

# Host-Pathogen Wars: New Weapons from Biotechnology and Genomics

Mychele B. da Silva<sup>1,2</sup>, Richard F. Davis<sup>3</sup>, Andrew H. Paterson<sup>4</sup>, Shavannor M. Smith<sup>1</sup>, Nelson Dias Suassuna<sup>5</sup>, Peng W. Chee<sup>2\*</sup>

<sup>1</sup>Department of Plant Pathology, University of Georgia, Tifton, GA, USA

<sup>2</sup>Institute of Plant Breeding, Genetics, and Genomics, University of Georgia, Tifton, GA, USA

<sup>3</sup>Crop Protection and Management Research Unit, USDA-ARS, Tifton, GA, USA

<sup>4</sup>Plant Genome Mapping Lab, University of Georgia, Athens, GA, USA

<sup>5</sup>EMBRAPA Algodão, Campina Grande, Brazil

Email: \*pwchee@uga.edu

**How to cite this paper:** da Silva, M.B., Davis, R.F., Paterson, A.H., Smith, S.M., Suassuna, N.D. and Chee, P.W. (2019) Host-Pathogen Wars: New Weapons from Biotechnology and Genomics. *American Journal of Plant Sciences*, 10, 402-416. <https://doi.org/10.4236/ajps.2019.103029>

**Received:** January 7, 2019

**Accepted:** March 8, 2019

**Published:** March 11, 2019

Copyright © 2019 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

Pathogens are imminent threats to crop production. Among the management tools available to protect crops from diseases, the use of host-plant resistance had been hindered by a lack of tools and resources to identify resistance genes (R-genes). Genomic technologies have empowered acquisition of a new level and quality of information on plant-pathogen interactions. Next generation sequencing, differential transcriptome analysis, gene editing, and use of bioinformatics have greatly expanded the numbers of R-genes identified, enriched understanding of R-avirulence gene interactions, and disease diagnosis. In this review, we highlight the application of genomic technologies to identification of pathogen machinery for future improvement of host plant resistance.

## Keywords

Host-Plant Resistance, R-Genes, Genomics, Crop Improvement

---

## 1. Introduction

Since the emergence of the first land plants around 700 million years ago [1], molecular interactions have occurred with microorganisms, symbiotic, epiphytic or pathogenic. Such interactions have shaped the evolution of the plant immune response, evolving a broadly primary mechanism to recognize common features of microbial pathogens and a more specific secondary mechanism based on resistance proteins (R proteins) to detect the presence of the pathogen effector proteins [2].

In agroecosystems, natural ecosystems that have been modified to produce food and fiber, cultivated plants are more exposed to sudden alterations imposed by weather and man to pest damage and catastrophic outbreaks owing to lack of diversity in species of plants and microorganisms. A plant disease epidemic develops when three factors occur simultaneously over a relatively long period of time: favorable environmental conditions, a virulent pathogen, and susceptible host plants [3]. Plant disease management practices try to modify one or more factors aiming to reduce initial inoculum (delay disease onset) or reduce disease progress curve in order to achieve economical control. Environment can be modified through crop management, as in a no-tillage system, use of crop rotation, adjustment of planting density, early planting or avoidance of crops susceptible to pathogens that predominate in specific locations. In some cases, when the pathogen is widespread, the only plausible options to slow disease progression are chemical control and the use of resistant genotypes.

The objectives of most plant breeding programs are to develop varieties that produce greater yields or better quality. However, there are several successful examples of breeding for disease resistance worldwide [4] [5] [6]. Plant genetic resistance can be defined as a set of mechanisms which interfere with and/or reduce the growth and/or development of economically important parasites [7], and disease resistance can be categorized as complete resistance or immunity conditioned by a single gene or incomplete resistance conditioned by multiple genes of partial effects [8], respectively named qualitative disease resistance and quantitative disease resistance. Breeding programs have identified and used R-genes widely, achieving complete resistance mediated by resistance (R) proteins. However, quantitative disease resistance, resistance that is expressed as a reduction in disease, rather than as the absence of disease, would contribute to the design and deployment of durably resistant crop cultivars [9].

Qualitative disease resistance is based on R-genes and occurs when there is an incompatible interaction between host plant and pathogen. In the evolutionary host-pathogen arm race, the pathogens impose a selective pressure and force the plant population to evolved post-invasion resistance mechanisms, often controlled by dominant R-genes, whose proteins detect specific pathogen effectors and trigger effective defense responses [10]. Resistance (R) and avirulence (avr) proteins may have direct recognition as stated in the gene-for-gene model proposed by Flor (1956) [11] or indirect interaction, as per the guard or decoy model [12], providing resistance to the plant. However, R-genes can be deflected by the pathogen and allowing the pathogen to successfully establish in the host.

To facilitate the study of R-genes and their interaction, it is necessary to know where the gene is located in the genome, when and how its expression is regulated, and the phenotypic reaction resulting from its expression. Genetic mapping to locate the chromosome location of a specific R-gene or the distribution of R-gene families across the genome has been greatly facilitated with improvement of genetic markers technology over the last three decades. The develop-

ment of genetic markers transitioned from plant morphology and proteins to the DNA level with the appearance of Restriction Fragment Length Polymorphisms or RFLPs [13]. Cleavage of specific sequences in the DNA using restriction enzymes produces thousands of markers that can be evaluated by segregation of the corresponding loci along the chromosomes of an organism [14]. The advent of the polymerase chain reaction (PCR) [15], simple-sequence repeat (SSR) [16] and Single Nucleotide Polymorphism (SNP) markers provide detection of subtle changes in DNA between closely related individuals, detecting even changes in single nucleotides throughout the genome. SNPs can be detected in expressed sequence tags (ESTs) or newly sequenced DNA [17] [18], with their genome-wide distribution potentially reducing marker bias [19].

The whole genome sequences of both crop plants and pathogens have increased the ability to study the regulation of R-genes at the whole genome level. While Sanger-based sequencing has produced most of the high-quality reference genomes to date, next generation sequencing (NGS) methods have dramatically improved the cost-effectiveness of whole genome sequencing and overtaken Sanger sequencing even to produce reference-quality genome sequences. Evolution and gene expression studies benefit from re-sequencing of ESTs or conserved regions in the genome [20] [21] [22]. As a result of genome sequencing, gene cloning becomes trivial, as does the identification of potentially functional R-genes that contain functional domains known to be diagnostic [23] [24] [25]. Further, gene editing offers the opportunity to reveal the effects of R-genes that cause interactions with the pathogen, perhaps disrupting virulence which drives a mutation to high frequency within a pathogen population [26].

Effector proteins have been a useful tool for the identification of candidate resistance genes as they act as molecular markers [25] [27]. Avr genes have corresponding R-genes that when activated in the plants can trigger immune responses. As such, effectors can be used to track their R-gene complements in many plant species and provide information about their functionality [25] [27]. Although there is an extensive database available of effector genes [28] [29], there will be many yet to be discovered as pathogens evolve quickly. Some class of pathogens have few distinct characteristics that are helpful to locate avr genes in the genome. For example, oomycete effectors are easily detected through the genome by their secretion signal peptide and RXLR motifs (Arg-X-Leu-Arg) required for translocation into the cell [23], while fungi have different signal sequences but no known specific signature sequence motifs to identify effectors.

Host resistance is an important component of crop disease management programs. Genetic markers have greatly advanced our understanding on the chromosome location and distribution of R-genes across plant genomes. However, the development of crops with host resistance is still hindered by the lack of genomic tools that can further dissect the mechanisms of host defense responses and how they are overcome by pathogens. In this review we discuss how available genomic technologies can advance progress in understanding host plant-pathogen interactions to better control and diagnose diseases.

## 2. Application of Genomic Tools to Improving Host-Plant Resistance

Genes conferring durable resistance to a pathogen are often difficult to find, deploy, and maintain. Crop plants frequently have extensive, although sometimes cryptic, genetic variation available in germplasm collections. Monoculture has arguably increased the rate at which new pathogens emerge [30] with associated selection pressure imposed on both the plant and pathogen potentially contributing to chromosome rearrangement of pathogens [31], horizontal gene transfer [32] [33], and recombination or mutation events. Selection pressure on R-genes is constantly present in the field and forces R-genes to evolve to defeat pathogens, while pathogens in turn experience selection pressures to defeat the plant immune system (arms race). Since the *Pto* gene in tomato [34], several R-genes have been cloned and most of them encode intracellular multi-domain proteins [35]. R proteins have a C-terminal leucine-rich repeat domain (LRR) linked to a central nucleotide-binding domain (NB), forming the NB-LRR receptor proteins which recognize effectors during pathogen infections. The nucleotide-binding fold in NB-LRR proteins is part of a larger domain, the NBS-apoptosis repressor with caspase recruitment (ARC) domain [36].

Genetic mapping has become a key element in studying plant disease resistance and identifying DNA markers diagnostic of disease resistance. First based on isozymes, the development of more abundant markers progressing from RFLPs to SSRs and eventually to SNPs has increased the efficiency for identifying diagnostic markers based on differences between genotypes of as little as a single base pair [37]. As a result, markers can be developed to narrow down a genomic region of interest, 'fine mapping' a trait to identify plausible candidate gene(s) [38]. Characteristic features of R-genes such as the NBS-LRR or other motifs provide clues by which they might be tentatively identified. Routine whole genome sequencing now accelerates searches for candidate R-genes likely to participate in salient biological pathways, utilizing similarities to known genes to deduce their possible functions in the cell, and providing DNA sequence from which to develop effector-assisted markers [21] [25].

### 1) Next generation sequencing applications

An important way to identify effector genes is through conserved gene regions of pathogens or fingerprints. For phyto bacteria, a typical effector sequence has *hrp* promoters, signal peptide, T3SS chaperone site, and skewed GC content in pathogenicity islands [39]. For fungi there is a signal peptide, transmembrane domain, and secretomes predicting effector location on the genome, however, sequence tags need to be introduced to confirm effector function since there are no universal translocational motifs for fungi as it is for oomycetes [40]. Although the RXLR translocational motif may also be found in fungi, it is the fingerprint for oomycete effectors [21] [25]. Until recent years, plant parasitic nematode genes had no known signature by which to predict effector proteins. The complete genome analysis of the yellow potato cyst nematode (*Globodera ros-*

*tochiensis*) revealed a promoter element motif called the DOG Box in most dorsal gland effectors that cause parasitism of this species [41]. The DOG BOX, as RXLR in oomycetes, is the beginning of discovering new effector motifs and possibly aid in genomic comparisons among plant parasitic nematodes to find new effector genes.

Due to the short life cycle of most pathogens, genes are usually fast evolving, giving rise to host specificity and pathogen speciation. A genome contains tens of thousands of genes that can mutate, recombine, or re-assort over time. Alterations in the genome can be identified, comparing individuals at species, genera, or family levels. For example, *Fusarium verticillioides*, *F. oxysporum* and *F. graminearum* have high genetic similarity. *F. oxysporum* share 91% homology of orthologous genes with *F. verticillioides* while both species share 85% homology with *F. graminearum* [33]. This finding indicates that *F. oxysporum* and *F. verticillioides* are most likely to have had a common ancestor more recently than either one did with *F. graminearum*. Additionally, re-sequencing data of *Lep-tosphaeria maculans* fungi, etiological agent of black leg disease of canola, showed a predicted effector and a previously cloned gene were in fact the same sequence with just three SNPs differentiating the virulent and avirulent alleles [20]. Sequencing and re-sequencing is important for, but not limited to, finding variations that cause changes in pathogenicity and tracking avr genes evolution from one species to another.

Building phylogenetic trees is a useful way to depict DNA sequence information in the study of pathogens and to find microorganism related with disease control. DNA sequences of closely related strains of *Bacillus amyloliquefaciens*, a biocontrol agent, were aligned and searched for associations with bacterial activity against pathogens. Strains of *B. amyloliquefaciens* contained many beneficial genes related to plant health such as genes that elicit plant basal defenses (*flgK*, *fliD*, and *Hag*) [42]. Genomic comparisons of pathogenic and non-pathogenic strains may provide clues toward new methods of disease control.

NGS is also routinely being used for RNA sequencing to provide a means to gain insight into the gene repertoire and expression patterns of an organism. The resulting sequences of gene fragments which are isolated en masse and reverse transcribed from messenger RNA are called expressed sequence tags (ESTs). Until recently, there are only a few complete genome sequences of plant parasitic nematodes, therefore much of our knowledge of their genomes derives from ESTs. Comparative genomics of ESTs may implicate specific transcripts in pathogenicity. For example, analysis of ESTs in *Bursaphelenchus xylophilus*, which attacks pine trees, revealed potentially unique pathogenicity genes compared with non-pathogenic species of *Bursaphelenchus* [43]. Comparing the number, timing, and level of expression of ESTs in pathogenic versus non-pathogenic species or biotypes can provide leads toward identifying pathogenicity genes or associated pathways and contribute to the study of gene evolution. Identification and functional characterization of effectors in *Globodera pallida*, a pathogen of potato,

based on analysis of more than 9242 ESTs revealed similarities with soybean pathogen *Heterodera glycines* secretion proteins, cellulose-binding protein, and cell signaling SPRYSEC with functional roles in early stages of nematode development [44]. These findings suggested that both pathogens are genetically closely related, and their effector genes may share common evolutionary pathways that lead to host specificity. Similarly, a comparison between non-pathogenic and pathogenic fungi revealed ESTs unique to pathogenicity [45]. The predicted protein functions of such transcripts might help in identifying the actual R-genes. These genetic differences between pathogenic and non-pathogenic organisms can further be screened for domains that are characteristic of R-genes such as NBS and ARC-2 domains [24].

Simple sequence repeats within ESTs typically vary between 2 - 6 nucleotides in the repeating unit(s). It is useful to design markers because the number of repeats can evolve quickly and create a great number of polymorphic loci [18]. The repeats can often be found close to a gene of interest, which can be efficiently utilized to track R-genes. However, ESTs databases are not a complete repertoire of expressed genes on the genome because the expression of R-genes can be dependent on environmental condition, what plant part the elicitor is being applied, and/or the time pathogen infection is occurring, making their detection difficult. In partial summary, ESTs remain of high value in the study of plant disease, as will be further elaborated in this article.

From another perspective, quarantine of pathogens is highly important to prevent introduction into areas that are pathogen free. While seed lots and nurseries are often screened for quarantined pathogens, small quantities of a pathogen in a large seed lot or nursery may pass undetected due to lack of sensitivity of diagnostic methods. A survey completed in Italy using 454 pyrosequencing together with meta-barcoding revealed *Phytophthora ramorum*, a quarantined pathogen for ornamental plants, in the soil of one of the nurseries [22]. This early detection has prevented the spread of *P. ramorum* to nurseries nearby, exemplifying how genomic techniques applied in plant pathology may prevent disastrous consequences. Although the identification of *P. ramorum* was possible, many pathogens have poor or no database available to rely on detection through sequencing. Oxford Nanopore technology is the most recent third generation DNA sequencing with ability to produce long reads length which facilitate genome assembly and analysis of structural variants [46]. It can generate 100 kilobase reads at a low cost. However, the use of this technology has not been fully explored in plant pathology. The MinIon is the mobile sequencer equipment that can quickly sequence a whole genome in about 24 hrs. The MinIon contains one cell and it was able to reconstruct *Escherichia coli* gene order without reference sequence or platforms [47]. The ability to sequence the genome at a fast pace will open doors to gather more pathogen information and detect diseases on site.

## 2) Differential gene expression

RNA sequencing and microarray transcriptome analysis are often used to

identify differentially expressed gene(s). Many expressed genes are up- or down-regulated in the infection-defense response. Both approaches can scan thousands of genes at the same time, with qRT-PCR targeting specific genes providing an important validation to remove analysis bias. The transcriptomes of two genotypes (e.g., resistant vs. susceptible) can be compared by quantifying gene expression. Detection of differential expression may help to identify genes involved in pathogenicity. In *Phytophthora sojae*, genes with effector signatures (RXLR) were up-regulated when cysts were germinating on soybean leaves. However, three hours after infection RXLR protein genes were down-regulated, as well as in zoospore stages [48]. Importantly, this information highlighted steps at which the pathogen may be prompted to infect, invade and reproduce *in planta*.

High throughput microarray methods have been developed to profile transcriptomes [49]. Although microarrays are an older technology than RNA sequencing (RNA-Seq), they are still widely used in plant science due to low cost. For instance, the life cycle of a pathogen may be related to its gene expression at different steps of infection. *Phytophthora infestans* gene expression profiling based on more than 18,000 unigenes gathered from previously identified ESTs, showed different expressed genes throughout the pathogen life stages with almost 90% of the sequence selected expressed at least once during the *P. infestans* life cycle. Groups of genes expressed during different stages of the pathogen life cycle may indicate genes involved in infection or disease development. Examples of such genes include catalase/oxidase and RXLR proteins, which are related to plant pathogenesis and flagellar components, which improve pathogen mobility and dispersion [21].

Although microarrays are useful for gene expression and quantification, they can miss important genes that are unknown in EST databases. Microarrays are unable to identify alternative splicing, may or may not distinguish gene isoforms, and are only a relative quantification method. In comparison, RNA-Seq provides a broader view of the transcriptome with coverage of coding and non-coding genes, besides being able to analyze just mRNA or poly-A RNA with elimination of rRNA. For example, comparative transcriptome profiling of two races of *Xanthomonas oryzae* pv. *oryzae* revealed some genes to be higher expressed in the virulent HB1009 (K3 race) strain than avirulent Xoo KACC10331 (K1 race), which were predicted to be involved in virulence or host-specificity [50].

The ability of RNA-Seq to recover small RNA is facilitating studies of gene silencing. Zheng *et al.* (2015) [51] were able to predict genes essential to *Heterodera avenae* survival using a siRNA system. Soaking *H. avenae* in a solution containing gene specific siRNA from *Caenorhabditis elegans* proved to be lethal to *H. avenae* larvae. In another study, RNA-Seq was utilized to examine changes in gene expression in *Globodera pallida*, a potato cyst nematode. High numbers of expressed genes were observed in juvenile stage 2 (J2) and male stage. Importantly, most of the predicted effectors in *G. pallida* had almost no similarity with root-knot nematodes which suggested different mechanisms of infection [31].

These studies permitted better understanding of gene expression throughout the pathogen life cycle and in host-pathogen interaction, identifying genes that could be important to plant defenses.

Despite RNA-Seq having gained popularity in determining transcriptome profile and quantifying gene expression levels, it is still a challenge to apply this method in organisms in which a reference genome is currently not available. However, as the costs of NGS continues to decrease, a reference quality genome sequence is expected to become available for most plant, animal, and microbial species by the next decade. The field of RNA-Seq based transcription profiling is progressing rapidly, and each new improvement in sequencing and data processing strategy brings us closer to a more complete understanding of the complex mechanisms involving host-plant defense responses and how they are overcome by pathogens.

### 3) Gene editing

The ability to inactivate or insert a gene sequence facilitates the work of plant pathologists in studies of devastating pathogens worldwide such as *Magnaporthe oryzae* that causes rice blast and accounts for more than 30% of world rice loss [52]. Techniques such as Zinc Finger, TALENs, and CRISPR-Cas9 are mainly used for this purpose. Both Zinc Fingers and TALENs (Transcription Activator-Like Effector Nucleases) are conceptually similar in that both systems use DNA binding motifs with customizable sequence specific nucleases to direct the same non-specific nuclease to cleave the genome at a specific site. CRISPR-Cas9 (Clustered, Regularly Interspaced, Short, Palindromic Repeats) system, on the other hand, uses a Cas endonuclease to cleave any target DNA sequence in the nucleus directed by a guide RNA (gRNA). Gene editing by any of the three systems may permit knock-out or inactivation of effector proteins that cause plant disease, providing insight into gene function. Understanding of gene function will lead to a better understanding of host-pathogen relationships, and perhaps lead to discovery of new treatments to the target problem. Zinc Fingers and TALENs constructs are more difficult to engineer than CRISPR-Cas9 because of how both techniques interact with the DNA target. By the Zinc Finger approach, each Zinc Finger interacts with three nucleotides in a specific manner, bonding to an endonuclease called FOKI. Because Zinc Finger proteins interact with three nucleotides at a time, it is more difficult to design the construct. On the other hand, TALEN protein domains interact with one nucleotide at a time but constructs also may be difficult to design due to specificity for each base pair. Moreover, the 5' end residue in the target affects DNA binding, decreasing the efficiency of the TALEN nucleases [53]. Comparatively, CRISPR-Cas9 uses bacterial cellular machinery to edit DNA, which is more cost effective and easier to make than Zinc Finger and TALENs. The CRISPR system was derived from a bacterial mechanism to defend against bacteriophage. It incorporates the virus sequence into the CRISPR locus and later transcribes small RNAs used to guide cas-9 endonuclease to cut DNA at a specific location in the genome. When the

bacteria detect the exogenous sequence in a subsequent attack by the same type of bacteriophage, the transcription machinery stored in the bacteria's CRISPR locus will create an instantaneous transcription response to neutralize the invading pathogen [54].

These gene editing technologies have opened new doors to the study and mitigation of plant diseases. TALENs conquered a change in rice susceptibility gene *Os11N3* that interacts with Avr7 or PhXo3 effectors, disabling the compatible interaction [26]. Similarly, CRISPR-cas9 was used in wheat and rice to knock-out mitogen-activated protein kinase (*OsMPK2*) genes that are related to plant defense response, and the phytoene desaturase (*OsPDS*) component of the abscisic acid production pathway. This research showed that mutation intensity was as high as 9% for the *OsPDS* gene, however, cleavage off-site may occur [55]. The newer gene editing technology is improving the prospects for producing plants engineered toward resistance and increased food production. Many resistant plants have been under selection pressure and pathogens may periodically overcome plant resistance. The ability to inactivate virulence genes or plant recognition sites by the pathogen, improvement of plant defenses, or gene stacking, may increase plant protection by decreasing selection pressure on the plant side.

#### 4) Bioinformatics

Bioinformatics is a complex field that combines biology with mathematic and computation science to understand biological functions and processes. Through computational analysis, bioinformaticians seek to improve sequence alignment, finding protein motif similarities, and determining pathways, predictions of effector genes, etc [25] [50] [56]. For each new sequencing platform developed, there is an improved software connected to it to analyze the sequence data more accurately [57]. Once a genome is sequenced, the genome assembly is annotated, and protein-based gene models are predicted. For example, gene predictions can be done by software such as FGENESH+ [17] or gene ontology can be done by the BLAST2GO program after RNA sequencing is performed [51]. Different bioinformatics programs were used to assemble the whole *Globodera* genome and predict effectors and their location, splicing sites, and identify horizontal gene transfer candidate genes [41].

Homology with known genes and proteins facilitates manual inspection for functional genes and introns, using gene ontology classifications to collect and analyze genetic data [56]. Many DNA sequences in international databases can be employed to search for the best match to new R or effector gene sequences, pathogen deployments and speciation, and genome and gene evolution. Even small changes in DNA sequence can distinguish one strain from another, making the use of SNP markers important for distinguishing populations with conserved region that might be essential to plant resistance. By aligning DNA sequences from resistant and susceptible wheat plants, a SNP marker was found and assessed for putative linkage to the leaf rust resistance *Lr67* gene, not found in susceptible plants [58]. Moreover, RFLP sequences of necrotrophic pathogens

from the *Cochliobolus* genus were anchored and many scaffolds were created to search for SNPs and pinpoint virulence genes [59]. In another example, by analyzing the DNA sequence of different isolates of *Fusarium oxysporum* f. sp. *vasinfectum*, pathogenic and non-pathogenic to cotton, Wang *et al.*, (2010) [60] were able to determine that the virulent strain of *Fusarium oxysporum* f. sp. *vasinfectum* found in Australia was evolved from a wild ancestor indigenous to Australia that is mildly pathogenic to cotton. Bioinformatics have infinite uses and is a necessary key tool in the study of plant-pathogen interaction and disease control.

### 3. Final Considerations

Advances in biotechnology and genomics have improved the potential for research discoveries in host-plant interactions. The emergence of new genomic tools such as the draft genome sequence of crop plants and their pathogens, and the new gene editing tools such as TALENs and CRISPR-cas9 are offering additional avenues for researchers to explore new approaches to discovering and incorporating host-plant resistance for plant improvement. The rapid declines in cost of new technology, together with continually expanding capabilities, promise a rich yield of new discoveries salient to mitigation of plant disease. The evolution of DNA technology has greatly benefitted the studies of new and alternative identification methods of pathogens. Faster pathogen diagnosis methods can help both farmers and researchers and will lead to improvement of disease management strategies. A better support in research and development to further understand pathogen biology needs to be addressed.

### Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

### References

- [1] Heckman, D.S., Geiser, D.M., Eidell, B.R., Stauffer, R.L., Kardos, N.L. and Hedges, S.B. (2001) Molecular Evidence for the Early Colonization of Land by Fungi and Plants. *Science*, **293**, 1129-1133. <https://doi.org/10.1126/science.1061457>
- [2] Chisholm, S.T., Coaker, G., Day, B. and Staskawicz, B.J. (2006) Host-Microbe Interactions: Shaping the Evolution of the Plant Immune Response. *Cell*, **124**, 803-814. <https://doi.org/10.1016/j.cell.2006.02.008>
- [3] Agrios, G.N. (2005) Plant Pathology. 5th Edition, Elsevier Academic Press, Amsterdam, 922 p.
- [4] Bird, L.S. (1982) The MAR (Multi-Adversity Resistance) System for Genetic Improvement of Cotton. *Plant Disease*, **66**, 172-176. <https://doi.org/10.1094/PD-66-172>
- [5] Bockus, W.W., Appel, J.A., Bowden, R.L., Fritz, A.K., Gill, B.S., Martin, T.J., Sears, R.G., Seifers, D.L., Brown-Guedira, G.L. and Eversmeyer, M.G. (2001) Success Stories: Breeding for Wheat Disease Resistance in Kansas. *Plant Disease*, **85**, 453-461. <https://doi.org/10.1094/PDIS.2001.85.5.453>

- [6] Saka, N. (2006) A rice (*Oryza sativa* L.) Breeding for Field Resistance to Blast Disease (*Pyricularia oryzae*) in Mountainous Region Agricultural Research Institute, Aichi Agricultural Research Center of Japan. *Plant Production Science*, **9**, 3-9. <https://doi.org/10.1626/pp.s.9.3>
- [7] Vale, F.X.R., Parlevliet, J.E. and Zambolim, L. (2001) Concepts in Plant Disease Resistance. *Fitopatologia Brasileira*, **26**, 577-589. <https://doi.org/10.1590/S0100-41582001000300001>
- [8] van der Plank, J.E. (1968) Disease Resistance in Plants. Academic Press, New York/London, 206 p.
- [9] Poland, J.A., Balint-Kurti, P.J., Wisser, R.J., Pratt, R.C. and Nelson, R.J. (2009) Shades of Gray: The World of Quantitative Disease Resistance. *Trends in Plant Science*, **14**, 21-29. <https://doi.org/10.1016/j.tplants.2008.10.006>
- [10] Jones, J.D.G. and Dangl, J.L. (2006) The Plant Immune System. *Nature*, **444**, 323-329. <https://doi.org/10.1038/nature05286>
- [11] Flor, H. (1956) The Complementary Genic Systems in Flax and Flax Rust. *Advances in Genetics*, **8**, 29-54. [https://doi.org/10.1016/S0065-2660\(08\)60498-8](https://doi.org/10.1016/S0065-2660(08)60498-8)
- [12] Van Der Hoorn, R.A. and Kamoun, S. (2008) From Guard to Decoy: A New Model for Perception of Plant Pathogen effectors. *The Plant Cell*, **20**, 2009-2017. <https://doi.org/10.1105/tpc.108.060194>
- [13] Botstein, R., White, R.L., Skolnick, M. and Davis, R.W. (1980) Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms. *American Journal of Human Genetics*, **32**, 314-331.
- [14] Tanksley, S.D., Young, N.D., Paterson, A.H. and Bonierbale, M.W. (1989) RFLP Mapping in Plant Breeding: New Tools for an Old Science. *Bio/Technology*, **7**, 257-264.
- [15] Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. (1986) Specific Enzymatic Amplification of DNA *in Vitro*: The Polymerase Chain Reaction. Cold Spring Harbor Symposia on Quantitative Biology. Cold Spring Harbor Laboratory Press, 263-273. <https://doi.org/10.1101/SQB.1986.051.01.032>
- [16] Tautz, D. (1989) Hypervariability of Simple Sequences as a General Source for Polymorphic DNA Markers. *Nucleic Acids Research*, **17**, 6463-6471. <https://doi.org/10.1093/nar/17.16.6463>
- [17] Lamour, K.H., Mudge, J., Gobena, D., Hurtado-Gonzales, O.P., Schmutz, J., Kuo, A., Miller, N.A., Rice, B.J., Raffaele, S. and Cano, L.M. (2012) Genome Sequencing and Mapping Reveal Loss of Heterozygosity as a Mechanism for Rapid Adaptation in the Vegetable Pathogen *Phytophthora capsici*. *Molecular Plant-Microbe Interactions*, **25**, 1350-1360. <https://doi.org/10.1094/MPMI-02-12-0028-R>
- [18] Lopez, C.E., Quesada-Ocampo, L.M., Bohorquez, A., Duque, M.C., Vargas, J., Tohme, J. and Verdier, V. (2007) Mapping EST-Derived SSRs and ESTs Involved in Resistance to Bacterial Blight in *Manihot esculenta*. *Genome*, **50**, 1078-1088. <https://doi.org/10.1139/G07-087>
- [19] Schlötterer, C. (2004) The Evolution of Molecular Markers—Just a Matter of Fashion? *Nature Reviews Genetics*, **5**, 63-69. <https://doi.org/10.1038/nrg1249>
- [20] Ghanbarnia, K., Fudal, I., Larkan, N.J., Links, M.G., Balesdent, M.H., Profotova, B., Fernando, W., Rouxel, T. and Borhan, M.H. (2015) Rapid Identification of the *Lep-tosphaeria maculans* Avirulence Gene AvrLm2 Using an Intraspecific Comparative Genomics Approach. *Molecular Plant Pathology*, **16**, 699-709. <https://doi.org/10.1111/mpp.12228>
- [21] Judelson, H.S., Ah-Fong, A.M., Aux, G., Avrova, A.O., Bruce, C., Cakir, C., Da

- Cunha, L., Grenville-Briggs, L., Latijnhouwers, M. and Ligterink, W. (2008) Gene Expression Profiling during Asexual Development of the Late Blight Pathogen *Phytophthora infestans* Reveals a Highly Dynamic Transcriptome. *Molecular Plant-Microbe Interactions*, **21**, 433-447. <https://doi.org/10.1094/MPMI-21-4-0433>
- [22] Prigigallo, M.I., Abdelfattah, A., Cacciola, S.O., Faedda, R., Sanzani, S.M., Cooke, D.E. and Schena, L. (2016) Metabarcoding Analysis of *Phytophthora* Diversity Using Genus-Specific Primers and 454 Pyrosequencing. *Phytopathology*, **106**, 305-313. <https://doi.org/10.1094/PHYTO-07-15-0167-R>
- [23] Dou, D., Kale, S.D., Wang, X., Chen, Y., Wang, Q., Wang, X., Jiang, R.H.Y., Arredondo, F.D., Anderson, R.G., Thakur, P.B., Mcdowell, J.M., Wang, Y. and Tyler, B.M. (2008) Conserved C-Terminal Motifs Required for Avirulence and Suppression of Cell Death by *Phytophthora sojae* Effector Avr1b. *The Plant Cell*, **20**, 1118-1133. <https://doi.org/10.1105/tpc.107.057067>
- [24] Liu, X., Lin, F., Wang, L. and Pan, Q. (2007) The *in silico* Map-Based Cloning of P36, a Rice Coiled-Coil-Nucleotide-Binding Site-Leucine-Rich Repeat Gene That Confers Race-Specific Resistance to the Blast Fungus. *Genetics*, **176**, 2541-2549. <https://doi.org/10.1534/genetics.107.075465>
- [25] Vleeshouwers, V.G., Rietman, H., Krenek, P., Champouret, N., Young, C., Oh, S.-K., Wang, M., Bouwmeester, K., Vosman, B. and Visser, R.G. (2008) Effector Genomics Accelerates Discovery and Functional Profiling of Potato Disease Resistance and *Phytophthora infestans* Avirulence Genes. *PLoS ONE*, **3**, e2875. <https://doi.org/10.1371/journal.pone.0002875>
- [26] Li, T., Liu, B., Spalding, M.H., Weeks, D.P. and Yang, B. (2012) High-Efficiency TALEN-Based Gene Editing Produces Disease-Resistant Rice. *Nature Biotechnology*, **30**, 390-392. <https://doi.org/10.1038/nbt.2199>
- [27] Vleeshouwers, V.G., Finkers, R., Budding, D., Visser, M., Jacobs, M.M., Van Berloo, R., PeL, M., Champouret, N., Bakker, E. and Krenek, P. (2011) Sol R Gene: An Online Database to Explore Disease Resistance Genes in Tuber-Bearing Solanum Species. *BMC Plant Biology*, **11**, 1. <https://doi.org/10.1186/1471-2229-11-116>
- [28] NCBI (2017) Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, **44**, D7-D19.
- [29] Wang, Y., Huang, H., Sun, M.A., Zhang, Q. and Guo, D. (2012) T3DB: An Integrated Database for Bacterial Type III Secretion System. *BMC Bioinformatics*, **13**, 66. <https://doi.org/10.1186/1471-2105-13-66>
- [30] Mcdonald, B.A. and Linde, C. (2002) Pathogen Population Genetics, Evolutionary Potential, and Durable Resistance. *Annual Review of Phytopathology*, **40**, 349-379. <https://doi.org/10.1146/annurev.phyto.40.120501.101443>
- [31] Cotton, J.A., Lilley, C.J., Jones, L.M., Kikuchi, T., Reid, A.J., Thorpe, P., Tsai, I.J., Beasley, H., Blok, V. and Cock, P.J. (2014) The Genome and Life-Stage Specific Transcriptomes of *Globodera pallida* Elucidate Key Aspects of Plant Parasitism by a Cyst Nematode. *Genome Biology*, **15**, 1-17. <https://doi.org/10.1186/gb-2014-15-3-r43>
- [32] Haegeman, A., Jones, J.T. and Danchin, E.G. (2011) Horizontal Gene Transfer in Nematodes: A Catalyst for Plant Parasitism? *Molecular Plant-Microbe Interactions*, **24**, 879-887. <https://doi.org/10.1094/MPMI-03-11-0055>
- [33] Ma, L.-J., Van Der Does, H.C., Borkovich, K.A., Coleman, J.J., Daboussi, M.-J., Di Pietro, A., Dufresne, M., Freitag, M., Grabherr, M. and Henrissat, B. (2010) Comparative Genomics Reveals Mobile Pathogenicity Chromosomes in *Fusarium*. *Nature*, **464**, 367-373. <https://doi.org/10.1038/nature08850>

- [34] Martin, G.B., Brommonschenkel, S.H., Chunwongse, J., Frary, A., Ganai, M.W., Spivey, R., Wu, T., Earle, E.D. and Tanksley, S.D. (1993) Map-Based Cloning of a Protein Kinase Gene Conferring Disease Resistance in Tomato. *Science*, **262**, 1432-1436. <https://doi.org/10.1126/science.7902614>
- [35] Martin, G.B., Bogdanove, A.J. and Sessa, G. (2003) Understanding the Functions of Plant Disease Resistance Proteins. *Annual Review of Plant Biology*, **54**, 23-61. <https://doi.org/10.1146/annurev.arplant.54.031902.135035>
- [36] van der Biezen, E.A. and Jones, J.D.G. (1998) The NB-ARC Domain: A Novel Signalling Motif Shared by Plant Resistance Gene Products and Regulators of Cell Death in Animals. *Current Biology*, **8**, R226-R227. [https://doi.org/10.1016/S0960-9822\(98\)70145-9](https://doi.org/10.1016/S0960-9822(98)70145-9)
- [37] Varshney, R.K., Chabane, K., Hendre, P.S., Aggarwal, R.K. and Graner, A. (2007) Comparative Assessment of EST-SSR, EST-SNP and AFLP Markers for Evaluation of Genetic Diversity and Conservation of Genetic Resources Using Wild, Cultivated and Elite Barleys. *Plant Science*, **173**, 638-649. <https://doi.org/10.1016/j.plantsci.2007.08.010>
- [38] Jander, G., Norris, S.R., Rounsley, S.D., Bush, D.F., Levin, I.M. and Last, R.L. (2002) *Arabidopsis* Map-Based Cloning in the Post-Genome Era. *Plant Physiology*, **129**, 440-450. <https://doi.org/10.1104/pp.003533>
- [39] Petnicki-Ocwieja, T., Schneider, D.J., Tam, V.C., Chancey, S.T., Shan, L., Jamir, Y., Schechter, L.M., Janes, M.D., Buell, C.R., Tang, X., Collmer, A. and Alfano, J.R. (2002) Genomewide Identification of Proteins Secreted by the Hrp Type III Protein Secretion System of *Pseudomonas syringae* pv. *tomato* DC3000. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 7652-7657. <https://doi.org/10.1073/pnas.112183899>
- [40] Sonah, H., Deshmukh, R.K. and Bélanger, R.R. (2016) Computational Prediction of Effector Proteins in Fungi: Opportunities and Challenges. *Frontiers in Plant Science*, **7**, 126. <https://doi.org/10.3389/fpls.2016.00126>
- [41] Eves-Van Den Akker, S., Laetsch, D.R., Thorpe, P., Lilley, C.J., Danchin, E.G., Da Rocha, M., Rancurel, C., Holroyd, N.E., Cotton, J.A. and Szitenberg, A. (2016) The Genome of the Yellow Potato Cyst Nematode, *Globodera rostochiensis*, Reveals Insights into the Basis of Parasitism and Virulence. *Genome Biology*, **17**, 1-23. <https://doi.org/10.1186/s13059-016-0985-1>
- [42] Magno-Pérez-Bryan, M., Martínez-García, P., Hierrezuelo, J., Rodríguez-Palenzuela, P., Arrebola, E., Ramos, C., De Vicente, A., Pérez-García, A. and Romero, D. (2015) Comparative Genomics within the *Bacillus* Genus Reveal the Singularities of Two Robust *Bacillus amyloliquefaciens* Biocontrol Strains. *Molecular Plant-Microbe Interactions*, **28**, 1102-1116. <https://doi.org/10.1094/MPMI-02-15-0023-R>
- [43] Kikuchi, T., Aikawa, T., Kosaka, H., Pritchard, L., Ogura, N. and Jones, J.T. (2007) Expressed Sequence Tag (EST) Analysis of the Pine Wood Nematode *Bursaphelenchus xylophilus* and *B. mucronatus*. *Molecular and Biochemical Parasitology*, **155**, 9-17. <https://doi.org/10.1016/j.molbiopara.2007.05.002>
- [44] Jones, J.T., Kumar, A., Pylypenko, L.A., Thirugnanasambandam, A., Castelli, L., Chapman, S., Cock, P.J., Grenier, E., Lilley, C.J. and Phillips, M.S. (2009) Identification and Functional Characterization of Effectors in Expressed Sequence Tags from Various Life Cycle Stages of the Potato Cyst Nematode *Globodera pallida*. *Molecular Plant Pathology*, **10**, 815-828. <https://doi.org/10.1111/j.1364-3703.2009.00585.x>
- [45] Soanes, D.M. and Talbot, N.J. (2006) Comparative Genomic Analysis of Phytopathogenic Fungi Using Expressed Sequence Tag (EST) Collections. *Molecular Plant Pathology*, **7**, 61-70. <https://doi.org/10.1111/j.1364-3703.2005.00317.x>

- [46] Deamer, D., Akeson, M. and Branton, D. (2016) Three Decades of Nanopore Sequencing. *Nature Biotechnology*, **34**, 518. <https://doi.org/10.1038/nbt.3423>
- [47] Loman, N.J., Quick, J. and Simpson, J.T. (2015) A Complete Bacterial Genome Assembled *de novo* Using Only Nanopore Sequencing Data. *Nature Methods*, **12**, 733. <https://doi.org/10.1038/nmeth.3444>
- [48] Ye, W., Wang, X., Tao, K., Lu, Y., Dai, T., Dong, S., Dou, D., Gijzen, M. and Wang, Y. (2011) Digital Gene Expression Profiling of the *Phytophthora sojae* Transcriptome. *Molecular Plant-Microbe Interactions*, **24**, 1530-1539. <https://doi.org/10.1094/MPMI-05-11-0106>
- [49] Tachibana, C. (2015) Transcriptomics Today: Microarrays, RNA-seq, and More. *Science*, **349**, 544-546. <https://doi.org/10.1126/science.349.6247.544>
- [50] Noh, T.-H., Song, E.-S., Kim, H.-I., Kang, M.-H. and Park, Y.-J. (2016) Transcriptome-Based Identification of Differently Expressed Genes from *Xanthomonas oryzae* pv. *oryzae* Strains Exhibiting Different Virulence in Rice Varieties. *International Journal of Molecular Sciences*, **17**, 259. <https://doi.org/10.3390/ijms17020259>
- [51] Zheng, M., Long, H., Zhao, Y., Li, L., Xu, D., Zhang, H., Liu, F., Deng, G., Pan, Z. and Yu, M. (2015) RNA-Seq Based Identification of Candidate Parasitism Genes of Cereal Cyst Nematode (*Heterodera avenae*) during Incompatible Infection to *Aegilops variabilis*. *PLoS ONE*, **10**, e0141095. <https://doi.org/10.1371/journal.pone.0141095>
- [52] Talbot, N.J. (2003) On the Trail of a Cereal Killer: Exploring the Biology of *Magnaporthe grisea*. *Annual Reviews in Microbiology*, **57**, 177-202. <https://doi.org/10.1146/annurev.micro.57.030502.090957>
- [53] Nemudryi, A., Valetdinova, K., Medvedev, S. and Zakian, S. (2014) TALEN and CRISPR/Cas Genome Editing Systems: Tools of Discovery. *Acta Naturae*, **6**, 22.
- [54] Doudna, J.A. and Charpentier, E. (2014) The New Frontier of Genome Engineering with CRISPR-Cas9. *Science*, **346**, Article ID: 1258096. <https://doi.org/10.1126/science.1258096>
- [55] Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z. and Gao, C. (2013) Targeted Genome Modification of Crop Plants Using a CRISPR-Cas System. *Nature Biotechnology*, **31**, 686. <https://doi.org/10.1038/nbt.2650>
- [56] Alemu, K. (2015) The Role and Application of Bioinformatics in Plant Disease Management. *Advances in Life Science and Technology*, **28**, 28-33.
- [57] Deschamps, S. and Campbell, M.A. (2010) Utilization of Next-Generation Sequencing Platforms in Plant Genomics and Genetic Variant Discovery. *Molecular Breeding*, **25**, 553-570. <https://doi.org/10.1007/s11032-009-9357-9>
- [58] Forrest, K., Pujol, V., Bulli, P., Pumphrey, M., Wellings, C., Herrera-Foessel, S., Huerta-Espino, J., Singh, R., Lagudah, E., Hayden, M. and Spielmeier, W. (2014) Development of a SNP Marker Assay for the Lr67 Gene of Wheat Using a Genotyping by Sequencing Approach. *Molecular Breeding*, **34**, 2109-2118. <https://doi.org/10.1007/s11032-014-0166-4>
- [59] Condon, B.J., Leng, Y., Wu, D., Bushley, K.E., Ohm, R.A., Ollilar, R., Martin, J., Schackwitz, W., Grimwood, J., Mohdzainudin, N., Xue, C., Wang, R., Manning, V.A., Dhillon, B., Tu, Z.J., Steffenson, B.J., Salamov, A., Sun, H., Lowry, S., Labutti, K., Han, J., Copeland, A., Lindquist, E., Barry, K., Schmutz, J., Baker, S.E., Ciuffetti, L.M., Grigoriev, I.V., Zhong, S. and Turgeon, B.G. (2013) Comparative Genome Structure, Secondary Metabolite, and Effector Coding Capacity across *Cochliobolus* Pathogens. *PLoS Genetics*, **9**, e1003233. <https://doi.org/10.1371/journal.pgen.1003233>

- [60] Wang, B., Brubaker, C.L., Summerell, B.A., Thrall, P.H. and Burdon, J.J. (2010) Local Origin of Two Vegetative Compatibility Groups of *Fusarium oxysporum* f. sp. *vasinfectum* in Australia. *Evolutionary Applications*, **3**, 505-524.  
<https://doi.org/10.1111/j.1752-4571.2010.00139.x>

### Abbreviation Note List

ARC	Apoptosis Repressor with Caspase Recruitment Domain
Avr	Avirulence
BLAST2GO	Basic Local Alignment Search Tool Gene Ontology
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
ESTs	Expressed Sequence Tags
LRR	Leucine-Rich Repeat
NB	Nucleotide-Binding
NGS	Next Generation Sequencing
PCR	Polymerase Chain Reaction
qRT-PCR	Quantitative Reverse Transcription Polymerase Chain Reaction
R	Resistance
RFLPs	Restriction Fragment Length Polymorphisms
RXLR motifs	Argenine-X-Leucine-Argenine Motifs
SNP	Single Nucleotide Polymorphism
SSR	Simple-Sequence Repeat
TALENs	Transcription Activator-Like Effector Nuclease