crystal structural data in the Protein Data Bank (PDB) of at least a few effectors of *P. infestans*. Computational studies allow mapping of effector amino acid residues involved in interactions with NBS-LRRs receptors. Modeling three-dimensional structures of both effector and host R proteins that are docked help to predict pathogen/enzyme binding sites, and map the key amino acid residues or alleles guilty for possible loss of resistance functions in R genes (Fig. 14.1).

Fig. 14.1 Identification of sequence motifs, W, Y and L showing conserved pattern of amino acids in RXLR domains of *Avr3a* effectors in *P. infestans* (MEME tool output) (unpublished, personal communication)

Late blight pathogen *P. infestans* thrives on the host potato plant for nutrition and survival. To infect the host, *P. infestans* secretes extra-cellular effector proteins that enter the host plant and alter its metabolism to promote its own growth. Potato plants in turn have developed a complex and multilayered resistance mechanism to counter the invading pathogen. The outcome of this cross-talk between these two actors defines the incidence or avoidance of the disease. Host plant resistance to microbial diseases comprises two mechanisms. The first is basal defense, manifesting actions of the plant immune system, activated in response to elicitors or microbe-associated molecular patterns (MAMPS), such as fungal chitins, polysaccharides, elongation factors or heptaglucosides. These elicitors are the gene products of the attacking pathogen (Jones and Takemoto 2004). The second mechanism is based on the actions of the triggered adaptive immune system after the detection of the cytoplasmic effectors called nucleotide-binding and Leu-rich repeat (NBS-LRR) protein domains, coded by resistance (R) genes, and it represents R genes where they deploy their gene products against *Avr* genes as products of the pathogen. The basal defense of the host immune system can be activated upon recognition of the pathogen elicitor. The second part of the host immune system includes activation of the resistance (R) genes. These R proteins initiate host immune responses conferring a resistant phenotype on the plant. A series of interactions taking place between the host R genes and the pathogen-secreted *Avr* effectors basically subscribes to the

gene-for-gene theory (Flor 1974). Continuously co-evolving against pathogen effectors, host plants also have developed an innate adaptive immune system. On the other hand, the pathogen also undergoes adaptive selection to outsmart the host and induce successful infection (Fig. 14.2).

Historically, the race-specific resistance based on gene-for-gene relationship was initially identified in hexaploid wild species of *S. dimessum*. The identified genes exhibited a hyper-sensitive tissue response to all races of *P. infestans* that did not possess the corresponding virulence to the R gene. By the middle of the last century, 11 R genes from *S. dimessum* (Black et al. 1953) had been identified. Many of these genes were extensively used in commercial breeding programmes world-wide, including India, and resulted in development of an array of resistant potato varieties. However, deployed resistance was overcome in due course by the development of new complex matching virulence in *P. infestans* owing to its diversity, plasticity and adoptive ability. This led to a change in breeding strategies and potato breeders started looking for horizontal resistance conferred by a group of genes (including minor genes with adoptive affects). Potato breeders even resorted to the development of R gene-free populations throughout the second half of the last century. Of course, this strategy was found to be unsuccessful and late blight resistance remains a challenge to scientists across the globe (Fig. 14.3).

However, our understanding of the various mechanisms underlying the pathogen-host plant interaction is supported by the molecular and genomic investigations undertaken during the last two decades. A few mapping populations were created, quantitative trait loci (QTLs) for late blight resistance were identified, the genome of *S. Phureja* was sequenced. Consequently, nearly 21 R genes have been identified, conferring resistance to different races of *P. infestans*

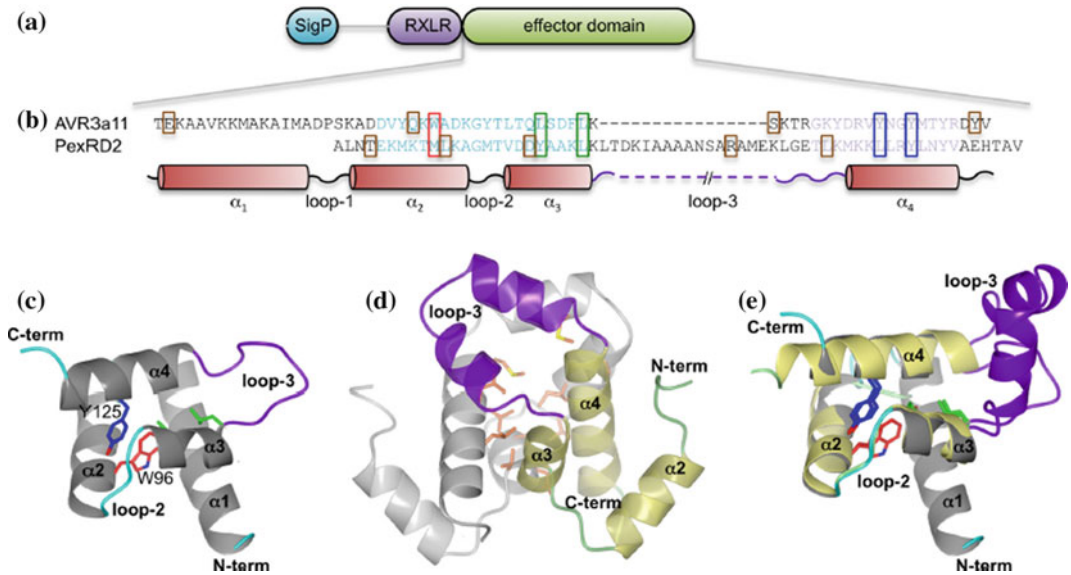

Fig. 14.2 RXLR effectors AVR3a11 and PexRD2 adopt a structurally conserved but adaptable fold. **a** Oomycete RXLR effectors are modular proteins comprising a secretion signal (*cyan*), RXLR translocation motif (*purple*), and an effector domain (*green*). **b** Structural alignment and secondary structure elements of the effector domains of AVR3a11 and PexRD2. W-motifs (*cyan*) and Y-motifs (*lilac*) are colored, with key residues (as discussed under "Results and Discussion") boxed. **c** The structure of AVR3a11 is a monomeric four-helix bundle with a hydrophobic core. Carbon atoms of key residues in the W- and Y-motifs are colored as boxed in **b**; loop-3 is shown in purple. **d** PexRD2 is a dimer with a hydrophobic interface, including residues Val[73], Asp[74], Ala[77], Thr[83], Ile[86], Ala[90], Met[96], Gly[100], Met[105], Leu[108], Leu[109], and Leu[112] (shown for one monomer only). α-Helices are labeled to correspond to equivalent positions in AVR3a11. **e** The AVR3a11/PexRD2-monomer overlay generated using SSM, showing the conserved fold. Protein structures are colored as in **c** and **d** with key residues of the W- and Y-motifs colored as in **b**. (Published with permission from Boutemy et al. 2011)

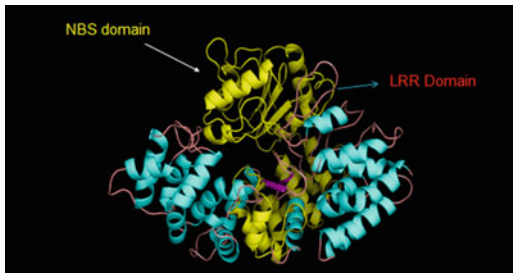

Fig. 14.3 3-D structure of R3 resistant protein from *S. tuberosum*, P-loops belonging to LRR are shown in *light brown*. Energy Values: 34962.695 kJ/mol (Modeler output) (unpublished, personal communication)

(Spooner and Bamberg 1994; Douches et al. 2001; Hijmans et al. 2003). Recently a new R gene was isolated simultaneously—RB (GenBank accession number AAP45164; Song et al. 2003) and *Rpi-blb1* (GenBank accession number AY42659; Van der Vossen et al. 2005) from *Solanum bulbocastanum* by employing map-based cloning and PCR techniques. The RB gene also comprises NBS-LRR functional domains typical of other R genes and was found to possess broader resistance against many races of *P. infestans* (Fig. 14.4). Potato breeders around the world are presently working on strategies to incorporate the RB gene into potato cultivars by using both conventional breeding as well as modern biotechnological approaches. Recently the RB gene was introduced into a popular Indian cultivar using transgenic methods, resulting in high levels of resistance against late blight (Shandil et al. 2017). Of late, the RB gene also has started becoming susceptible to late blight. This development again forced potato breeders and pathologists to explore other strategies. Recent sequencing of the genome of *P. infestans* (T-30; Haas et al. 2009) and the availability of functional data on various effector proteins (targeting especially the structure of regions known for virulence) assisted by in silico structural studies appears to be promising.

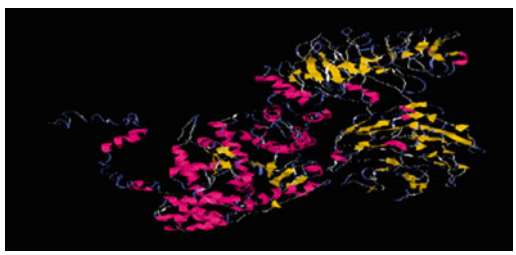

Fig. 14.4 3-D structure of, *Rpi-blb1,* R gene isolated from *S. bulbocastratum*. Functional domains coil-coil (shown in *pink*), NBS (*yellow*) and LRR (*blue*) are shown (Modeler output)

All the identified R genes belong to the NBS-LRR superfamily (Marone et al. 2013). They are one of the largest gene families in plants with more than 650 R genes mapped and sequenced so far. NBS-LRRs are known for their pathogen recognition functions inside the cell and are very adaptive in tune with the selective pathogen pressure. The nucleotide binding site along with other domains belongs to the NB-ARC complex. NBS-LRR proteins are classified into two types: TIR-NBS-LRR (TNL) that contain a Toll-like domain and CC-NBS-LRR (CNL) proteins. CNLs are characterized by their unique coiled-coil nature of the structure (see Fig. 14.4).

A major function of the LRR domain is the recognition of specific effectors secreted by *P. infestans*. The modeled structure of the R-proteins reveals new functions of LRR domains with identified P-loops, Kinase-2 and Gly-Leu-Pro-Leucin and other motifs that are associated with signaling during effector identification (Marone et al. 2013). LRR domains are structurally very dynamic and facilitate high levels of adaptation to suit the demands of the host immune system as part of co-evolution. Adoptive plasticity of LRR domains arises from their selective protein-protein interactions that yield different active site binding conformations. Predicted solvent-exposed residues of LRR domains provide space for variable binding sites, thus allowing variability and diversity to compete with evolving pathogen effectors (Figs. 14.5 and 14.6). Sub-cellular placement of NBS-LRR proteins is essential for resistance phenotypes and plays an important role in the identification of individual *P. infestans* and conferring resistance against them. NBS-LRR proteins are shown to be localized mostly in nucleus and cytoplasm (Shen 2007; Meyers et al. 2003). Apart from *P. infestans*, different alleles of

Fig. 14.5 3-D structure of *P. infestans* effector *Avr3a1*, Ribbon display of Modeler output, showing P-loop from LRR region four helix bundles

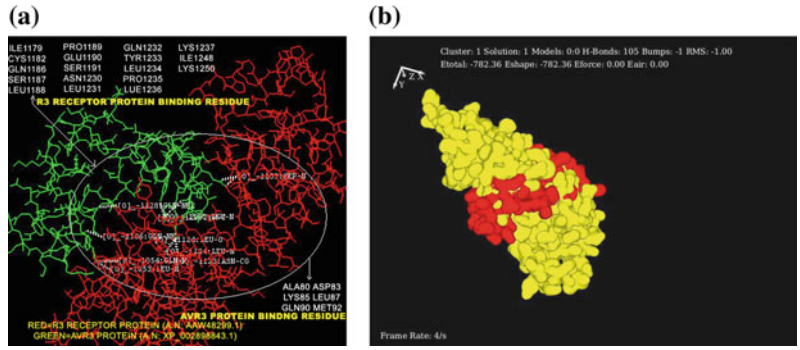

Fig. 14.6 Molecular docking of *P. infestans* effector *Avr3a* and R3 proteins **a** Mapped amino acids residues from binding site (shown inside the circle), **b** Diagram showing presence of RXLR motif in red color. HEX output, (Marla 2016) (Color figure online)

NBS-LRR provide resistance also to other pathogens such as viruses and nematodes.

Recent reports provide information about the underlying molecular mechanisms that activate NBS-LRR proteins upon pathogen invasion. It is interesting to know what molecular mechanism facilitates NBS-LRR host receptors to recognize an individual pathogen. Of late, data obtained from modeled three-dimensional structures of NBS-LRR proteins have given some clues and throw some light on this recognition mechanism. For example, effector *Avr3* corresponds to the popular R3a gene in potato. Of the two alleles of *Avr3*, AVR3a^{KI} only is reported to be recognized, and the other allele AVR3a^{EM} avoids recognition by the potato immune receptor. Modeled three-dimensional structures of both *Avr3a* and the NBS-LRR protein may potentially illustrate this, while docking involving both these allelic forms thus explains recognition or evasion. According to the proposed "jack-knife" model, the *Avr* effector disrupts the existing intra-molecular associations and liberates the CC, NB and or LRR domains, allowing them to interact with other proteins. An exchange of NBS-ARC domains between related other proteins and resulting in new conformational changes enables the freshly bound nucleotide to recognize the attacking *Avr* protein (Dodds et al. 2006). Nucleotide-dependent conformational changes and the subsequent oligomerization activate, forming a scaffold of signaling

components that finally trigger cell death (Mestre and Baulcombe 2006).

14.2 *P. infestans* Genome Organization and Composition

The genome of *P. infestans* is 240 Mb in size and was sequenced by the Broad Institute in the USA in 2012 (Haas et al. 2009). Nearly 18,000 genes were predicted from the assembled genome. Mapped predicted genes revealed interesting facts about their composition and organization. *P. infestans* contains a total of 560 RXLR and 180 CRN effector genes secreted (Haas et al. 2009). Most of the pathogenesis-related genes were found to be located in gene-sparse areas of the chromosomes. An Indian strain of *P. infestans* HP 1031 was sequenced and computationally analyzed in our lab at ICAR.NBPGR, New Delhi, employing NGS techniques. A few effectors, *AVR1*, *AVR2*, *AVR3a*, *AVR4*, IPI-O1 (AVR-blb1), AVR-blb2 and AVR-vnt1, were identified (Vleeshouwers et al. 2011). The majority of secreted effectors contain one RXLR motif, located on the N-terminal (Fig. 14.2), that is only activated upon recognition by the corresponding R gene from the host plant, and two W motifs and one Y motif on the C-terminal domain. The located RXLR motif on the N-terminal is only activated upon recognition by the corresponding R gene and the C-terminal motifs define the activity of the latter (Rehmany et al. 2005).

14.3 Plasticity and Adoptability in *P. Infestans* Effectors

MEME tool was employed to identify similar patterns in selected RXLR effectors in predicted *P. infestans* strain HP 103. Three conserved sequence motifs, W, Y and L, amino acids are arranged in a strict conserved pattern (Fig. 14.1). As reported, these repeat-rich, gene-sparse regions housing W, Y and L sequence patterns are supposed to promote evolutionary plasticity and variation in *Avr* genes, facilitating higher pathogenocity and host-specificity (Raffaels et al. 2010).

At present, a total of 11 *P. infestans* race-specific R genes have been mapped, isolated and cloned from wild potato species *S. demissum* that correspond to *P. infestans* secreted effectors.

14.4 Genomic Organization of NBS-LRR Genes in Potato

Genome sequencing of nuclear and organellar genomes of *S. tuberosum*, *Phureja* DM has reported on the prediction and mapping of 500 NBS candidate genes (Potato Genome Sequencing Consortium 2011). The sequenced DM genome (844 Mb) contains 408 NBS-LRR, 51 TIR and other non-TIR domains of R genes. NBS-LRR genes are not uniformly distributed across all the chromosomes in potato. A large number of them (nearly 15%) are mapped onto chromosomes 4 and 11 and a few are located (1%) on chromosome 3 (Jupe et al. 2012; Lozano et al. 2012). In the potato genome, NBS-LRR genes are organized either as isolated genes or in clusters, presumed to be the result of rapid R gene evolution against the challenged stress factors (Hulbert et al. 2001). In potato, 73% of mapped genes are found to be located in 63 clusters (Wan 2008). Compared to the TNL class, the CNL class of NBS-LRR genes are in the majority (85%) and mostly located in different clusters across the chromosomes.

14.5 R and *Avr* Pairs

14.5.1 R1-*Avr1*

R1 is the first R gene isolated from *S. demissum* using map-based cloning. The gene was isolated, and cloned using map-based cloning and mapped onto the short arm of chromosome V (Ballova et al. 2002). R1 gene codes a polypeptide of 1293 amino acids that belong to the CC-NBS-LRR class of plant R genes (Dongl and Jones 2001). The gene has three haplotypes. Although R1 was

one of the first R genes to be introduced into potato cultivars, it soon was overcome by *P. infestans*. The *Avr1* gene matches the R1 gene in *Solanum* and was isolated using map-based cloning and codes 208 amino acids (Guo 2008).

14.5.2 R2-*Avr2*

R2 gene (Genbank ac. JX313794) is a member of a wider family providing resistance to several strains of *P. infestans*. Eleven orthologues of R2 gene were isolated and mapped onto chromosome 4 from various wild species of potato, *S. demissum, S. bulbocastanum, S. bjertingii, S. edinense, S. scbenckii* and *S. microdonatum*. They include *R2, R2-like, Rpi-blb3, Rpi-abpt, Rpi-mcd1, Rpi-snk1.1, Rpi-snk 1.2, Rpi-edn1.1, Rpi-bjt1.1, Rpi-bjt1.2, Rpi-* and *Rpi-bjt1.3* (Hermsen 1973; Champouret 2010). Interestingly, most of these R2 gene homologues (R2GH) were isolated using effector screens. In no time, almost all the R2 genes homologues were defeated in the field by various strains of *P. infestans*.

14.5.3 R3A-*Avr3a*

R3 was isolated from *S. dimessum* and mapped onto the short arm of chromosome XI (Huang et al. 2004). Detailed fine mapping studies further revealed that the phenotypic resistance was due to two tightly linked R3 genes, *R3a* and *R3b*. R3 gene-coded protein is made up of 1283 amino acids and belongs to CC-NB-LRR class (Fig. 14.3). Other homologues of R3 were also isolated from *S. stoloniferum* (*Rpi-sto2*) using *Avr3* functional screens. However, R3 was also defeated by fast-evolving races of *P. infestans*.

Avr3a was isolated from *P. infestans* using potato genotypes carrying various R genes from *S. demissum* (Armstrong et al. 2005). *Avr3* encodes typical RXLR cytoplasmic effectors. Homologues of *Avr3a* were also identified in *P. sojae* and *P. capsici*. *P. infestans* populations have two alleles of *Avr3a* expressing two haplotypes, *Avr3a KI* and *Avr3a FM*, two secretary protein effectors with the ability to suppress cell death immune response reaction in potato host plants (Armstrong et al. 2005). These alleles identified in the *Avr3* locus differ in their amino acid positions 80 and 103 respectively. The crystal structure of *Avr3a* protein is available in the Protein databank (PDB# 2NAR; Fig. 14.5). *Avr3a* was found to suppress immunity by binding to host E3 ubiquitin ligase CMPG1 (Yaeno et al. 2011). It appears to be a promising way to unlock (engineer) this binding so as to achieve durable resistance against *P. infestans*.

Today the main strategy is targeted search for new R genes and integrate them to confer broad-spectrum resistance to late blight (Park et al. 2009; Song et al. 2003; Van der Vossen et al. 2005) from Mexican wild species *S. bulbocastanum* (2n = 24). The RB gene is reported to confer broad resistance against a majority of screened strains of *P. infestans*. The RB gene encodes a protein of 970 amino acids long and belongs to the CC-NBS-LRR class (Song et al. 2003). Incidentally RB also belongs to the same cluster on chromosome V along with R1 and other R genes. Figure 14.5 shows the presence of the four layered RXLR bundles and the beta sheet recognition receptor fold comprising beta sheets as shown in modeled structures of *Rpi-blb1* (Fig. 14.4).

The vast majority of effectors secreted by oomycetes belong to the RXLR class. Each individual effector is specific and matches a corresponding R gene in the host potato plant. They are characterized by the presence of RXLR motif, Arg-Xaa-Leu-Arg. So far, 563 RXLR effectors have been predicted from the sequenced genome of *P. infestans,* T-30-4 (Haas et al. 2009). By genome analysis of the *P. infestans* strain HP-1031 (ICAR.IIPR, Shimla), using the RXLR motif signature from T-304, we identified 380 RXLR effectors in our laboratory. Assembly of 42,000 contigs yielded a total of nearly 13,000 *Avr* genes. Annotation of predicted *Avr* genes in *P. infestans* strain HP 1031 revealed their functions.

RXLRs are a major class of *P. infestans* effectors that mediate delivery to the host cytoplasm. However, the molecular mechanism

underlying the entry of pathogen into the host cell is yet to be understood. Bioinformatics and structural analysis describing the mapping surface topology of amino acid residues and their changed conformations during interaction, and recognition of effector molecules by host receptors, and resulting in the induction of immune responses, may potentially enrich our understanding of late blight pathogenesis.

Crystal structures of five oomycete RXLR effector proteins are deposited in the RCSB Protein databank in the public domain. Among them we selected three templates. AVR3a4 (PDB# 2LC2), AVR3a11 (PDB# 3ZRG), PexRD2 from *Phytopthora infestans* (PDB# 3ZRG.1). All the reported structures comprise a three alpha-helix fold called the WY domain, containing the conserved Trp and Tyr residues which interact to form a stable hydrophobic core (Fig. 14.6). When checked against the target structures, only a few structures deposited in PDB showed < 20% sequence similarity. However, the modeled structures (using the found targets) showed that the WY fold is well conserved but in monomeric forms (in AVR3a4 and AVR3a11) as a four-helix model (Boutemy et al. 2011). It is reported that residues present on the C-terminal play vital roles in both avoidance by host receptor and also later in suppressing the immune responses of the latter (Boutemy et al. 2011). While analyzing the RXLR effector *AVR3* (# HP.1.468) identified from *P. infestans* strain HP 1031, the observed WY fold may provide a flexible, stable scaffold that supports existing polymorphisms, reflecting the structural diversity among different effectors. Plasticity and adoptability are gained chiefly by resorting to changes in core folds (amino acid replacements) in WY in the alpha helices, thus enabling avoidance of host recognition. The process involves adoption of new loop regions in the alpha helices, by virtue of acquired insertions or depletions or point mutations resulting from replacement of surface amino acids *viz*. tryptophan, leucin and argenin (Boutemy et al. 2011).

Another key question seeking answers is how the effector protein after entry into the host cytoplasm interacts with host cell targets at the molecular level. Structural studies, for example, direct recognition of fungal and oomycete effectors by NBS-LRR receptors can be investigated to answer some of these questions. In our study we mapped polymorphic residues on the surface of *P. infestans Avr3a* and depicted one site responsible for recognition of this effector by NBLRR proteins of R3. Polymorphic residues are mapped to the four helix bundles as expected. The mapped surface region in *Avr3a* is important for recognition by NBS-LRR proteins of R3a. A structural understanding of how NB-LRRs directly interact with effectors is important in the prediction of active binding sites.

The availability of sequenced information of the *P. infestans* genome allows elucidation of the functions of various avirulent molecules involved in late blight infection and the suppression of host plant immune activity. For example, a *P. infestans* T30-4 strain has been sequenced and assembled, however, of the suspected presence of ~17,000 coded effector genes, only a few hundred have been annotated, revealing their molecular and biological functions (http://www.broadinstitute.org/scientific-community/software). The information available on genome organization and structure helped isolate many avirulent factors involved in late blight infection. The availability of literature reporting various investigations on functional genomics, structures of modeled molecules and efficient computational ontology tools and pipelines facilitate further elucidation of functions of different effector genes involved in late blight pathogenesis. In our laboratory we analyzed the sequence raw data and identified 390 effectors. The identified effectors belong to both types—apoplastic and cytoplasmic—presumed to be involved in early cyst germination and entry and establishment of zoospores, cell wall degradation, host enzyme inhibition, and nacrosis.

In the present study, raw sequence data of *Phytophthora infestans* strain HP 1031, sequenced at the Central Potato Research Institute (CPRI), Shimla, was assembled and searched for potential effectors among the predicted coding genes. Sequenced genome analysis was done employing next-generation sequencing

tools: CLC Bio Work bench, Velvet and Augustus, yielded 13,262 coded proteins from the 42,000 assembled contigs. We predicted a significant number of RXLR class of effectors, i.e. 392 in *P. infestans* HP 1031 as against 563 RXLR effectors reported with *P. infestans* T-30-4. Of these, we selected three effectors (*AVR3*, *AVR3a* and *AVR2*) belonging to the RXLR class among the predicted effector sequence of *P. infestans* strain HP 1031 and conducted sequence, structure analysis. Similarly seven *Solanum* R proteins (R2, R3, R3a and SHO 10 from *S. demissum* and BSL.1 from *S. tuuberosum* and *Rpi-blb3* from *S. bulbocastodium*) were downloaded from NCBI and used for sequence and structure analysis. Effector protein domain analysis confirmed the presence of a conserved N-terminus that carries a secretary signal peptide and two repeating sequence motifs RXLR and dEER, combined with a highly divergent C-terminus, housing the vital elicitor domains. RXLR (Arf-Xaa-Leu-Arg) is found to be a strictly conserved domain, as reported earlier (Vleeshowers et al. 2008) in the analyzed *P. infestans* genome data. This domain is reported to be chiefly responsible for translocation inside host plant cells and primarily influences virulence. Further, sequence analysis of RXLR effectors revealed the presence of repeating sequence motifs "W", "Y" and "L" representing amino acids conserved at a certain position in the C-terminal region. Occurrence of these repeating sequence motifs in a characteristic fashion is supposed to contribute to the rapid evolution and expansion of the RXLR superfamily under selection pressure and aids in adaptation (Vleeshowers et al. 2008).

Protein 3D structures of both effector proteins and *Solanum* R proteins, *Avr3a* and R3a respectively, were modeled and refined, employing structural parameters. The modeled ligand (effector molecule and receptor Solanum R proteins) were docked using CDOCK and AutoDOCK and HEX software tools. The obtained elicitor and host receptor docking was verified with energy parameters to define the binding strength (the efficiency of virulence) and were computed using Schroedinger Package. The obtained docking results helped us to successfully establish a viable effector, Solanum R protein interactions occurring between pairing sets *AVR3* and R3a (S. *demissum*); *AVR3a* and R3 *S. tuberosum*); *AVR* and BSL.1 (*S.tuberosum*); and *AVR2* and RPiblb3 (*S. bulbocastodium*). We demonstrated the working of the developed effector structure-based profiling of *Solanum* R proteins for screening individual *P. infestans* strains (Fig. 14.7).

14.6 Identification Effector Functions

As described above, *P. infestans* secretes several extra-cellular effectors to infect and establish itself in host potato plant. We annotated the predicted effectors employing Blast2Go, InterProScan tools to deduce their molecular and biological functions. We identified several effectors responsible for such vital functions as plant cell wall-degrading enzyme functions,

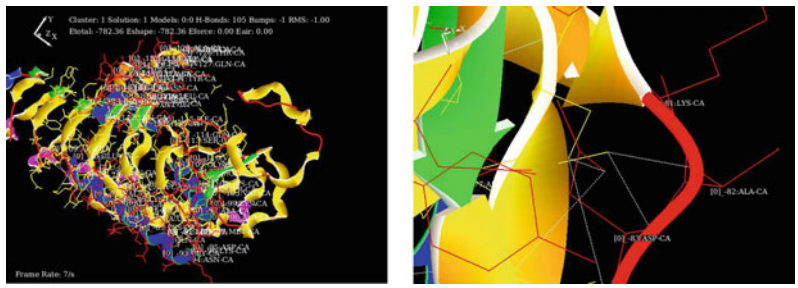

Fig. 14.7 *Avr3a* and R3 Docking with H bonds showing functional amino acids, suspected to be linked to SNPs to gain adaptation, HEX output

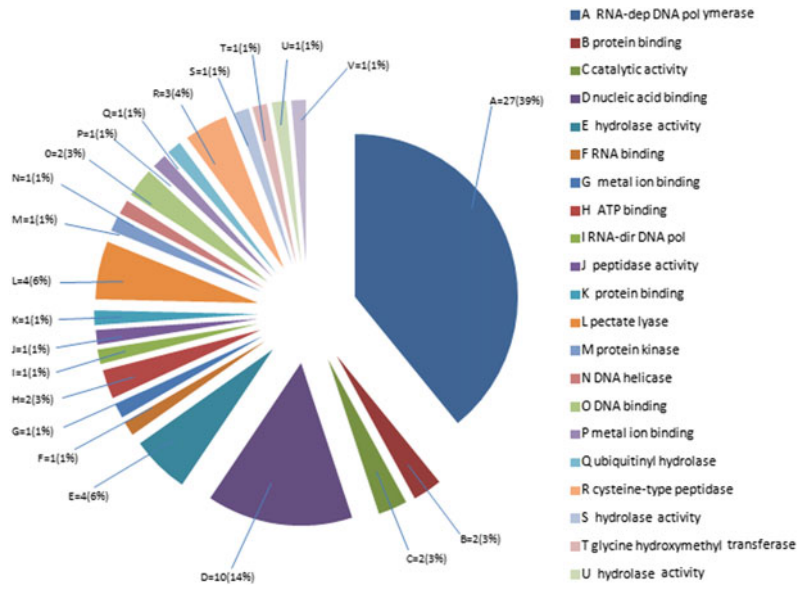

Fig. 14.8 Function annotation of sequenced genome of *P. infestans* strain HP 1031 (India), predicted protein domains and other functions

endoglucanases, PGIP, poly galacturanasesinose, endopolygalactinases, Pectate lysae (the polysaccharide group), cellulose binding genes; plant enzyme inhibitors belonging to Cystins, serines; the UIP1 protease family (involved in host plant tissue degeneration), sodium/potassium transport with ATPase signature; Myb-like chromatin binding proteins (FDNA); histone methylation (for cell cycle regulation), Crinkler family proteins (related to RXLRs for spore transmission), NEP1 orthologues for necrosis induction, protein kinase-related calcineurin-like phosphoesterase-1 (vital for plasticity and adoption), CRN protein family effectors (play a major role in host plant infection and manipulation of metabolic pathways) and several others (Fig. 14.8).

For example, the predicted CRN effectors, after transport into the host cytoplasm start manipulating its cellular processes, resulting in cell death, chlorosis and tissue browning (Boutemy et al. 2011). Effectors belonging to various families are annotated from sequenced *P. infestans* strain HP. A detailed biochemical study coupled with pathogenic verification of these effector functions may enrich our present understanding of late blight pathogenesis and the interplay between *P. infestans* and its host.

14.7 Databases and Other Computational Resources

Thanks to the untiring efforts of potato researchers around the world, a wealth of information is being deposited to disseminate the results to the potato research community. Some of the databases provide information relating to maps of all 12 potato chromosomes with mapped candidate R genes, mapped markers, sequence data, putative gene functions, SNP and Indel information from different diploid and tetraploid potato genotypes, and publication references. Apart from the widely available GenBank (http://www.ncbi.nlm.nih.gov/) and SGN (*Solanaceae* Genomics Network, http://www.sgn.cornell.edu/), many other resources are being reported in public domain repositories. Some of the notable resources include SolGene (https://solgenomics.net/), an online database to explore disease-resistant genes in *Solanum* species. Apart from providing data on R genes, QTLs, and molecular markers, SolRgene also has information on phylogeny, crossability among *Solanum* species, genotype-based resistance information (though only for a few important ones), genetic maps, allele mining, and above all facilitates

graphical visualization of these available resources. Other notable resources are SPUD (Hirsch et al. 2014) and PoMaMO (https://gabi.rzpd.de/PoMaMo.html; Meyer et al. 2005), a comprehensive database for potato genome data. Embedded in PoMaMO is YAMB, a genome browser that lets the user explore the genetic maps and visualize them. PFGD is integrated with the *Solanacae* Genomics database (SolGD; http://www.solgd.org) and provides a comparison between effector and their matching receptor molecules from the *Solanacae* species, enabling the exploration of host-pathogen interactions. The *Phytophthora* database (https://data.nal.usda.gov/dataset/phytophthora-database/resource/) supported by USDA. AFRI is a wonderful resource providing information on world collections of *Phytopthora* species, a referral global atlas. The Phytophthora Functional Genome Database PFGD (http://www.pfgd.org) is another information resource on *P. infestans* containing transcripts, genomics, gene expression, and functional assay data.

14.8 Effectors in R Gene Screens

A viable option to perform late blight management in potato, while incurring minimal losses, appears to be by building durable broad spectrum resistance pyramiding multiple R candidate genes as well as other genes with minor complementary effects. The sequenced genomic information and the associated gene lists with specific trait functions provide a logical platform for selection and comparison of effectors and their corresponding host resistance partners. SolGD is one such computational platform described above. Sequencing of a native strain HP of *P. infestans* enabled the identification of effectors for various functions involved in pathogenesis, such as host cell wall degradation, kinase signaling, potassium transport, and serine protease inhibition (see Fig. 14.8). Many of the predicted effectors are supposed to possess remarkable ability to manipulate biochemical, physiological and morphological processes in host plants. Use of effectors to screen *Solanum* germplasm to identify matching resistance molecules is a promising approach in late blight diagnostics. Apart from the use of effectors for generation of quantitative resistance data (Hunag et al. 2005), they are successfully employed in the identification of candidate R genes from wild *Solanum* species. R genes conferring resistance to *P. infestans* (*Rpi*) have been successfully isolated from *S. bulbocastanum* and *S. venture* (Van der Voosen et al. 2005; Foster et al. 2009).

14.9 Conclusion

Computational structural biology has provided key advances in our understanding of plant-pathogen interactions in recent years, there has already been available data generated to enrich our current understanding of identification of new and corresponding candidate R genes, and understand the molecular mechanisms underlying infection. Structural studies of modeled effector molecules provide information especially about effector diversity, how effector plasticity leads to diversity and aids to avoid host recognition. Some of the computational tools facilitate graphical visualization of virulence and many other aspects related to adoption, host recognition and other interesting aspects. Some of the structure-related studies may lead to identification of new genetic targets (pathogen recognition binding sites) for the efficient management of late blight, to help in the development of resistant potato varieties.

Acknowledgements Author greatfully acknowledges the support from ICAR-CPRI, Shimla, and the financial support from NAIP-ICAR for this project.

References

Armstrong MR, Whisson SC, Pritchard L, Bos JI, Venter E, Avrova AO, Rehmany AP et al (2005) An ancestral oomycete locus contains late blight avirulence gene *Avr3a*, encoding a protein that is

recognized in the host cytoplasm. Proc Nat Acad Sci USA 102:7766–7771

Ballvora A et al (2002) The *R1* gene for potato resistance to late blight (*Phytophthora infestans*) belongs to the leucine zipper/NBS/LRR class of plant resistance genes. Plant J 30:361–371

Black W, Mastenbroek C, Mills WR, Peterson LC (1953) A proposal for an international nomenclature of races of *Phytophthora infestans* and of genes controlling immunity in *Solanum demissum* derivatives. Euphytica 2:173–178

Boutemy LS, Stuart R, King F, Win J, Richard K, Hughes K, Thomas A, Clark MT, Blumenschein A, Kamoun S, Banfield MJ (2011) Structures of *Phytophthora* RXLR effector proteins: a conserved but adoptable fold underpins functional diversity. J Biol Chem 286:35834–35842

Champouret N (2010) Functional genomics of *Phytophthora infestans* effectors and *Solanum* resistance genes. Ph.D. Thesis University of Wageningen

Dodds PN, Lawrence GJ, Catanzariti AM, Teh T, Wang CI, Ayliffe MA, Kobe B, Ellis JG (2006) Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. Proc Natl Acad Sci USA 103:8888–8893

Dongl JL, Jones JD (2001) Plant pathogens and integrated defense responses to infection. Nature 411:826–833

Douches D, Bamberg J, Ir W, Jastrzebasi K, Neimira B (2001) Evaluation of wild *Solanum* species for resistance to the US-8 genotype of *Phytophtera infestans* utilizing a fine screening technique. Am J Potato Res 78:159–165

Flor FH (1974) Current status of the gene-for-gene concept. Annu Rev Phytopathol 78:275–298

Foster SJ, Park TH, Pel M, Brigneti G, Sliwka J, Jagger L, Van der Vossen EAG, Jones JDG (2009) *Rpi-Vnt1.1*, a *TM-2*^{2} a homolog from *Solanum venturii*, confers resistance to potato late blight. Mol Plant-Microbe Interact 22:601–615

Guo J (2008) *Phytophthora infestans* avirulence genes; mapping, cloning and diversity in field isolates. Ph.D. thesis, Wageningen University, The Netherlands, p. 138

Haas et al (2009) Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature 461:393–398

Hermsen JGTRM (1973) Double-bridge hybrids of *Solanum bulbocastanum* and cultivars of *Solanum tuberosum*. Euphytica 22:457–466

Hijmans RJ, Jacobs M, Bamberg JB, Spooner DM (2003) Frost resistance in wild potato species: assessing the predictability of taxonomic, geographic, ecological factors. Euphytica 130:47–59

Hirsch CD, Hamilton JP, Childs KL, Cepela J, Crisovan E, Vaillancourt B, Hirsch CN, Habermann M, Neal B, Buell CR (2014) Spud DB: a resource for mining sequences, genotypes, and phenotypes to accelerate potato breeding. Plant Genome 7(1):0

Huang S, Van der Vossen EAG, Kuang H, Vleeshouwers VGAA, Zhang N, Borm TJA, Van ECKHJ, Backer B, Jackbsen E, Visser RGF (2005) Comparative genomics enabled the isolation of *R3a* late blight resistance gene in potato. Plant J 45:251–265

Huang S, Vleeshouwers VG, Werij JS, Hutten RC, van Eck HJ, Visser RG, Jacobsen E (2004) The *R3* resistance to *Phytophthora infestans* in potato is conferred by two closely linked *R* genes with distinct specificities. Mol Plant-Microbe Interact 17:428–435

Hulbert SH, Webb CA, Smith SM, Sun Q (2001) Resistance gene complexes: evolution and utilization. Annu Rev Phytopathol 39:285–312

Jones DA, Takemoto D (2004) Plant innate immunity-direct and indirect recognition of general and specific pathogen-associated molecules. Curr Opin Immunol 16:48–62

Jupe F, Pritchard L, Ehrington GJ (2012) Identification and localization of the NBS-LRR gene family within the potato genome. BMC Genom 13:75

Lozano R, Ponce O, Ramirez M, Mostazo N, Orzeda G (2012) Genome wide identification and mapping of NBS encoding resistance genes in *Solanum tuberosum* group Phureja. PLoS ONE 7(4):e3475

Marone D, Russo MA, Laido A, Leonardis MD, Mastrangelo AM (2013) Plant nucleotide binding site–leucine-rich repeat (NBS-LRR) genes: active guardians in host defense responses. Inter J Mol Sci 14:7302–7326

Mestre P, Baulcombe DC (2006) Elicitor-mediated oligomerization of the tobacco n disease resistance protein. Plant Cell 18:491–501

Meyer M, Nagel A, Gebhardt C (2005) PoMaMo—a comprehensive database for potato genome data. Nucleic Acids Res 33:D666–D670

Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW (2003) Genome-wide analysis of NBS-LRR–encoding genes in Arabidopsis. Plant Cell 15:809–814

Park TH, Vleeshouwers VGAA, Jacobsen E, Visser RGF, Van der Vossen EAG (2009) Molecular breeding for resistance to *Phytophthora infestans* in potato. Plant Breed 128:109–117

Potato Genome Sequencing Consortium (2011) Genome sequence and analysis of the tuber crop potato. Nature 475:189–195

Raffaeles S, Win J, Cano I, Kamoun S (2010) Analysis of genome architecture and gene expression revel novel candidate virulence factors in the seccretome of *Phytophtera infestans*. BMC Genom 11:637

Rehmany AP, Gordon A, Rose LE, Allen RL, Armstrong MR, Whisson SC, Kamoun S, Tyler BM, Birch PRJ, Beynon JL (2005) Differential recognition of highly divergent downy mildew avirulence gene alleles by *RPP1* resistance genes from two Arabidopsis lines. Plant Cell 17:1839–1850

Shandil RK, Chakrabharti SK, Singh BP, Sharma S, Sunderasha S, Kaushik SK, Bhat KA, Sharma NN (2017) Genotypic background of the recipient plant is

crucial for conferring RB gene mediated late blight resistance in potato. BMC Genet 18:22

Shen QH (2007) Nuclear activiMLA immune receptors links isolate-specific and basal disease resistance responses. Science 315:1098–1103

Song J, Bradeen JM, Naes KS, Raasch JA, Weilgus SM, Haberlach T, Liu JL, Kuang H, Austin-Philips S, Buell RC, Helgeson JP, Jiang J (2003) Gene *RB* cloned from *Solanum bulbocastanum* confers broad spectrum resistance to potato late blight. Proc Natl Acad Sci USA 100:9128–9133

Spooner DM, Bamberg JB (1994) Potato genetic resources: sources of resistance and systematic. Am Potato J 71:325–327

Van der Vossen EA, Gros J, Sikkema A, Muskens M, Wouters D, Wolters P, Pereira A, Allefs S (2005) The *Rpi-blb2* gene from *Solanum bulbocastanumis* an *Mi-1* gene homolog conferring broad-spectrum late blight resistance in potato. Plant J 44:208–222

Vleeshouwers VGAA, Raffaels S, Vossen JH (2011) Understanding and exploiting late blight resistance in the age of effectors. Annu Rev Phytopathol 49:507–531

Vleeshouwers VGAA, Rietman H, Krenek P, Champouret N, Young C, Oh SK, Wang M, Bouwmeester K, Vosman B, Visser RGF et al (2008) Effector genomics accelerates discovery and functional profiling of potato disease resistance and *Phytophthora infestans* avirulence genes. PLoS ONE 3:e2875

Wan J (2008) A LysM receptor-lie kinase plays a critical role in chitin signaling and fungal resistance in Arabidopsis. Plant Cell 20:471–481

Yaeno T, Li H, Chapparo-Garcia A, Schornak S, Koshiba S, Watanabe S, Kigawa T, Kamoun S, Shirasu K (2011) Phosphatidylinositol monophosphate-binding interface in the oomycete RXLR effector *AVR3a* is required for its stability in host cells to modulate plant immunity. Proc Nat Acad Sci USA 108:14682–14687