Phytopathology® • 2023 • 113:637-650 • https://doi.org/10.1094/PHYTO-09-22-0337-KD

Phytopathology®
REVIEW

Effector Identification in Plant Pathogens

Amelia H. Lovelace,[1] Sara Dorhmi,[1,2] Michelle T. Hulin,[1] Yufei Li,[1] John W. Mansfield,[3] and Wenbo Ma[1,†]

[1] The Sainsbury Laboratory, Norwich, NR4 7UH, U.K.
[2] Department of Microbiology and Plant Pathology, University of California Riverside, CA 92521, U.S.A.
[3] Faculty of Natural Sciences, Imperial College London, London, SW7 2BX, U.K.
Accepted for publication 7 January 2023.

Abstract

Effectors play a central role in determining the outcome of plant−pathogen interactions. As key virulence proteins, effectors are collectively indispensable for disease development. By understanding the virulence mechanisms of effectors, fundamental knowledge of microbial pathogenesis and disease resistance have been revealed. Effectors are also considered double-edged swords because some of them activate immunity in disease resistant plants after being recognized by specific immune receptors, which evolved to monitor pathogen presence or activity. Characterization of effector recognition by their cognate immune receptors and the downstream immune signaling pathways is instrumental in implementing resistance. Over the past decades, substantial research effort has focused on effector biology, especially concerning their interactions with virulence targets or immune receptors in plant cells. A foundation of this research is robust identification of the effector repertoire from a given pathogen, which depends heavily on bioinformatic prediction. In this review, we summarize methodologies that have been used for effector mining in various microbial pathogens which use different effector delivery mechanisms. We also discuss current limitations and provide perspectives on how recently developed analytic tools and technologies may facilitate effector identification and hence generation of a more complete vision of host−pathogen interactions.

Keywords: genomics, host−pathogen interactions, microbial pathogenesis, structural modeling, virulence activities

Plants defend themselves from potential parasites by mounting a myriad of immune responses. However, successful pathogens can subvert plant defenses through the function of effectors, which are virulence proteins secreted from the pathogens during infection (Hogenhout et al. 2009). Although some effectors function in the extracellular space (apoplastic effectors) within plant tissue, many are delivered into host cells (termed cytoplasmic effectors) and directly manipulate host cellular processes. In addition to defeating plant immunity, effectors also contribute to the creation of a suitable environment for pathogen colonization and proliferation (Lovelace and Ma 2022). Collectively, effectors play a key role in inducing susceptibility and are indispensable for disease. As a consequence of host−pathogen co-evolution, plants have evolved immune receptors to recognize specific effectors and activate immunity (Jones and Dangl 2006). The dynamic interactions between effectors and the host immune system determine whether or not diseases will occur. Therefore, understanding the molecular basis of this interplay provides important insights into the governing principles of infectious diseases, thus setting the foundation for developing resistance. Over the years, substantial efforts have been invested in effector research. A foundation of this research area is the characterization of the effector repertoire in a given pathogen.

It is intriguing to think that the first gene encoding a microbial effector acting in plants, the avirulence gene *avrA* from the bacterial pathogen *Pseudomonas syringae* pv. *glycinea*, was cloned almost 9 years before the term "effector protein" was first used to describe virulence determinants delivered to host cells by a microbial pathogen (Ménard et al. 1993; Staskawicz et al. 1984). The first named effectors were characterized in *Shigella flexneri*, the cause of dysentery, as the IpaB, C, and D proteins which proved essential for bacterial entry into epithelial cells (Ménard et al. 1993). Paradoxically, the effector AvrA was discovered not because of its role in virulence but because when delivered into plant cells via the Type III secretion system (T3SS) it triggers a rapid hypersensitive resistance reaction (HR) (Cornelis 2006; Staskawicz et al. 1984). Because the HR leads to robust resistance to the pathogen, the effector was named AvrA as avirulence A. For the isolation of *avrA*, a cosmid library of race 6 of *P. syringae* pv. *glycinea* was conjugated into race 5, which was unable to activate HR in the soybean cultivar Harosoy. The transconjugants were screened for gain of HR-triggering activity in Harosoy and the clones further

†Corresponding author: W. Ma; wenbo.ma@tsl.ac.uk

A. H. Lovelace, S. Dorhmi, M. T. Hulin, and Y. Li contributed equally.

Funding: Support for W. Ma was provided by Gatsby Charitable Foundation.

The author(s) declare no conflict of interest.

Copyright © 2023 The Author(s). This is an open access article distributed under the CC BY-NC-ND 4.0 International license.

characterized using Tn5-based mutagenesis, eventually leading to the sequencing of *avrA* (Napoli and Staskawicz 1987).

This elegant breakthrough approach that led to the cloning of *avrA* by function was subsequently used to clone additional *avr* genes from several *P. syringae* species and *Xanthomonas campestris*. In these pathosystems, different races of the pathogen are differentiated by their ability to infect susceptible varieties of the host plant, whereas the presence of single resistance genes confers resistance in an *avr* gene/resistance gene specific interaction between pathogen and host (Mansfield 2009). Such gene for gene interactions are now considered examples of effector-triggered immunity (ETI) (Jones and Dangl 2006). Functional cloning also allowed the identification of *avr* genes triggering resistance in nonhost plants (Dangl et al. 1992). It was not until 1999 that Jackson et al. (1999) cloned an effector, *virPphA* (*hopAB2*), based on its virulence function, i.e., suppressing the HR caused by a strain of *P. syringae* pv. *phaseolicola* in its host bean. We now know that the genes initially identified from functional screening as avirulence factors have fundamental roles in microbial pathogenicity. The term avirulence can therefore be misleading and is now generally replaced by effectors.

Through the years, many molecular tools have been developed to investigate the virulence contributions by individual effectors, and therefore virulence mechanisms have been extensively characterized. However, the major bottleneck in effector biology is effector prediction, especially the genome-wide determination of a pathogen's entire effector repertoire. This is mainly due to the diversity in effector delivery mechanisms and a lack of understanding of these mechanisms in most pathogens. Therefore, methods that can be used to increase the likelihood of identifying effectors with specific functions in disease development are needed to gain a holistic view of mechanisms underlying microbial pathogenesis, plant immunity, and host–pathogen coevolution. Such knowledge is essential for the development of resistance to economically important diseases of crops.

In this review, we summarize features related to effectors in various bacterial, fungal, and oomycete pathogens of plants. We discuss bioinformatic analysis tools that have been used to identify these features and recent analytic/technological breakthroughs that will facilitate effector detection. We apologize for not being able to include all the exciting, related research on effector biology in nematodes, insects, and parasitic plants. However, many methodologies and tools discussed in this review are also applicable to the identification and characterization of effectors from these parasites.

Effector Prediction Based on Protein Sequences

Different pathogens use different effector secretion/delivery mechanisms. The best studied example of effector delivery machinery is the T3SS of Gram-negative bacterial pathogens. T3SS encodes a injectosome, which transports T3 secreted effectors (T3SEs) directly into host cells (Green and Mecsas 2016). Features related to T3SS-dependent secretion and host cell translocation have been characterized in a small number of model plant pathogens such as *P. syringae*, *Xanthomonas* spp., *Erwinia amylovora*, and *Ralstonia solanacearum*, but these features are not conserved in other bacteria. Gram-positive pathogens and certain phloem- and xylem-colonizers, such as *Candidatus* liberibacter and *Xylella* spp., do not encode the T3SS. In these bacteria, effector delivery is dependent on the presence of the N-terminal signal peptide (SP), which is required for protein secretion through the general Sec secretion system (Natale et al. 2008). However, in contrast to the T3SS, it is likely that not all of these Sec-secreted proteins are effectors. Similarly, identification of the N-terminal SP is the major criterion used for effector prediction in eukaryotic pathogens, from which specialized secretion machinery for effectors is unknown (Petre and Kamoun 2014).

Secretion motifs

A few motifs or conserved sequences that are related to effector secretion have been used for effector identification in certain pathogens. In Gram-negative bacterial pathogens that use the T3SS, the first 50 to 100 amino acids of effector proteins are thought to contain the sequence required for secretion and translocation. Although this secretion/translocation-related region does not have a defined consensus sequence, some characteristics have been discovered in plant and human pathogens such as *Pseudomonas* and *Salmonella* species based on extensive functional screens. In *Pseudomonas syringae*, these features include a high percentage (~17%) of serine residues, an aliphatic amino acid (Asp or Glu) at position three, and the lack of a negatively charged amino acid in the first 10 residues (Guttman et al. 2002) (Fig. 1A). Using these amino acid composition bias features, the complete complement of T3SEs can be predicted from any *P. syringae* genome with a relatively high confidence. A Mann-Whitney test can be used to compare amino acid frequencies from the whole sequence and the N-termini region of selected proteins to reveal enrichments and depletions of certain amino acids (Arnold et al. 2009). However, these features are not found in other bacteria even though the T3SS machinery is conserved. Nonetheless, these T3SE-related features can serve as training sets for machine learning-based identification of effectors from related bacteria, such as rhizobia (Yang et al. 2010).

The N-terminal sequence bias has been instrumental in T3SE prediction. By contrast, in pathogens lacking the T3SS, effectors are presumably secreted through the general secretion system. Therefore, a key criterion for effector prediction is the presence of an N-terminal secretion SP (Fig. 1A). Identification of candidate secreted proteins was greatly enhanced through computational tools predicting SPs (Table 1). The first publicly available program for SP prediction, SignalP, was released in 1996 as an artificial neural network (Nielsen et al. 1996). SignalP was later improved with hidden Markov models added in version 2.0. Since then, newer versions of SignalP have added new capabilities and improved efficiency, including the most recent version 6.0 (Teufel et al. 2022) that can detect all five known types of SPs. These SPs are related to protein secretion through the Sec secretion system or Tat translocon and cleaved by signal peptidase I or II. Expansion of the types of SPs within the training set allows increased accuracy of effector prediction. Over the years, additional algorithms based on SignalP, such as PexFinder (Torto et al. 2003), Phobius (Kall et al. 2007), TargetP (Almagro Armenteros et al. 2019) and PrediSi (Hiller et al. 2004), have also been released, further strengthening the prediction of SP-based secreted proteins (Table 1). Given the wide-ranging functions of proteins secreted through the Sec pathway, additional analysis is required to confirm genuine effectors.

Domains related to eukaryote-specific functions

One indication that a bacterial protein may be an effector is the presence of a eukaryote-specific functional domain. This suggests that the proteins might function to manipulate the host by mimicking the targets of host proteins. For instance, the GALA family of *R. solanacearum* T3SEs contains F-box motif-like sequences (Angot et al. 2006). F-box proteins are related to protein ubiquitination-based proteosome degradation, which is a eukaryote-specific activity (Fig. 1A). Another example is the YopJ family of Ser/Thr acetyltransferases, which require the eukaryote-specific compound inositol hexakisphosphate for activation (Ma and Ma 2016; Zhang et al. 2016, 2017). Therefore, these bacterial proteins are expected to have specific functions in the host cells, consistent with their role as effectors. Sequence homology based-methods can be used to identify eukaryotic domains within bacterial proteins. For instance, BLASTP and Pfam provide large collections of eukaryotic protein families that can serve as a database from which effectors can be searched. Several scanning tools can be used to screen bacte-

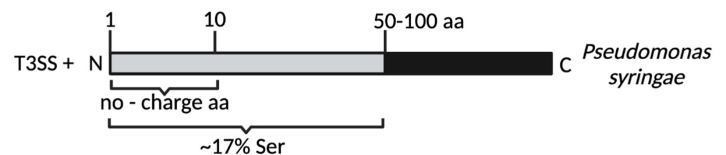
rial proteome for proteins containing motifs of interest. These tools include the MEME suite (Bailey et al. 2009) and MotifScan (Sun et al. 2018) (Table 1).

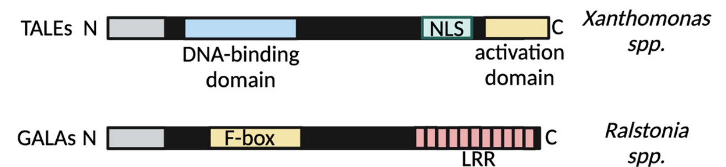
Another feature that could be used to identify bacterial effectors is that they often target host cellular compartments such as the nucleus, mitochondria, and chloroplasts. Therefore, these effectors will likely contain plant cell compartmentalization signals or transit peptides that exploit the host machinery to enter the specific organelles. Many effectors have been predicted to contain the nuclear localization signals (NLS). The best examples are the transcription activator-like effectors (TALEs) encoded by *Xanthomonas* spp. (Mak et al. 2013). TALEs bind to specific promoter sequences in the plant genome and manipulate host gene expression. The NLS is essential for their entry into the plant cell nucleus (Fig. 1A). Other bacterial pathogens, such as the phloem-colonizing phytoplasmas, also produce effectors that target the plant nucleus. The predicted NLS in phytoplasma effectors is required for their manipulation of plant transcription factors (Sugio et al. 2014). Some effectors carry transit peptides that relocate the effector to a specific organelle after secretion (Hicks and Galan 2013). For examples, the *P. syringae* T3SEs HopI1, AvrRps4, and HopK1 have predicted chloroplast targeting signals (Jelenska et al. 2007; Li et al. 2014). Valuable prediction tools, such as LOCALIZER (Sperschneider et al. 2017), DeepLoc 2.0 (Thumuluri et al. 2022), and TargetP 2.0 (Almagro Armenteros et al. 2019), use a sliding window approach to predict chloroplast/mitochondrial transit peptides and/or NLS in a protein (Table 1).

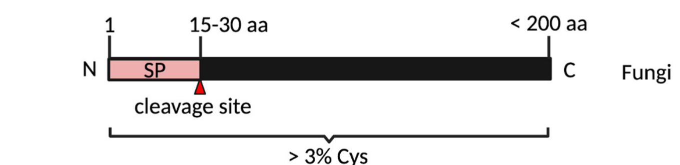
Prediction of secreted proteins in eukaryotic pathogens

The vast majority of the effectors from filamentous eukaryotic pathogens including fungi and oomycetes are defined by N-terminal SP using bioinformatic programs discussed above such as SignalP and PexFinder. In fungi, most effectors have been found to be small, secreted proteins that are shorter than 150 to 200 amino acids and rich (>3%) in cysteine residues (Sperschneider et al. 2015). However, these features have not been linked directly to entry into host cells. Nevertheless, the identification of "small, cysteine-rich" secreted proteins has been widely used to predict fungal effectors in a pipeline which uses cumulative features of known effectors (Fig. 1B). Once a fungal secretome (proteins predicted to have an SP) is defined, the pipeline can be used to filter out proteins that are too large or cysteine-poor to be considered effector candidates (Saunders et al. 2012).

In fungal pathogens, other than the SP, only one motif that is potentially associated with effector secretion has been identified. Highly expressed genes at the specialized infection structure, haustoria, formed by the barley powdery mildew pathogens, *Blumeria graminis* f. sp. *hordei*, were found to encode proteins sharing a Y/F/WxC (x is any amino acid) motif in the N-terminus, downstream of the predicted secretion SP (Godfrey et al. 2010). Although this motif has yet to be functionally characterized, it can be used to identify effectors in newly sequenced *Blumeria graminis* f. sp. *hordei* strains or sister species using hidden Markov model (HMM)-based programs such as HMMER (Eddy 2009) (Table 1).







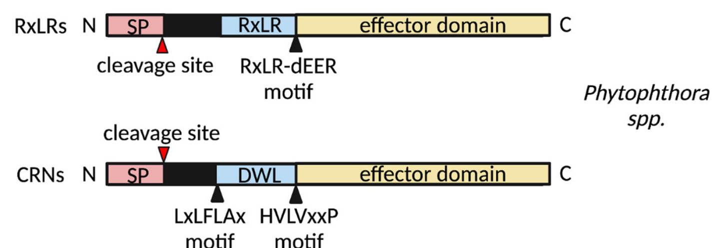


FIGURE 1 Protein sequence features in effectors from bacterial and filamentous plant pathogens. **A,** N-terminal features of bacterial effectors secreted via the Type III secretion system (T3SS+) in *Pseudomonas syringae* or secreted via other means (T3SS−). Domain and motif features of *Xanthomonas* Transcription activator-like effectors (TALEs) and *Ralstonia* "GALA" effectors. **B,** Sequence and domain/motif features of filamentous plant pathogen effectors including those belonging to the RxLR and CRN family in *Phytophthora*. Different domains and motifs are highlighted in separate colors. SP = signal peptide, aa = amino acid, Cys = cysteine, and Ser = serine. Figures are not drawn to scale.

TABLE 1
Bioinformatic tools/resources used for effector prediction

Tool[a]	Description	Availability	Reference
Sequence motif analysis			
MEME-suite	Motif-based sequence analysis	Web API	Bailey et al. (2009) https://meme-suite.org/meme/
HOMER	Promoter motif identification	Stand-alone	Heinz et al. (2010) http://homer.ucsd.edu/homer/motif/
HMMER	HMM profile-based search	Web Stand-alone	Eddy (2009) http://hmmer.org/
MotifScan	Identification of known motifs in sequences	Web	Sun et al. (2018) https://myhits.sib.swiss/cgi-bin/motif_scan
SignalP	Secretion signal prediction	Web Stand-alone	Teufel et al. (2022) https://services.healthtech.dtu.dk/service.php?SignalP
PrediSi	Secretion signal prediction	Web Stand-alone	Hiller et al. (2004) http://www.predisi.de/
Phobius	Secretion signal prediction	Web Stand-alone	Kall et al. (2007) https://phobius.sbc.su.se/
Interproscan	Domain identification	Web Stand-alone	Apweiler et al. (2001) https://www.ebi.ac.uk/interpro/search/sequence/
Protein disorder prediction			
IUPred2A	Prediction of disordered regions in protein sequence	Web Stand-alone	Mészáros et al. (2018) https://iupred2a.elte.hu/
Effector subcellular localization			
LOCALIZER	Intracellular locations of eukaryotic effectors	Web Stand-alone	Sperschneider et al. (2017) https://localizer.csiro.au/
DeepLoc 2.0	Intracellular locations of eukaryotic effectors	Web Stand-alone	Thumuluri et al. (2022) https://services.healthtech.dtu.dk/service.php?DeepLoc-2.0
TargetP 2.0	Intracellular locations of eukaryotic effectors	Web Stand-alone	Almagro Armenteros et al. (2019) https://services.healthtech.dtu.dk/service.php?TargetP-2.0
Genomic feature identification			
ISFinder	Identification of insertion sequences	Web	Siguier et al. (2006) https://isfinder.biotoul.fr/
MGEfinder	Identification of MGE in bacterial genomes	Stand-alone	Durrant et al. (2020) https://github.com/bhattlab/MGEfinder
Islandviewer4	Identification of genomic islands in bacterial genomes	Web API	Bertelli et al. (2017) https://www.pathogenomics.sfu.ca/islandviewer/
PHASTER	Prophage identification	Web API	Arndt et al. (2016) https://phaster.ca/
REPET	Identify TEs in eukaryotic genomes	Stand-alone	Quesneville et al. (2005) https://bio.tools/repet
RepeatMasker	Identification of repetitive regions in genomes	Stand-alone	Smit et al. (2013-2015) https://www.repeatmasker.org/
Protein 3D structure prediction			
MODELLER	Homology modelling	Stand-alone	Webb and Sali (2016) https://salilab.org/modeller/
SWISS-MODEL	Homology modelling	Web	Waterhouse et al. (2018) https://swissmodel.expasy.org/
Phyre2	Homology modelling	Web	Kelley et al. (2015) http://www.sbg.bio.ic.ac.uk/phyre2/
I-TASSER	Homology modelling	Web Stand-alone	Roy et al. (2010) https://zhanggroup.org/I-TASSER/
HHPred	Homology modelling	Web Stand-alone	Söding et al. (2005) https://toolkit.tuebingen.mpg.de/tools/hhpred
RoseTTAFold	Template-free	Web Stand-alone	Baek et al. (2021) https://robetta.bakerlab.org/
AlphaFold2	Template-free	Collabfold (web) Stand-alone	Jumper et al. (2021) https://www.deepmind.com/open-source/alphafold
Machine/deep learning effector prediction			
MacSyfinder	Bacterial secretion systems	Stand-alone	Abby et al. (2016) https://github.com/gem-pasteur/macsyfinder
EffectiveT3	T3, T4 effectors	Web Stand-alone	Arnold et al. (2009) https://effectivedb.org/method/effectivet3
Effectidor	T3 effectors	Web	Wagner et al. (2022) https://effectidor.tau.ac.il/
T3SEpp	T3 effectors	Web Stand-alone	Hui et al. (2020) http://www.szu-bioinf.org/T3SEpp/
Preffector	T1-6 effectors	Web	Dhroso et al. (2018) http://korkinlab.org/preffector
Deepredeff	Bacteria, fungi, oomycetes	Stand-alone	Kristianingsih and MacLean (2021) https://ruthkr.github.io/deepredeff/
EffectorP	Fungal, oomycete	Web Stand-alone	Sperschneider et al. (2016) https://effectorp.csiro.au/
EffectorO	Oomycete	Web Stand-alone	Nur et al. (2021) https://github.com/mjnur/oomycete-effector-prediction

[a] This table provides a list of representative software for each task but it does not include all possible software available.

Although morphologically similar to fungi, oomycetes are structurally and evolutionarily distinctive. Effector prediction is more confident in oomycete pathogens due to the discovery of host-targeting motifs in the N-terminal region of cytoplasmic effector proteins following the secretion SP. The majority of these effectors carry the RxLR-dEER motif (Rehmany et al. 2005), which is similar to a signal required for translocation of protein from the malaria parasites (Hiller et al. 2004) (Fig. 1B). Another class of cytoplasmic effectors are CRNs (for crinkling and necrosis), which are widespread in oomycetes (Amaro et al. 2017; Torto et al. 2003). In addition to an N-terminal host-targeting motif LxLFLAx, CRNs have a DWL domain, which harbors a conserved HVLVxxP motif (Fig. 1B). Predictions combining the secretion SP and these motifs have been widely used in defining effector repertoires in oomycetes, especially *Phytophthora* species where these motifs were best characterized. Again, HMMER is a useful tool in these analyses.

Effector Prediction Using Structural Features

Protein structural information can provide deeper insights into function than sequences alone. For effector prediction, structural analysis is particularly useful because effectors are notorious for not sharing sequence homology with known function proteins and they often also lack known sequence motifs. Structural information was traditionally only available through crystallography studies, which causes a major bottleneck due to the tremendous effort required. The idea of protein structural modeling based on similarity to experimentally determined structures dates back to as early as 1969 (Browne et al. 1969). MODELLER was the first available program for 3D structure prediction (Šali and Blundell 1993). Since the early 2000s, a range of tools has been developed (Table 1). Many of these structure prediction programs such as SWISS-MODEL, Phyre2, I-TASSER, and Hhpred (Kelley et al. 2015; Roy et al. 2010; Söding et al. 2005) use homology modeling. This involves identifying known experimental structures with sequence similarity, which can be used to build a model of the unknown protein. If the effector of interest is structurally similar to a protein with a known domain, this method can allow us to infer that the effector may also possess this domain and its associated functions. However, these analyses are limited when no homologous proteins with experimentally determined structures can be identified. Recent innovations in template-free structural modeling using artificial intelligence (AI), including AlphaFold2 (Jumper et al. 2021) and RoseTTAFold (Baek et al. 2021), revolutionized effector identification using structural features because they do not require a solved homologous protein structure to act as the template. These AI-based prediction typically involve the generation of a multiple sequence alignment (MSA), which is then used to predict features such as secondary structure, backbone torsion angles, and a residue-residue contact map. These features then dictate the prediction of a 3D model. If sequence homologs are not available, the program can also run without an MSA. Many new approaches combine strengths of both template and template-free modelling to further increase the robustness of structure prediction (Kuhlman and Bradley 2019).

Effectors with divergent sequences may form similar structural folds or include domains with similar folds (recently reviewed by Mukhi et al. [2020] and Outram et al. [2022]). This could be the consequence of convergent evolution of unrelated proteins that have evolved to have the same fold and hence function, or loss of sequence similarity between distant homologs due to the rapid evolution during host−pathogen arms race (Seong and Krasileva 2021). The enrichment of the same folds/domains in multiple effectors suggests they are related to effector functions such as secretion and/or host manipulation. Therefore, the presence of this fold can be a useful indicator that the protein may be an effector. For example, the MAX (*Magnaporthe* Avrs and ToxB like) fold has been found in effectors produced by ascomycete fungal pathogens. Although often unrelated in sequences, these effectors share the common "MAX" fold, which is made up of six β-sheets stabilized by a conserved disulfide bridge (de Guillen et al. 2015) (Fig. 2B). RNAse-like folds are also commonly identified in fungal effectors (Fig. 2B). These were first identified in *Blumeria graminis* using Interproscan and 3D structure prediction and are the most abundant effector type in this species (Pederson et al. 2012; Pennington et al. 2019).

Similarly, many RxLR effectors in *Phytophthora* and their relative species contain the WY and LWY fold despite limited sequence conservation (Boutemy et al. 2011; Jiang et al. 2008; Wood et al. 2020). The WY fold consists of a three or four α-helical bundle and the LWY fold forms a five α-helical bundle (He et al. 2019; King et al. 2014; Maqbool et al. 2016; Win et al. 2012). The (L)WY units are often arranged as tandem repeats. In particular, the LWY repeats enable a non-globular protein structure, attributed to the rigid linkages between adjacent units (He et al. 2019) (Fig. 2A). Indeed, it has been shown that effector prediction based on the presence of the WY fold may be more efficient than using the RxLR sequence motif in the lettuce pathogen *Bremia lactucae* (Wood et al. 2020). It is unquestionable that by utilizing AlphaFold- or RoseTTAFold-based structural modeling of secreted proteins and identifying the presence of these effector-enriched folds will strengthen effector predictions in fungal and oomycete pathogens respectively.

An additional example of a shared fold is the *Leptosphaeria* avirulence and suppressing (LARS) fold, which was characterized in the oilseed rape stem canker pathogen *Leptosphaeria maculans.* The LARS fold was identified in the crystal structures of *L. maculans* effectors AvrLm4-7 and AvrLm5-9, which both possess a well-defined antiparallel β-sheet and a set of disulfide bonds despite showing limited sequence similarity (Fig. 2B). Prediction based on the presence of a LARS fold using HMM and AlphaFold2 identified 13 new effector candidates that would not have been identified by sequence-based analysis (Blondeau et al. 2015; Lazar et al. 2022). The LARS fold has recently also been identified in the rice blast fungus *Magnaporthe oryzae* (Seong and Krasileva 2021), indicating that it can be used for effector prediction from a wide range of fungal species.

Structural features overrepresented in effector proteins have also been found in bacteria. For example, T3SEs may share folds associated with mechanistical lability to allow structural flexibility during translocation (LeBlanc et al. 2021). It has also been observed that T3SEs often have a structurally disordered N-terminus (Buchko et al. 2010) and may possess structural motifs required for chaperone-binding important for the secretion process (Costa et al. 2012). Studies across bacteria have shown that intrinsic disorder is uncommon in bacterial proteomes and therefore the enrichment of this feature in T3SEs can be used as a criterion for effector prediction (Chen and Xia 2021; Dunker et al. 2000). Intrinsic disorder within protein sequences can be predicted using webservers such as IUPred2A (Mészáros et al. 2018; Necci et al. 2021). Although intrinsic disorder is not specifically predicted during 3D structural prediction, the regions modeled with lower confidence (represented by a lower per-residue confidence score or pLDTT score in AlphaFold2 models) could represent disordered regions (Wilson et al. 2022).

As discussed above, the presence of eukaryote-specific functional domains in a bacterial protein is suggestive of their activity in the host cell and therefore can be used as for effector prediction. Structure- but not sequence-based homology to eukaryotic proteins has been found in bacterial effectors. For example, the *P. syringae* T3SE AvrPtoB contains a C-terminal domain that forms a structural mimic of RING-finger and U-box proteins, which are E3 ubiquitin ligases that are specific to eukaryotes (Janjusevic et al. 2006). Importantly, the sequence of this domain in AvrPtoB has no similarity to known E3 ligases. With the revolution in structural

model prediction by AlphaFold2, it will be possible to identify bacterial effector candidates by predicting structural folds that may be related to host manipulation. Furthermore, virulence mechanisms of effectors with unknown functions will also be illuminated from the structural features. It is now possible to perform genome-wide structural prediction for individual plant pathogen proteomes. It will likely soon also be computationally possible to perform large-scale comparative structural genomics within and between species. This will enable the identification of novel folds to enhance effector predictions in known and newly characterized pathogens.

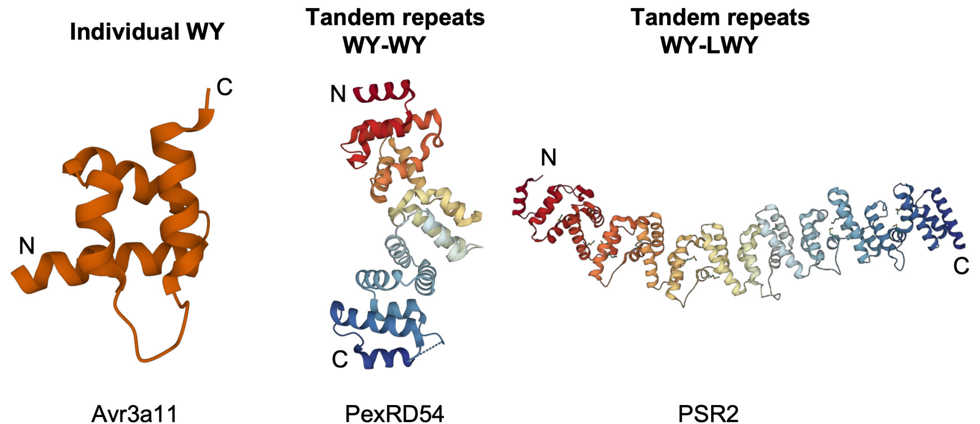
Effector Prediction Facilitated by Genomic Analysis

Next generation sequencing technologies have allowed fast and economical determination of genome sequences, which enables research to move away from identification of single effectors with defined virulence or immunity-activating phenotypes to studying effector repertoires as a whole. In particular, long-read sequencing technologies such as Pacific BioSciences (PacBio), single-molecule real-time (SMRT), and Oxford Nanopore (ONT) (Besser et al. 2018) have allowed the accurate resolution of repeat rich regions which may contain effectors. It is now common practice to generate complete bacterial genomes (Smits 2019) and probably soon chromosome-level eukaryotic genomes using these technologies.

Thanks to the advancement in sequencing technologies and the affordability of full genome sequencing, plant pathology research has advanced towards studying species pangenomes rather than individual model strains. The pangenome is defined as the entire set of genes within a species. This covers both the core genome, which consists of conserved genes encoding proteins with essential functions and the accessory genome, which consists of genes not present in all members of the species and may be dispensable (Tettelin et al. 2005). Effector genes are often within the accessory genome and exhibit presence-absence polymorphism even between closely related pathogen strains (Dillon et al. 2019; Jones et al. 2021). Pangenomic analyses aim to saturate the gene content and therefore fully survey the effector diversity present in the species. Moving away from a single "representative" strain is significant because it allows the characterization of the whole effector repertoire in pathogens, therefore providing a more complete understanding of pathogenesis. Pangenomic analyses have been conducted in several bacterial species (Dillon et al. 2019; Sabbagh et al. 2019; Siani et al. 2021). Insight gained from these analyses improves understanding of the sequence and structural features that define effectors, which in turn benefits effector characterization in general.

Selection analysis to identify effectors

The dynamic interplay between effectors and host targets implies endless co-evolution between the pathogens and hosts. A hallmark of host−pathogen arms race is the accelerated evolution of effectors which frequently undergo positive (diversifying) selection (Menardo et al. 2017; Stukenbrock et al. 2011; Zhang et al. 2015). In comparison to the co-evolution model of single effector-immune receptor gene pairs, an alternative model involves the



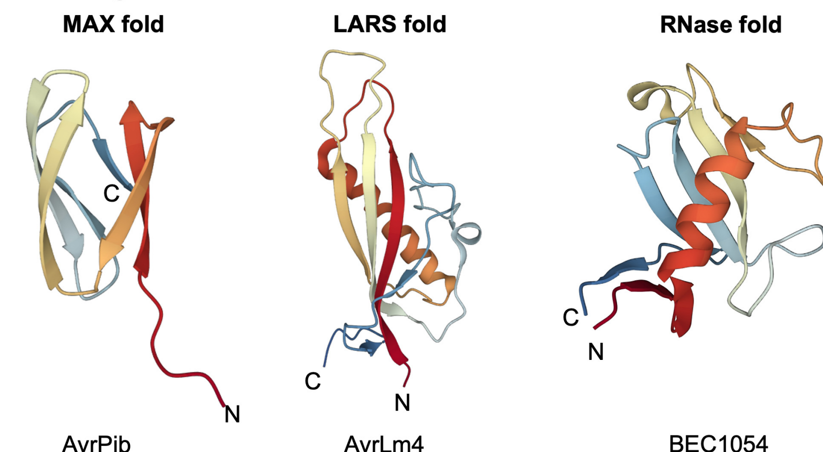


FIGURE 2 Structural folds enriched in effectors from oomycetes and fungi. **A,** The structures of oomycete WY domain effectors. Structures shown are defined and individual WY (Avr3a11), and tandem repeats (PSR2). **B,** Structures of fungal effectors displaying MAX fold (AvrPib), LARS fold (AvrLm4), and RNase fold (BEC1054). Structures downloaded from Protein DataBase.

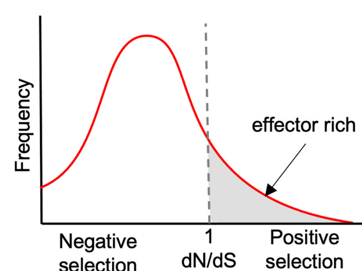
"trench-warfare" scenario where effector and host immune genes undergo rapid recycling with gain/loss and mutation occurring to the repertoire as a whole (Derbyshire 2020). Population genomics approaches used to identify loci undergoing positive selection include statistical methods that measure dN/dS ratio and Tajima's D selective sweeps. The dN/dS ratio measures nonsynonymous (dN) to synonymous (dS) substitutions per site and predicts fixation events between species assuming that a certain level of sequence diversity is present. A ratio of >1 indicates positive selection, whilst <1 indicates negative selection (Fig. 3A). The Tajima's D statistic is used in combination with other statistical methods such as the composite likelihood ratio (CLR) (Derbyshire 2020; Hartmann et al. 2018). These methods are applied within a species to identify drops in nucleotide diversity across the genome which may represent a selective sweep, which is represented by an increase in the frequency of a particular allele in a population (Fig. 3A).

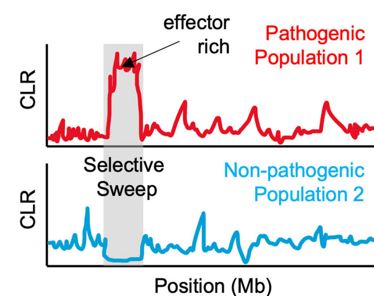
Signatures of positive selection have been used as a criterion for effector identification with those under greater positive selection being scored more likely as effector candidates (Syme et al. 2018) (Fig. 3A). An example of positive selection analyses is the genomic analysis of sister species *Phytophthora infestans* and *P. mirabilis*, which revealed that effectors show dN/dS > 1, suggesting that they undergo positive selection (Haas et al. 2009; Raffaele et al. 2010). Studies of fungal pathogens have also used selective sweep analyses to identify effectors (Badouin et al. 2017; Richards et al. 2019). In the bacterial pathogen *P. syringae*, the effector HopZ1 shows strong evidence of positive selection at positions important for host recognition (Ma et al. 2006; Zhou et al. 2009). In addition, these methods have been applied to bacterial pathogens to identify genomic regions under positive selection (Singh and Khan 2019; Zhang et al. 2015), which could be used in combination with gene-based selection analysis to strengthen effector identification.

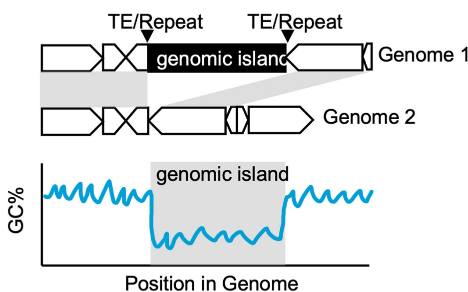
It should be noted that not all effectors are necessarily under positive selection. Conserved effectors with essential virulence activities may undergo purifying selection, leading to dN/dS ratio < 1, due to their indispensable roles in pathogenicity (McCann and Guttman 2008). In addition, different region(s) of the same effector genes may exhibit different evolution patterns with certain sequences under purifying selection whilst others under diversifying selection (Win et al. 2007). It has been shown in the *P. syringae* species complex that T3SEs are not all undergoing diversifying selection; rather, the identification of those exhibiting evolution noncongruent with the core genome (having undergone recombination or horizontal gene transfer [HGT]) was a better way to detect effectors (Dillon et al. 2019).

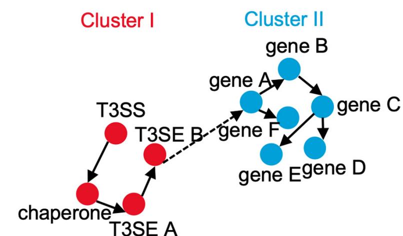
Genome compartmentalization of effector loci

An important mechanism of effector evolution is through HGT. Effectors are often associated with mobile genetic elements (MGEs) or located in genomic regions that promote HGT. In bacteria, effectors are commonly located with transposable elements (TEs), integrative conjugative elements, plasmids, prophage sequences or









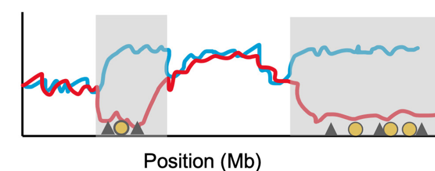


FIGURE 3 Genomic analysis for effector identification. **A,** Identification of regions of positive selection using dN/dS ratio (left) and selective sweep (right) analyses at the protein and genome-level respectively. **B,** Genome compartmentalization analyses to identify mobile genetic elements such as pathogenicity islands through GC content and transposable element (TE) abundance. Co-occurrence analysis identifies virulence genes in clusters such as those linked to the Type III secretion system (T3SS) and Type III effectors (T3SEs). Arrows indicate direction of gene orientation. Accessory or mini chromosomes contain gene sparse regions (GSR), gene dense regions (GDR) and repeat regions enriched in effectors and TEs. CLR = composite likelihood ratio.

hybrids of these such as genomic islands (Jackson et al. 2011). Mobile elements not only play a major role in the movement of effectors between and within genomes, but also in gene loss through inactivation through pseudogenization/promoter loss, or the birth of new effectors through recombination of effector promoters/signal peptide sequences with novel sequences (Jackson et al. 2011; Stavrinides et al. 2006). The proximity of mobile elements is therefore a useful parameter for identifying effector genes. Recent import of effectors into bacterial genomes on mobile elements leads to characteristic signatures such as a change in GC content in comparison to flanking regions (Fig. 3B). Therefore, bioinformatic tools developed for the identification of mobile elements such as MGEfinder, which is used to identify insertion sites of MGEs and determine gene ontology of nearby genes, is useful to identify putative effectors (Durrant et al. 2020). In addition, tools for the identification of genomic islands (Bertelli et al. 2017), prophages (Arndt et al. 2016), and insertion sequences (Siguier et al. 2006) can also be used for this purpose (Table 1).

Advanced genome sequencing of various pathogens suggests a bipartite genome organization, in which genes responsible for housekeeping and virulence functions are enriched in different compartments. Pathogen virulence genes are often flanked by TEs in the genome as TEs provide a genetic source for sequence variation and therefore facilitate rapid gene evolution. The genomic region coding for these virulence genes displays a negative correlation between TEs and gene density. Based on intergenic distances between genes, pathogen genomic regions can be defined as gene dense region (GDR) and gene sparse regions (GSR) (Dong et al. 2015; Faino et al. 2016; Wang et al. 2017). In *Phytophthora*, many RxLR and CRN effectors are enriched in GSR and the association with TEs in these regions have been proposed to serve as a stimulator that promotes fast evolution (Dong et al. 2015) (Fig. 3B). In some pathogens, effector enriched compartments have higher levels of sequence polymorphisms and/or positive selection than genomic segments with higher gene density (Faino et al. 2016; Raffaele et al. 2010). This phenomenon was described as a “two-speed genome,” which indicates uneven evolutionary rates between these genomic compartments. In addition, in some fungi, TE activities are constrained by repeat-induced point (RIP) mutations, resulting in AT-rich isochores (Grandaubert et al. 2014). Effector genes are often found in these AT-rich regions such as in *L. maculans* (Rouxel et al. 2011).

While TE and effector enriched genomic regions can reside on core chromosomes (chromosomes harbour housekeeping genes), these contents are also present on chromosomes that are not shared by all members of species (namely accessory chromosomes) (Goodwin et al. 2011; Miao et al. 1991; Peng et al. 2019; Schotanus et al. 2015). These variable chromosomes are also referred to as lineage specific (LS), conditionally dispensable (CD), supernumerary, and mini chromosomes (Ma et al. 2010; Vlaardingerbroek et al. 2016). Like accessory genomic regions, TE-rich landscapes and numerous effectors are often found on accessory chromosomes and were later proven to contribute to virulence (Bao et al. 2017; Ma et al. 2010; Peng et al. 2019; Schmidt et al. 2013). In the case of mini chromosomes (chromosomes with size less than 3 mb), high proportions of repetitive elements and putative effector genes have been discovered in *F. solani*, *M. oryzae*, and many other filamentous pathogens (Bao et al. 2017; Han et al. 2001; Peng et al. 2019; Temporini and VanEtten 2004) (Fig. 3B). Tools available for identifying repeat-rich regions include REPET and RepeatMasker (Flutre et al. 2011; Quesneville et al. 2005; Smit et al. 2013−2015) (Table 1).

In addition to co-occurrence with TEs, virulence genes can sometimes form clusters in bacteria, fungi, and oomycete genomes. For example, 12 gene clusters encoding secreted proteins are found in the fungal pathogen *Ustilago maydis*. This pathogen is devoid of repetitive elements and thus clustering of effector genes provides an alternative strategy for candidate identification (Kämper et al. 2006). In bacteria, effector genes located as gene clusters can also be co-regulated as operons. The proximity to known effectors and their cognate chaperones is therefore another characteristic useful for effector prediction. In bacteria possessing the T3SS, such as *P. syringae*, *Erwinia amylovora*, and specific pathovars of *Pantoea agglomerans*, the T3SS machinery gene cluster is adjacent to the highly conserved effector gene family *avrE* (Alfano et al. 2000). In these diverse bacterial species other effectors are also often located in this conserved locus, such as *hopM1* and *hopAA1* in *P. syringae* (Alfano et al. 2000; Xin et al. 2018). In the future, co-occurrence analyses could be performed to identify effector candidates that are more often closely located with known or predicted effector genes in the genome as an additional feature to facilitate effector prediction (Fig. 3B).

Effector Prediction Based on Gene Expression Patterns

In practice, induced expression in planta is often considered as a feature of genes encoding effectors. This information is beneficial to narrow down candidates for further validation or functional characterization. It is also important to note that some effector genes may be silenced as a mechanism to avoid host recognition during specific interactions (Dong and Ma 2021). Nonetheless, induced expression is suggestive of an active role in host manipulation, consistent with the prediction that these genes may encode effector proteins.

Co-regulation of effector genes during infection was well documented in T3SS and related effector genes in bacterial pathogens. These so called “*hrp*” gene clusters are commonly regulated by specific sigma factors, such as HrpL, which interact with defined cis-regulatory elements in their promoters (O’Malley and Anderson 2021). In bacterial pathogens with known virulence regulators, effectors can be identified using genomic screenings of the cis-regulatory elements in combination with high throughput expression screenings (Furutani et al. 2006). Additionally, in vivo expression technology (IVET) and inducible FACS assays capitalized on gene expression dependent on the alternative sigma factors has also been used to identify novel effectors (Boch et al. 2002; Chang et al. 2005). However, pathogenicity genes that may not necessarily encode effectors may also be pulled out from these analyses.

In bacterial pathogens that do not rely on the T3SS or pathogens in which the regulatory proteins required for virulence are poorly understood, RNA-Seq has been widely used to identify effector candidates. For bacterial pathogens, a major limiting factor of performing RNA-Seq on infected tissue is the underrepresentation of bacterial RNAs, which lack 3′-polyA, in total RNA extracts, resulting in poor coverage of the pathogen transcriptome. Recently, cost-effective methods have been developed to enrich bacterial RNAs through physical separation of the bacteria cells from infected tissues, thus enabling global expression analysis including effector genes (De Francesco et al. 2022; Lovelace et al. 2018; Nobori et al. 2018). Filamentous pathogens have complicated disease cycles that involve various cell types and status which adds further complexity to determine specific gene expression during infection. Several studies have determined stage-specific transcriptomes from infected tissues for identifying putative effectors in fungi such as the wheat rust pathogen *Puccinia triticina* (Hu et al. 2007), the rice blast pathogen *M. oryzae* (Mosquera et al. 2009; Yan et al. 2022), and barley powdery mildew *Blumeria graminis* f. sp. *hordei* (Godfrey et al. 2010). Furthermore, laser microdissection has been used to separate plant material in close proximity to infecting filamentous pathogens (Fosu-Nyarko et al. 2010; Tang et al. 2022), providing a more precise way of evaluating gene expression at the plant−pathogen interface.

Gene expression clustering analyses can be used to predict new effectors by locating co-expressed genes with known effectors. For example, effectors arranged as gene clusters on the genome of some filamentous pathogens showed a co-expression pattern and were also induced during infection (Kämper et al. 2006; Keller and Hohn 1997). Based on the co-expression dataset, motif enrichment analysis of the promoters of co-expressed genes and effectors can then identify putative cis-regulatory elements of transcriptional regulators. The MEME-suite provides various programs for the enrichment and identification of sequence motifs (Table 1). To date, only a few fungal transcriptional regulators have been identified to contribute to effector expression in planta (Jones et al. 2019; Tang et al. 2022; Tollot et al. 2016). These regulatory elements can be used to identify their regulons using techniques such as chromatin immunoprecipitation followed by sequencing (ChIP-Seq), thus expanding the identification of novel effectors.

Validation of Predicted Effectors

Predicted effectors should be validated by experimental approaches. The gold standard is demonstration of virulence function, but this has been achieved with few effector proteins. It is also not feasible to test all the effector candidates by functional validation. Considerable success in demonstrating avirulence functions in effectors from oomycetes and fungi has been achieved by the use of *Agrobacterium*-mediated co-expression of plant resistance and effector proteins in leaves of *Nicotiana benthamiana.* Induction of the HR and also other effector functions such as organelle targeting using microscopy can be assessed. However, visualization of effectors to validate translocation during infection is extremely challenging. Although examples of microscopy demonstrating effector entry have been reported, such evidence is rather uncommon. The first effectors to be visualized using fluorescent fusion proteins included the *Ralstonia solanacearum* T3SE PopP2 during the infection of *Arabidopsis thaliana* (Deslandes et al. 2003) and the *Uromyces fabae* rust transferred protein 1 in the host *Vicia faba* (Kemen et al. 2005). To augment the fluorescence signals, introduction of an NLS into effectors has been used to enhance confirmation of translocation. This approach has been successfully applied to several fungal pathogens, allowing the visualization of effector cell-to-cell movement (Khang et al. 2010). The average size of a fluorescent protein is around 28 kDa which may interfere with effector translocation; therefore, an assay using a split fluorescent protein system has been used as a work-around. Green fluorescent protein (GFP) is comprised of 11 beta barrel strands; the 11th strand (GFP_{11}) was fused to two *P. syringae* T3SEs, which allowed their visualization during infection of transgenic Arabidopsis plants expressing the remainder of the GFP protein ($GFP_{1\text{-}10}$) (Henry et al. 2017).

Assays to validate effector translocation indirectly are also available. For T3SE candidates, the N-terminal region can be fused to the C-terminal region of a known T3SE AvrRpt2, which triggers the HR in Arabidopsis plants containing the cognate immune receptor Rps2. If a candidate effector-AvrRpt2 fusion can trigger immunity, it indicates that the N-terminal region of this candidate can lead the translocation of the fusion effector (Guttman et al. 2002). A bacterial translocation assay based on an adenylate cyclase (Cya) reporter can also be used to determine whether a candidate protein is translocated by the T3SS in a semiquantitative manner (Schechter et al. 2004). This assay fuses effectors to the Cya domain whereby after translocation through the T3SS, plant host calmodulin converts adenosine triphosphate into cyclic adenosine monophosphate (cAMP), which is then measured using an enzyme-linked immunosorbent assay. In the fungal pathogen *Ustilago maydis*, a translocation assay for effectors has been described based on in vivo biotinylation in maize (Lo Presti et al. 2017). It should, however, be stressed that translocation to the host cell per se does not alone confirm effector function.

Machine learning to predict effectors

Advances in machine learning are transforming many aspects of biological research. These techniques are often used as specific tools to predict protein features as mentioned above. The power of machine learning becomes apparent when multiple classes of information are incorporated. With the wealth of "omics" data available and a range of protein sequence and structural features that indicate effector identity, machine learning has become a useful tool for de novo identification of effectors (Fig. 4). So far, these approaches have mostly utilized genomic and protein features, however transcriptomic data and structural information could readily be incorporated into the pipeline and provide additional information for training the models to improve their accuracy and sensitivity. The genomes of model pathogens with well-defined effector repertoires can be used to train AI to differentiate effectors from noneffectors within proteomes. One of the first algorithms for identifying effectors within bacterial proteomes using machine learning was EffectiveT3 (Arnold et al. 2009) (Table 1). Since then, various tools have been developed relying on distinct machine learning techniques and different protein features to identify mostly T3SEs (Sperschneider 2020). The tool Preffector can predict effectors from secretion systems Type 1-6 in bacterial proteomes (Dhroso et al. 2018). For fungi and oomycete effectors, various tools such as EffectorP (Sperschneider et al. 2015) and EffectorO (Nur et al. 2021) are available. EffectorP has been optimized over time to be trained on further diverse effectors and distinguish apoplastic from cytoplasmic effectors. In addition, a deep-learning tool that does not rely on specific feature selection, Deepredeff (Kristianingsih and MacLean 2021), has recently been developed (Table 1).

Most tools utilize protein sequences, particularly trained using the N-terminal region features such as amino acid composition, position-specific scoring matrix and secondary structure informa-

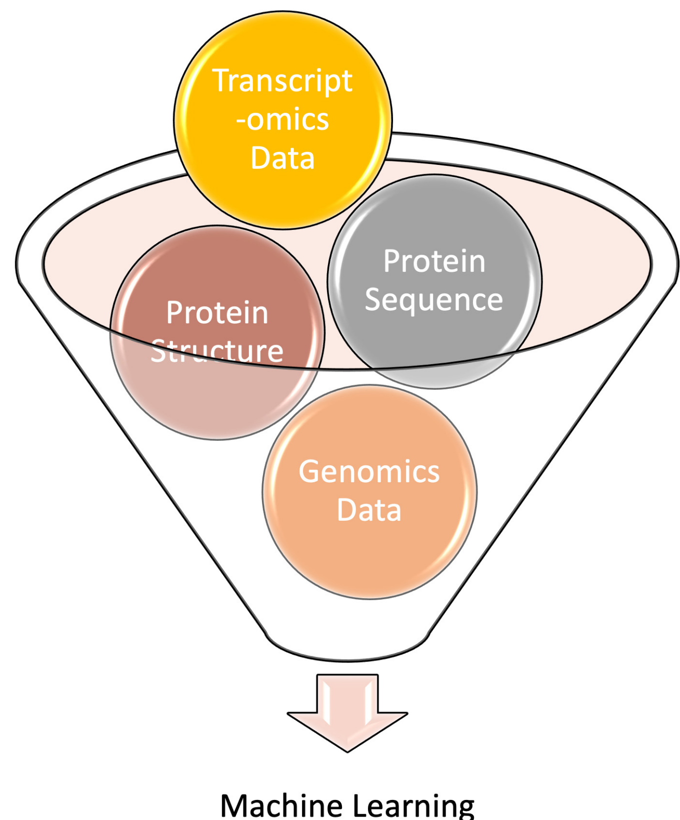


FIGURE 4
Integration of "omics" data along with structural and protein sequence features for de novo identification of effectors through machine learning.

tion. However, a recently released program, Effectidor, can integrate other information such as sequence motifs in the promoter and genomic location (Wagner et al. 2022). Another program T3Sepp also includes an option to add promoter sequences to facilitate effector prediction (Hui et al. 2020) (Table 1). The addition of transcriptional patterns would also provide a useful addition, particularly for oomycete and fungal effectors that show distinct expression profiles during infection (Evangelisti et al. 2017; Yan et al. 2022).

There are some shortcomings to machine learning-based analysis. Existing methods are tailored towards specific species or certain subclasses of effectors; therefore, it is challenging to predict novel effectors lacking these features. For example, bacterial pathogens lacking the T3SS, such as phloem/xylem-limited species that utilize secretion systems on which these tools have not been trained. In addition, these methods are highly reliant on accurate annotation of genomes, meaning that candidates may be missed if they are not well annotated, for example when they are misannotated on the N-terminal secretion signals.

Conclusions and Future Perspectives

Effector identification should now be one of the first steps in the study of pathogenesis. This is particularly important for biotrophic pathogens, which are mostly unculturable in artificial media and hence greatly limit the experimental approaches applicable. Individual investigation of effector functions in planta and in vitro has been used as a strategy to circumvent this limitation. In addition, the functional screening of effectors that can activate immunity in resistant cultivars or wild relatives of crops (avirulence factors) has also been proven useful to identify immune receptors that can be incorporated into elite cultivars to achieve disease resistance (Lin et al. 2022). An important foundation of this research is a robust prediction of proteins as effectors. Up until recently, except for a small number of model species, defining the effector repertoire of a pathogen of interest has remained challenging. However, recent breakthroughs in computational and technological tools, as discussed here, are expected to revolutionize effector identification.

A breakthrough that has been witnessed is template-free protein structural modeling, which will switch effector prediction from sequence-based to structure-based searches. Genome-wide structural modeling of putative secreted proteins in diverse, well-studied pathogens will reveal enriched folds, which will, in turn, facilitate effector prediction. These folds could be related to protein secretion, chaperone association, and on a larger scale, virulence activities. Furthermore, the protein structure information will feed into AI-based methods that can combine information reflecting genome and transcriptome features to further increase the robustness of effector prediction.

Another technical advancement is single cell-based transcriptomics. Infection-specific expression can provide guidance to effector identification and candidate selection for further functional characterizations. However, the RNA content of the pathogen is much lower compared to the host, especially at the critical early infection stages that establish the success or failure of colonization. Furthermore, gene expression changes may be masked in bulk transcriptomic analysis in which pathogens cells have heterogeneity in terms of developmental stage and host interaction. Single cell-based methods can significantly improve this analysis by focusing on pathogen cells that are in direct contact with the host. Clustering analysis based on gene expression profiles will also facilitate the establishment of co-expression patterns in particular cell types. Moreover, newer sequence-based imaging methods (e.g., PhytoMap) hold great promise to impart spatial information to transcriptomic data (Cole et al. 2021; Nobori et al. 2022). With these new methods and technologies, effector research is expected to enter an exciting new stage.

Acknowledgments

We thank Thorsten Langner for valuable comments.

Literature Cited

Abby, S. S., Cury, J., Guglielmini, J., Neron, B., Touchon, M., and Rocha, E. P. 2016. Identification of protein secretion systems in bacterial genomes. Sci. Rep. 6:23080.

Alfano, J. R., Charkowski, A. O., Deng, W. L., Badel, J. L., Petnicki-Ocwieja, T., van Dijk, K., and Collmer, A. 2000. The *Pseudomonas syringae* Hrp pathogenicity island has a tripartite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in plants. Proc. Natl. Acad. Sci. U.S.A. 25:4856-4861.

Almagro Armenteros, J. J., Salvatore, M., Emanuelsson, O., Winther, O., von Heijne, G., Elofsson, A., and Nielsen, H. 2019. Detecting sequence signals in targeting peptides using deep learning. Life Sci. Alliance 2:e201900429.

Amaro, T. M. M. M., Thilliez, G. J. A., Motion, G. B., and Huitema, E. 2017. A perspective on CRN proteins in the genomics age: Evolution, classification, delivery and function revisited. Front. Plant Sci. 8:99.

Angot, A., Peeters, N., Lechner, E., Vailleau, F., Baud, C., Gentzbittel, L., Sartorel, E., Genschik, P., Boucher, C., and Genin, S. 2006. *Ralstonia solanacearum* requires F-box-like domain-containing type III effectors to promote disease on several host plants. Proc. Natl. Acad. Sci. U.S.A. 103:14620-14625.

Apweiler, R., Attwood, T., Bairoch, A., Bateman, A., Birney, E., Biswas, M., Bucher, P., Cerutti, L., Corpet, F., Croning, M., Durbin, R., Falquet, L., Fleischmann, W., Gouzy, J., Hermjakob, H., Hulo, N., Jonassen, I., Kahn, D., Kanapin, A., Karavidopoulou, Y., Lopez, R., Marx, B., Mulder, N., Oinn, T., Pagni, M., Servant, F., Sigrist, C., and Zdobnov, E. 2001. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. Nucleic Acids Res. 29:37-40.

Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., and Wishart, D. S. 2016. PHASTER: A better, faster version of the PHAST phage search tool. Nucleic Acids Res. 44:W16-W21.

Arnold, R., Brandmaier, S., Kleine, F., Tischler, P., Heinz, E., Behrens, S., Niinikoski, A., Mewes, H. W., Horn, M., and Rattei, T. 2009. Sequence-based prediction of type III secreted proteins. PLoS Pathog. 5:e1000376.

Badouin, H., Gladieux, P., Gouzy, J., Siguenza, S., Aguileta, G., Snirc, A., Le Prieur, S., Jeziorski, C., Branca, A., and Giraud, T. 2017. Widespread selective sweeps throughout the genome of model plant pathogenic fungi and identification of effector candidates. Mol. Ecol. 26:2041-2062.

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G., Wang, J., Cong, Q., Kinch, L., Schaeffer, R., Millán, C., Park, H., C., A., Glassman, C., DeGiovanni, A., Pereira, J., Rodrigues, A., van Dijk, A., Ebrecht, A., Opperman, D., Sagmeister, T., Buhlheller, C., Pavkov-Keller, T., Rathinaswamy, M., Dalwadi, U., Yip, C., Burke, J., Garcia, K., Grishin, N., Adams, P., Read, R., and Baker, D. 2021. Accurate prediction of protein structures and interactions using a three-track neural network. Science 373:871-876.

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W., and Noble, W. S. 2009. MEME SUITE: Tools for motif discovery and searching. Nucleic Acids Res. 37:W202-W208.

Bao, J., Chen, M., Zhong, Z., Tang, W., Lin, L., Zhang, X., Jiang, H., Zhang, D., Miao, C., Tang, H., Zhang, J., Lu, G., Ming, R., Norvienyeku, J., Wang, B., and Wang, Z. 2017. PacBio sequencing reveals transposable elements as a key contributor to genomic plasticity and virulence variation in *Magnaporthe oryzae*. Mol. Plant 10:1465-1468.

Bertelli, C., Laird, M. R., Williams, K. P., Simon Fraser University Research Computing Group, Lau, B. Y., Hoad, G., Winsor, G. L., and Brinkman, F. S. L. 2017. IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Res. 45:W30-W35.

Besser, J., Carleton, H. A., Gerner-Smidt, P., Lindsey, R. L., and Trees, E. 2018. Next-generation sequencing technologies and their application to the study and control of bacterial infections. Clin. Microbiol. Infect. 24:335-341.

Blondeau, K., Blaise, F., Graille, M., Kale, S. D., Linglin, J., Ollivier, B., Labarde, A., Lazar, N., Daverdin, G., Balesdent, M. H., Choi, D. H., Tyler, B. M., Rouxel, T., van Tilbeurgh, H., and Fudal, I. 2015. Crystal structure of the effector AvrLm4-7 of *Leptosphaeria maculans* reveals insights into its translocation into plant cells and recognition by resistance proteins. Plant J. 83:610-624.

Boch, J., Joardar, V., Gao, L., Robertson, T., Lim, M., and Kunkel, B. 2002. Identification of *Pseudomonas syringae* pv. *tomato* genes induced during infection of *Arabidopsis thaliana*. Mol. Microbiol. 44:73-88.

Boutemy, L. S., King, S. R. F., Win, J., Hughes, R. K., Clarke, T. A., Blumenschein, T. M. A., Kamoun, S., and Banfield, M. J. 2011. Structures

of *Phytophthora* RXLR effector proteins: A conserved but adaptable fold underpins functional diversity. J. Biol. Chem. 286:35834-35842.

Browne, W., North, A., Phillips, D., Brew, K., Vanaman, T., and Hill, R. 1969. A possible three-dimensional structure of bovine alpha-lactalbumin based on that of hen's egg-white lysozyme. J. Mol. Biol. 42:65-86.

Buchko, G. W., Niemann, G., Baker, E. S., Belov, M. E., Smith, R. D., Heffron, F., Adkins, J. N., and McDermott, J. E. 2010. A multi-pronged search for a common structural motif in the secretion signal of *Salmonella enterica* serovar Typhimurium type III effector proteins. Mol. Biosyst. 6:2448-2458.

Chang, J. H., Urbach, J. M., Law, T. F., Arnold, L. W., Hu, A., Gombar, S., Grant, S. R., Ausubel, F. M., and Dangl, J. L. 2005. A high-throughput, near-saturating screen for type III effector genes from *Pseudomonas syringae.* Proc. Natl. Acad. Sci. U.S.A. 102:2549-2554.

Chen, Y. F., and Xia, Y. 2021. Structural profiling of bacterial effectors reveals enrichment of host-interacting domains and motifs. Front. Mol. Biosci. 8:626600.

Cole, B., Bergmann, D., Blaby-Haas, C. E., Blaby, I. K., Bouchard, K. E., Brady, S. M., Ciobanu, D., Coleman-Derr, D., Leiboff, S., Mortimer, J. C., Nobori, T., Rhee, S. Y., Schmutz, J., Simmons, B. A., Singh, A. K., Sinha, N., Vogel, J. P., O'Malley, R. C., Visel, A., and Dickel, D. E. 2021. Plant single-cell solutions for energy and the environment. Commun. Biol. 4:962.

Costa, S. C., Schmitz, A. M., Jahufar, F. F., Boyd, J. D., Cho, M. Y., Glicksman, M. A., and Lesser, C. F. 2012. A new means to identify type 3 secreted effectors: Functionally interchangeable class IB chaperones recognize a conserved sequence. mBio 3:e00243-11.

Cornelis, G. R. 2006. The type II secretion injectisome. Nat. Rev. Microbiol. 4:811-825.

Dangl, J. L., Ritter, C., Gibbon, M. J., Mur, L. A., Wood, J. R., Goss, S., Mansfield, J., Taylor, J. D., and Vivian, A. 1992. Functional homologs of the Arabidopsis RPM1 disease resistance gene in bean and pea. Plant Cell 4:1359-1369.

De Francesco, A., Lovelace, A. H., Shaw, D., Qiu, M., Wang, Y., Gurung, F., Ancona, V., Wang, C., Levy, A., Jiang, T., and Ma, W. 2022. Transcriptome profiling of '*Candidatus* Liberibacter asiaticus' in citrus and psyllids. Phytopathology 112:116-130.

de Guillen, K., Ortiz-Vallejo, D., Gracy, J., Fournier, E., Kroj, T., and Padilla, A. 2015. Structure analysis uncovers a highly diverse but structurally conserved effector family in phytopathogenic fungi. PLoS Pathog. 11:e1005228.

Derbyshire, M. C. 2020. Bioinformatic detection of positive selection pressure in plant pathogens: The neutral theory of molecular sequence evolution in action. Front. Microbiol. 11:644.

Deslandes, L., Olivier, J., Peeters, N., Feng, D. X., Khounlotham, M., Boucher, C., Somssich, I., Genin, S., and Marco, Y. 2003. Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. Proc. Natl. Acad. Sci. U.S.A. 100: 8024-8029.

Dhroso, A., Eidson, S., and Korkin, D. 2018. Genome-wide prediction of bacterial effector candidates across six secretion system types using a feature-based statistical framework. Sci. Rep. 8:17209.

Dillon, M. M., Thakur, S., Almeida, R. N. D., Wang, P. W., Weir, B. S., and Guttman, D. S. 2019. Recombination of ecologically and evolutionarily significant loci maintains genetic cohesion in the *Pseudomonas syringae* species complex. Genome Biol. 20:3.

Dong, S., and Ma, W. 2021. How to win a tug-of-war: The adaptive evolution of *Phytophthora* effectors. Curr. Opin. Plant Biol. 62:102027.

Dong, S., Raffaele, S., and Kamoun, S. 2015. The two-speed genomes of filamentous pathogens: Waltz with plants. Curr. Opin. Genet. Dev. 35:57-65.

Dunker, A., Obradovic, Z., Romero, P., Garner, E., and Brown, C. 2000. Intrinsic protein disorder in complete genomes. Genome Inform. Ser. Workshop Genome Inform. 11:161-171.

Durrant, M. G., Li, M. M., Siranosian, B. A., Montgomery, S. B., and Bhatt, A. S. 2020. A bioinformatic analysis of integrative mobile genetic elements highlights their role in bacterial adaptation. Cell Host Microbe 27:140-153.e9.

Eddy, S. 2009. New generation of homology search tools based on probabilistic inference. Genome Inform. 23:205-211.

Evangelisti, E., Gogleva, A., Hainaux, T., Doumane, M., Tulin, F., Quan, C., Yunusov, T., Floch, K., and Schornack, S. 2017. Time-resolved dual transcriptomics reveal early induced *Nicotiana benthamiana* root genes and conserved infection-promoting *Phytophthora palmivora* effectors. BMC Biol. 15:39.

Faino, L., Seidl, M. F., Shi-Kunne, X., Pauper, M., van den Berg, G. C., Wittenberg, A. H., and Thomma, B. P. 2016. Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen. Genome Res. 26:1091-1100.

Flutre, T., Duprat, E., Feuillet, C., and Quesneville, H. 2011. Considering transposable element diversification in de novo annotation approaches. PloS One 31:e16526.

Fosu-Nyarko, J., Jones, M., and Wang, Z. 2010. Application of laser microdissection to study plant–fungal pathogen interactions. Pages 153-164 in: Molecular and Cell Biology Methods for Fungi. A. Sharon, ed. Springer Protocols, New York.

Furutani, A., Nakayama, T., Ochiai, H., Kaku, H., Kubo, Y., and Tsuge, S. 2006. Identification of novel HrpXo regulons preceded by two cis-acting elements, a plant-inducible promoter box and a -10 box-like sequence, from the genome database of *Xanthomonas oryzae* pv. *oryzae*. FEMS Microbiol. Lett. 259:133-141.

Godfrey, D., Böhlenius, H., Pedersen, C., Zhang, Z., Emmersen, J., and Thordal-Christensen, H. 2010. Powdery mildew fungal effector candidates share N-terminal Y/F/WxC-motif. BMC Genom. 11:317.

Goodwin, S. B., M'Barek, S., B., D., B., W., A., H., Crane, C. F., Hane, J. K., Foster, A. J., Van der Lee, T. A., Grimwood, J., Aerts, A., Antoniw, J., Bailey, A., Bluhm, B., Bowler, J., Bristow, J., van der Burgt, A., Canto-Canche, B., Churchill, A. C., Conde-Ferraez, L., Cools, H. J., Coutinho, P. M., Csukai, M., Dehal, P., De Wit, P., Donzelli, B., van de Geest, H. C., van Ham, R. C., Hammond-Kosack, K. E., Henrissat, B., Kilian, A., Kobayashi, A. K., Koopmann, E., Kourmpetis, Y., Kuzniar, A., Lindquist, E., Lombard, V., Maliepaard, C., Martins, N., Mehrabi, R., Nap, J. P., Ponomarenko, A., Rudd, J. J., Salamov, A., Schmutz, J., Schouten, H. J., Shapiro, H., Stergiopoulos, I., Torriani, S. F., Tu, H., de Vries, R. P., Waalwijk, C., Ware, S. B., Wiebenga, A., Zwiers, L. H., Oliver, R. P., Grigoriev, I. V., and Kema, G. H. 2011. Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. PLos Genet. 7:e1002070.

Grandaubert, J., Lowe, R., Soyer, J., Schoch, C., Van de Wouw, A., Fudal, I., Robbertse, B., Lapalu, N., Links, M., Ollivier, B., Linglin, J., Barbe, V., Mangenot, S., Cruaud, C., Borhan, H., Howlett, B., Balesdent, M.-H., and Rouxel, T. 2014. Transposable element-assisted evolution and adaptation to host plant within the *Leptosphaeria maculans*−*Leptosphaeria biglobosa* species complex of fungal pathogens. BMC Genom. 15:891.

Green, E. R., and Mecsas, J. 2016. Bacterial secretion systems: An overview. Microbiol. Spectr. 4.

Guttman, D., Vinatzer, B., Sarkar, S., Ranall, M., Kettler, G., and Greenberg, J. 2002. A functional screen for the Type III (Hrp) secretome of the plant pathogen *Pseudomonas syringae*. Science 295:1722-1726.

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H., Handsaker, R. E., Cano, L. M., Grabherr, M., Kodira, C. D., Raffaele, S., Torto-Alalibo, T., Bozkurt, T. O., Ah-Fong, A. M., Alvarado, L., Anderson, V. L., Armstrong, M. R., Avrova, A., Baxter, L., Beynon, J., Boevink, P. C., Bollmann, S. R., Bos, J. I., Bulone, V., Cai, G., Cakir, C., Carrington, J. C., Chawner, M., Conti, L., Costanzo, S., Ewan, R., Fahlgren, N., Fischbach, M. A., Fugelstad, J., Gilroy, E. M., Gnerre, S., Green, P. J., Grenville-Briggs, L. J., Griffith, J., Grunwald, N. J., Horn, K., Horner, N. R., Hu, C. H., Huitema, E., Jeong, D. H., Jones, A. M., Jones, J. D., Jones, R. W., Karlsson, E. K., Kunjeti, S. G., Lamour, K., Liu, Z., Ma, L., Maclean, D., Chibucos, M. C., McDonald, H., McWalters, J., Meijer, H. J., Morgan, W., Morris, P. F., Munro, C. A., O'Neill, K., Ospina-Giraldo, M., Pinzon, A., Pritchard, L., Ramsahoye, B., Ren, Q., Restrepo, S., Roy, S., Sadanandom, A., Savidor, A., Schornack, S., Schwartz, D. C., Schumann, U. D., Schwessinger, B., Seyer, L., Sharpe, T., Silvar, C., Song, J., Studholme, D. J., Sykes, S., Thines, M., van de Vondervoort, P. J., Phuntumart, V., Wawra, S., Weide, R., Win, J., Young, C., Zhou, S., Fry, W., Meyers, B. C., van West, P., Ristaino, J., Govers, F., Birch, P. R., Whisson, S. C., Judelson, H. S., and Nusbaum, C. 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature 461:393-398.

Han, Y., Liu, X., Benny, U., Kistler, H., and VanEtten, H. 2001. Genes determining pathogenicity to pea are clustered on a supernumerary chromosome in the fungal plant pathogen *Nectria haematococca*. Plant J. 25: 305-314.

Hartmann, F. E., McDonald, B. A., and Croll, D. 2018. Genome-wide evidence for divergent selection between populations of a major agricultural pathogen. Mol. Ecol. 27:2725-2741.

He, J., Ye, W., Choi, D. S., Wu, B., Zhai, Y., Guo, B., Duan, S., Wang, Y., Gan, J., Ma, W., and Ma, J. 2019. Structural analysis of *Phytophthora* suppressor of RNA silencing 2 (PSR2) reveals a conserved modular fold contributing to virulence. Proc. Natl. Acad. Sci. U.S.A. 116:8054-8059.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., and Glass, C. K. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell 38:576-589.

Henry, E., Toruno, T. Y., Jauneau, A., Deslandes, L., and Coaker, G. 2017. Direct and indirect visualization of bacterial effector delivery into diverse plant cell types during infection. Plant Cell 29:1555-1570.

Hicks, S. W., and Galan, J. E. 2013. Exploitation of eukaryotic subcellular targeting mechanisms by bacterial effectors. Nat. Rev. Microbiol. 11: 316-326.

Hiller, K., Grote, A., Scheer, M., Munch, R., and Jahn, D. 2004. PrediSi: Prediction of signal peptides and their cleavage positions. Nucleic Acids Res. 32:W375-W379.

Hogenhout, S., Van der Hoorn, R., Terauchi, R., and Kamoun, S. 2009. Emerging concepts in effector biology of plant-associated organisms. Mol. Plant-Microbe Interact. 22:115-122.
Hu, G., Linning, R., McCallum, B., Banks, T., Cloutier, S., Butterfield, Y., Liu, J., Kirkpatrick, R., Stott, J., Yang, G., Smailus, D., Jones, S., Marra, M., Schein, J., and Bakkeren, G. 2007. Generation of a wheat leaf rust, *Puccinia triticina*, EST database from stage-specific cDNA libraries. Mol. Plant Pathol. 8:451-467.
Hui, X., Chen, Z., Lin, M., Zhang, J., Hu, Y., Zeng, Y., Cheng, X., Ou-Yang, L., Sun, M. A., White, A. P., and Wang, Y. 2020. T3SEpp: An integrated prediction pipeline for bacterial Type III secreted effectors. mSystems 5:e00288-20.
Jackson, R. W., Athanassopoulos, E., Tsiamis, G., Mansfield, J. W., Sesma, A., Arnold, D. L., Gibbon, M. J., Murillo, J., Taylor, J. D., and Vivian, A. 1999. Identification of a pathogenicity island, which contains genes for virulence and avirulence, on a large native plasmid in the bean pathogen *Pseudomonas syringae* pathovar *phaseolicola*. Proc. Natl. Acad. Sci. U.S.A. 96:10875-10880.
Jackson, R. W., Vinatzer, B., Arnold, D. L., Dorus, S., and Murillo, J. 2011. The influence of the accessory genome on bacterial pathogen evolution. Mob. Genet. Elements 1:55-65.
Janjusevic, R., Abramovitch, R., Martin, G., and Stebbins, C. 2006. A bacterial inhibitor of host programmed cell death defenses is an E3 ubiquitin ligase. Science 311:222-226.
Jelenska, J., Yao, N., Vinatzer, B., Wright, C., Brodsky, J., and Greenberg, J. 2007. A J domain virulence effector of *Pseudomonas syringae* remodels host chloroplasts and suppresses defenses. Curr. Biol. 17:499-508.
Jiang, R. H., Tripathy, S., Govers, F., and Tyler, B. M. 2008. RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving superfamily with more than 700 members. Proc. Natl. Acad. Sci. U.S.A. 105:4874-4879.
Jones, D. A. B., John, E., Rybak, K., Phan, H. T. T., Singh, K. B., Lin, S. Y., Solomon, P. S., Oliver, R. P., and Tan, K. C. 2019. A specific fungal transcription factor controls effector gene expression and orchestrates the establishment of the necrotrophic pathogen lifestyle on wheat. Sci. Rep. 9:15884.
Jones, D. A. B., Rybak, K., Bertazzoni, S., Tan, K.-C., Phan, H. T. T., and Hane, J. K. 2021. Pathogenicity effector candidates and accessory genome revealed by pan-genomic analysis of *Parastagonospora nodorum*. bioRxiv 2021.09.01.458590.
Jones, J. D., and Dangl, J. L. 2006. The plant immune system. Nature 444: 323-329.
Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P., and Hassabis, D. 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596:583-589.
Kall, L., Krogh, A., and Sonnhammer, E. L. 2007. Advantages of combined transmembrane topology and signal peptide prediction—The Phobius web server. Nucleic Acids Res. 35:W429-W432.
Kämper, J., Kahmann, R., Bolker, M., Ma, L. J., Brefort, T., Saville, B. J., Banuett, F., Kronstad, J. W., Gold, S. E., Muller, O., Perlin, M. H., Wosten, H. A., de Vries, R., Ruiz-Herrera, J., Reynaga-Pena, C. G., Snetselaar, K., McCann, M., Perez-Martin, J., Feldbrugge, M., Basse, C. W., Steinberg, G., Ibeas, J. I., Holloman, W., Guzman, P., Farman, M., Stajich, J. E., Sentandreu, R., Gonzalez-Prieto, J. M., Kennell, J. C., Molina, L., Schirawski, J., Mendoza-Mendoza, A., Greilinger, D., Munch, K., Rossel, N., Scherer, M., Vranes, M., Ladendorf, O., Vincon, V., Fuchs, U., Sandrock, B., Meng, S., Ho, E. C., Cahill, M. J., Boyce, K. J., Klose, J., Klosterman, S. J., Deelstra, H. J., Ortiz-Castellanos, L., Li, W., Sanchez-Alonso, P., Schreier, P. H., Hauser-Hahn, I., Vaupel, M., Koopmann, E., Friedrich, G., Voss, H., Schluter, T., Margolis, J., Platt, D., Swimmer, C., Gnirke, A., Chen, F., Vysotskaia, V., Mannhaupt, G., Guldener, U., Munsterkotter, M., Haase, D., Oesterheld, M., Mewes, H. W., Mauceli, E. W., DeCaprio, D., Wade, C. M., Butler, J., Young, S., Jaffe, D. B., Calvo, S., Nusbaum, C., Galagan, J., and Birren, B. W. 2006. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. Nature 444:97-101.
Keller, N., and Hohn, T. 1997. Metabolic pathway gene clusters in filamentous fungi. Fungal Genet. Biol. 21:17-29.
Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. 10:845-858.
Kemen, E., Kemen, A., Rafiqi, M., Hempel, U., Mendgen, K., Hahn, M., and Voegele, R. 2005. Identification of a protein from rust fungi transferred from haustoria into infected plant cells. Mol. Plant-Microbe Interact. 18:1130-1139.
Khang, C. H., Berruyer, R., Giraldo, M. C., Kankanala, P., Park, S. Y., Czymmek, K., Kang, S., and Valent, B. 2010. Translocation of *Magnaporthe oryzae* effectors into rice cells and their subsequent cell-to-cell movement. Plant Cell 22:1388-1403.
King, S. R., McLellan, H., Boevink, P. C., Armstrong, M. R., Bukharova, T., Sukarta, O., Win, J., Kamoun, S., Birch, P. R., and Banfield, M. J. 2014. *Phytophthora infestans* RXLR effector PexRD2 interacts with host MAPKKK epsilon to suppress plant immune signaling. Plant Cell 26:1345-1359.
Kristianingsih, R., and MacLean, D. 2021. Accurate plant pathogen effector protein classification ab initio with deepredeff: An ensemble of convolutional neural networks. BMC Bioinf. 22:372.
Kuhlman, B., and Bradley, P. 2019. Advances in protein structure prediction and design. Nat. Rev. Mol. Cell Biol. 20:681-697.
Lazar, N., Mesarich, C. H., Petit-Houdenot, Y., Talbi, N., Li de la Sierra-Gallay, I., Zelie, E., Blondeau, K., Gracy, J., Ollivier, B., Blaise, F., Rouxel, T., Balesdent, M. H., Idnurm, A., van Tilbeurgh, H., and Fudal, I. 2022. A new family of structurally conserved fungal effectors displays epistatic interactions with plant resistance proteins. PLoS Pathog. 18:e1010664.
LeBlanc, M. A., Fink, M. R., Perkins, T. T., and Sousa, M. C. 2021. Type III secretion system effector proteins are mechanically labile. Proc. Natl. Acad. Sci. U.S.A. 118:e2019566118.
Li, G., Froehlich, J. E., Elowsky, C., Msanne, J., Ostosh, A. C., Zhang, C., Awada, T., and Alfano, J. R. 2014. Distinct *Pseudomonas* type-III effectors use a cleavable transit peptide to target chloroplasts. Plant J. 77:310-321.
Lin, X., Jia, Y., Heal, R., Prokchorchik, M., Sindalovskaya, M., Olave-Achury, A., Makechemu, M., Fairhead, S., Noureen, A., Heo, J., Witek, K., Smoker, M., Taylor, J., Shrestha, R. - K., Lee, Y., Zhang, C., Park, S. J., Sohn, K. H., Huang, S., and Jones, J. D. G. 2022. The *Solanum americanum* pangenome and effectoromics reveal new resistance genes against potato late blight. bioRxiv 2022.08.11.503608.
Lo Presti, L., Zechmann, B., Kumlehn, J., Liang, L., Lanver, D., Tanaka, S., Bock, R., and Kahmann, R. 2017. An assay for entry of secreted fungal effectors into plant cells. New Phytol. 213:956-964.
Lovelace, A. H., and Ma, W. 2022. How do bacteria transform plants into their oasis? Cell Host Microbe 30:412-414.
Lovelace, A. H., Smith, A., and Kvitko, B. H. 2018. Pattern-triggered immunity alters the transcriptional regulation of virulence-associated genes and induces the sulfur starvation response in *Pseudomonas syringae* pv. *tomato* DC3000. Mol. Plant-Microbe Interact. 31:750-765.
Ma, K. W., and Ma, W. 2016. YopJ family effectors promote bacterial infection through a unique acetyltransferase activity. Microbiol. Mol. Biol. Rev. 80:1011-1027.
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., Henrissat, B., Houterman, P. M., Kang, S., Shim, W. B., Woloshuk, C., Xie, X., Xu, J. R., Antoniw, J., Baker, S. E., Bluhm, B. H., Breakspear, A., Brown, D. W., Butchko, R. A., Chapman, S., Coulson, R., Coutinho, P. M., Danchin, E. G., Diener, A., Gale, L. R., Gardiner, D. M., Goff, S., Hammond-Kosack, K. E., Hilburn, K., Hua-Van, A., Jonkers, W., Kazan, K., Kodira, C. D., Koehrsen, M., Kumar, L., Lee, Y. H., Li, L., Manners, J. M., Miranda-Saavedra, D., Mukherjee, M., Park, G., Park, J., Park, S. Y., Proctor, R. H., Regev, A., Ruiz-Roldan, M. C., Sain, D., Sakthikumar, S., Sykes, S., Schwartz, D. C., Turgeon, B. G., Wapinski, I., Yoder, O., Young, S., Zeng, Q., Zhou, S., Galagan, J., Cuomo, C. A., Kistler, H. C., and Rep, M. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature 464: 367-373.
Ma, W., Dong, F. F., Stavrinides, J., and Guttman, D. S. 2006. Type III effector diversification via both pathoadaptation and horizontal transfer in response to a coevolutionary arms race. PLos Genet. 2:e209.
Mak, A. N., Bradley, P., Bogdanove, A. J., and Stoddard, B. L. 2013. TAL effectors: Function, structure, engineering and applications. Curr. Opin. Struct. Biol. 23:93-99.
Mansfield, J. W. 2009 From bacterial avirulence genes to effector functions via the hrp delivery system: An overview of 25 years of progress in our understanding of plant innate immunity. Mol. Plant Pathol. 10:721-734.
Maqbool, A., Hughes, R. K., Dagdas, Y. F., Tregidgo, N., Zess, E., Belhaj, K., Round, A., Bozkurt, T. O., Kamoun, S., and Banfield, M. J. 2016. Structural basis of host autophagy-related protein 8 (ATG8) binding by the Irish potato famine pathogen effector protein PexRD54. J. Biol. Chem. 291:20270-20282.
McCann, H. C., and Guttman, D. S. 2008. Evolution of the type III secretion system and its effectors in plant−microbe interactions. New Phytol. 177: 33-47.
Ménard, R., Sansonetti, P. J., and Parsot, C. 1993. Nonpolar mutagenesis of the ipa genes defines IpaB, IpaC, and IpaD as effectors of *Shigella flexneri* entry into epithelial cells. J Bacteriol. 175:5899-5906.
Menardo, F., Praz, C. R., Wicker, T., and Keller, B. 2017. Rapid turnover of effectors in grass powdery mildew (*Blumeria graminis*). BMC Evol. Biol. 17:223.

Mészáros, B., Erdos, G., and Dosztanyi, Z. 2018. IUPred2A: Context-dependent prediction of protein disorder as a function of redox state and protein binding. Nucleic Acids Res. 46:W329-W337.

Miao, V., Covert, S., and VanEtten, H. 1991. A fungal gene for antibiotic resistance on a dispensable ("B") chromosome. Science 254:1773-1776.

Mosquera, G., Giraldo, M. C., Khang, C. H., Coughlan, S., and Valent, B. 2009. Interaction transcriptome analysis identifies *Magnaporthe oryzae* BAS1-4 as biotrophy-associated secreted proteins in rice blast disease. Plant Cell 21:1273-1290.

Mukhi, N., Gorenkin, D., and Banfield, M. J. 2020. Exploring folds, evolution and host interactions: Understanding effector structure/function in disease and immunity. New Phytol. 227:326-333.

Napoli, C., and Staskawicz, B. 1987. Molecular characterization and nucleic acid sequence of an avirulence gene from race 6 of *Pseudomonas syringae* pv. *glycinea*. J Bacteriol. 169:572-578.

Natale, P., Bruser, T., and Driessen, A. J. 2008. Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane-distinct translocases and mechanisms. Biochim. Biophys. Acta 1778:1735-1756.

Necci, M., Piovesan, D., Predictors, C., DisProt, C., and Tosatto, S. C. E. 2021. Critical assessment of protein intrinsic disorder prediction. Nat. Methods 18:472-481.

Nielsen, H., Engelbrecht, J., von Heijne, G., and Brunak, S. 1996. Defining a similarity threshold for a functional protein sequence pattern: The signal peptide cleavage site. Proteins 24:165-177.

Nobori, T., Oliva, M., Lister, R., and Ecker, J. R. 2022. PHYTOMap: Multiplexed single-cell 3D spatial gene expression analysis in plant tissue. bioRxiv 2022.07.28.501915.

Nobori, T., Velasquez, A. C., Wu, J., Kvitko, B. H., Kremer, J. M., Wang, Y., He, S. Y., and Tsuda, K. 2018. Transcriptome landscape of a bacterial pathogen under plant immunity. Proc. Natl. Acad. Sci. U.S.A. 115:E3055-E3064.

Nur, M., Wood, K., and Michelmore, R. 2021. EffectorO: Motif-independent prediction of effectors in oomycete genomes using machine learning and lineage specificity. bioRxiv 2021.03.19.436227.

O'Malley, M. R., and Anderson, J. C. 2021. Regulation of the *Pseudomonas syringae* Type III secretion system by host environment signals. Microorganisms. 9:1227.

Outram, M. A., Figueroa, M., Sperschneider, J., Williams, S. J., and Dodds, P. N. 2022. Seeing is believing: Exploiting advances in structural biology to understand and engineer plant immunity. Curr. Opin. Plant Biol. 67:102210.

Pederson, C., Ver Loren van Themaat, E., McGuffin, L. J., Abbot, J. C., Burgis, T. A., Barton, G., Bindschedler, L. V., Lu, X., Maekawa, T., Weßling, R., Cramer, R., Thordal-Christensen, H., Panstruga, R., and Spanu, P. D. 2012. Structure and evolution of barley powdery mildew effector candidates. BMC Genom. 13:694.

Peng, Z., Oliveira-Garcia, E., Lin, G., Hu, Y., Dalby, M., Migeon, P., Tang, H., Farman, M., Cook, D., White, F. F., Valent, B., and Liu, S. 2019. Effector gene reshuffling involves dispensable mini-chromosomes in the wheat blast fungus. PLos Genet. 15:e1008272.

Pennington, H. G., Jones, R., Kwon, S., Bonciani, G., Thieron, H., Chandler, T., Luong, P., Morgan, S. N., Przydacz, M., Bozkurt, T., Bowden, S., Craze, M., Wallington, E. J., Garnett, J., Kwaaitaal, M., Panstruga, R., Cota, E., and Spanu, P. D. 2019. The fungal ribonuclease-like effector protein CSEP0064/BEC1054 represses plant immunity and interferes with degradation of host ribosomal RNA. PLoS Pathog. 15:e1007620.

Petre, B., and Kamoun, S. 2014. How do filamentous pathogens deliver effector proteins into plant cells? PLoS Biol. 12:e1001801.

Quesneville, H., Bergman, C., Andrieu, O., Autard, D., Nouaud, D., Ashburner, M., and Anxolabehere, D. 2005. Combined Evidence Annotation of Transposable Elements in Genome Sequences. PLoS Com. Biol. 1:e22.

Raffaele, S., Farrer, R., Cano, L., Studholme, D., MacLean, D., Thines, M., Jiang, R., Zody, M., Kunjeti, S., Donofrio, N., Meyers, B., Nusbaum, C., and Kamoun, S. 2010. Genome evolution following host jumps in the Irish potato famine pathogen lineage. Science 330:1540-1543.

Rehmany, A. P., Gordon, A., Rose, L. E., Allen, R. L., Armstrong, M. R., Whisson, S. C., Kamoun, S., Tyler, B. M., Birch, P. R., and Beynon, J. L. 2005. Differential recognition of highly divergent downy mildew avirulence gene alleles by RPP1 resistance genes from two *Arabidopsis* lines. Plant Cell 17:1839-1850.

Richards, J. K., Stukenbrock, E. H., Carpenter, J., Liu, Z., Cowger, C., Faris, J. D., and Friesen, T. L. 2019. Local adaptation drives the diversification of effectors in the fungal wheat pathogen *Parastagonospora nodorum* in the United States. PLos Genet. 15:e1008223.

Rouxel, T., Grandaubert, J., Hane, J. K., Hoede, C., van de Wouw, A. P., Couloux, A., Dominguez, V., Anthouard, V., Bally, P., Bourras, S., Cozijnsen, A. J., Ciuffetti, L. M., Degrave, A., Dilmaghani, A., Duret, L., Fudal, I., Goodwin, S. B., Gout, L., Glaser, N., Linglin, J., Kema, G. H., Lapalu, N., Lawrence, C. B., May, K., Meyer, M., Ollivier, B., Poulain, J., Schoch, C. L., Simon, A., Spatafora, J. W., Stachowiak, A., Turgeon, B. G., Tyler, B. M., Vincent, D., Weissenbach, J., Amselem, J., Quesneville, H., Oliver, R. P., Wincker, P., Balesdent, M. H., and Howlett, B. J. 2011. Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by repeat-induced point mutations. Nat. Commun. 2:202.

Roy, A., Kucukural, A., and Zhang, Y. 2010. I-TASSER: A unified platform for automated protein structure and function prediction. Nat. Protoc. 5: 725-738.

Sabbagh, C. R. R., Carrere, S., Lonjon, F., Vailleau, F., Macho, A. P., Genin, S., and Peeters, N. 2019. Pangenomic type III effector database of the plant pathogenic *Ralstonia* spp. PeerJ 7:e7346.

Šali, A., and Blundell, T. 1993. Comparative protein modeling by satisfaction of spatial restraints. J. Mol. Biol. 234:779-815.

Saunders, D. G., Win, J., Cano, L. M., Szabo, L. J., Kamoun, S., and Raffaele, S. 2012. Using hierarchical clustering of secreted protein families to classify and rank candidate effectors of rust fungi. PLoS One 7:e29847.

Schechter, L. M., Roberts, K. A., Jamir, Y., Alfano, J. R., and Collmer, A. 2004. *Pseudomonas syringae* type III secretion system targeting signals and novel effectors studied with a Cya translocation reporter. J. Bacteriol. 186: 543-555.

Schmidt, S. M., Houterman, P. M., Schreiver, I., Ma, L., Amyotte, S., Chellappan, B., Boeren, S., Takken, F. L. W., and Rep, M. 2013. MITEs in the promoter of effector genes allow prediction of novel virulence genes in *Fusarium oxysporum*. BMC Genom. 14:119.

Schotanus, K., Soyer, J. L., Connolly, L. R., Grandaubert, J., Happel, P., Smith, K. M., Freitag, M., and Stukenbrock, E. H. 2015. Histone modifications rather than the novel regional centromeres of *Zymospetoria tritici* distinguish core and accessory genomes. Epigenetics Chromatin 8:41.

Seong, K., and Krasileva, K. V. 2021. Computational structural genomics unravels common folds and novel families in the secretome of fungal phytopathogen *Magnaporthe oryzae*. Mol. Plant-Microbe Interact. 34: 1267-1280.

Siani, R., Stabl, G., Gutjahr, C., Schloter, M., and Radl, V. 2021. Acidovorax pangenome reveals specific functional traits for plant beneficial and pathogenic plant-associations. Microbiol. Genom. 7:000666.

Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., and Chandler, M. 2006. ISfinder: The reference centre for bacterial insertion sequences. Nucleic Acids Res. 34:D32-D36.

Singh, J., and Khan, A. 2019. Distinct patterns of natural selection determine sub-population structure in the fire blight pathogen, *Erwinia amylovora*. Sci. Rep. 9:14017.

Smit, A. F. A., Hubley, R., and Green, P. 2013−2015. RepeatMasker Open-4.0. http://www.repeatmasker.org

Smits, T. H. M. 2019. The importance of genome sequence quality to microbial comparative genomics. BMC Genom. 20:662.

Söding, J., Biegert, A., and Lupas, A. N. 2005. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. 33:W244-W248.

Sperschneider, J. 2020. Machine learning in plant−pathogen interactions: Empowering biological predictions from field scale to genome scale. New Phytol. 228:35-41.

Sperschneider, J., Catanzariti, A. M., DeBoer, K., Petre, B., Gardiner, D. M., Singh, K. B., Dodds, P. N., and Taylor, J. M. 2017. LOCALIZER: Subcellular localization prediction of both plant and effector proteins in the plant cell. Sci. Rep. 7:44598.

Sperschneider, J., Dodds, P. N., Gardiner, D. M., Manners, J. M., Singh, K. B., and Taylor, J. M. 2015. Advances and challenges in computational prediction of effectors from plant pathogenic fungi. PLoS Pathog. 11:e1004806.

Sperschneider, J., Gardiner, D. M., Dodds, P. N., Tini, F., Covarelli, L., Singh, K. B., Manners, J. M., and Taylor, J. M. 2016. EffectorP: Predicting fungal effector proteins from secretomes using machine learning. New Phytol. 210:743-761.

Staskawicz, B. J., Dahlbeck, D., and Keen, N. T. 1984 Cloned avirulence gene of *Pseudomonas syringae* pv. *glycinea* determines race-specific incompatibility on *Glycine max* (L.) Merr. Proc. Natl. Acad. Sci. U.S.A. 81:6024-6028.

Stavrinides, J., Ma, W., and Guttman, D. S. 2006. Terminal reassortment drives the quantum evolution of type III effectors in bacterial pathogens. PLoS Pathog. 2:e104.

Stukenbrock, E. H., Bataillon, T., Dutheil, J. Y., Hansen, T. T., Li, R., Zala, M., McDonald, B. A., Wang, J., and Schierup, M. H. 2011. The making of a new pathogen: Insights from comparative population genomics of the domesticated wheat pathogen *Mycosphaerella graminicola* and its wild sister species. Genome Res. 21:2157-2166.

Sugio, A., MacLean, A. M., and Hogenhout, S. A. 2014. The small phytoplasma virulence effector SAP11 contains distinct domains required for nuclear targeting and CIN-TCP binding and destabilization. New Phytol. 202:838-848.

Sun, H., Wang, J., Gong, Z., Yao, J., Wang, Y., Xu, J., Yuan, G. C., Zhang, Y., and Shao, Z. 2018. Quantitative integration of epigenomic variation and

transcription factor binding using MAmotif toolkit identifies an important role of IRF2 as transcription activator at gene promoters. Cell Discov. 4:38.

Syme, R. A., Tan, K. C., Rybak, K., Friesen, T. L., McDonald, B. A., Oliver, R. P., and Hane, J. K. 2018. Pan-*Parastagonospora* comparative genome analysis-effector prediction and genome evolution. Genome Biol. Evol. 10: 2443-2457.

Tang, B., Yan, X., Ryder, L. S., Cruz-Mireles, N., Soanes, D. M., Molinari, C., Foster, A. J., and Talbot, N. J. 2022. Rgs1 is a regulator of effector gene expression during plant infection by the rice blast fungus *Magnaporthe oryzae*. bioRxiv 2022.09.04.506535.

Temporini, E. D., and VanEtten, H. D. 2004. An analysis of the phylogenetic distribution of the pea pathogenicity genes of *Nectria haematococca* MPVI supports the hypothesis of their origin by horizontal transfer and uncovers a potentially new pathogen of garden pea: *Neocosmospora boniensis*. Curr. Genet. 46:29-36.

Tettelin, H., Masignani, V., Cieslewicz, M. J., Donati, C., Medini, D., Ward, N. L., Angiuoli, S. V., Crabtree, J., Jones, A. L., Durkin, A. S., Deboy, R. T., Davidsen, T. M., Mora, M., Scarselli, M., Margarit y Ros, I., Peterson, J. D., Hauser, C. R., Sundaram, J. P., Nelson, W. C., Madupu, R., Brinkac, L. M., Dodson, R. J., Rosovitz, M. J., Sullivan, S. A., Daugherty, S. C., Haft, D. H., Selengut, J., Gwinn, M. L., Zhou, L., Zafar, N., Khouri, H., Radune, D., Dimitrov, G., Watkins, K., O'Connor, K. J., Smith, S., Utterback, T. R., White, O., Rubens, C. E., Grandi, G., Madoff, L. C., Kasper, D. L., Telford, J. L., Wessels, M. R., Rappuoli, R., and Fraser, C. M. 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial "pan-genome". Proc. Natl. Acad. Sci. U.S.A. 102:13950-13955.

Teufel, F., Almagro Armenteros, J. J., Johansen, A. R., Gislason, M. H., Pihl, S. I., Tsirigos, K. D., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. 2022. SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat. Biotechnol. 40:1023-1025.

Thumuluri, V., Almagro Armenteros, J. J., Johansen, A. R., Nielsen, H., and Winther, O. 2022. DeepLoc 2.0: Multi-label subcellular localization prediction using protein language models. Nucleic Acids Res. 50:W228-W234.

Tollot, M., Assmann, D., Becker, C., Altmuller, J., Dutheil, J. Y., Wegner, C. E., and Kahmann, R. 2016. The WOPR protein Ros1 is a master regulator of sporogenesis and late effector gene expression in the maize pathogen *Ustilago maydis*. PLoS Pathog. 12:e1005697.

Torto, T. A., Li, S., Styer, A., Huitema, E., Testa, A., Gow, N. A., van West, P., and Kamoun, S. 2003. EST mining and functional expression assays identify extracellular effector proteins from the plant pathogen *Phytophthora*. Genome Res. 13:1675-1685.

Vlaardingerbroek, I., Beerens, B., Schmidt, S. M., Cornelissen, B. J., and Rep, M. 2016. Dispensable chromosomes in *Fusarium oxysporum* f. sp. *lycopersici*. Mol. Plant Pathol. 17:1455-1466.

Wagner, N., Avram, O., Gold-Binshtok, D., Zerah, B., Teper, D., and Pupko, T. 2022. Effectidor: An automated machine-learning based web server for the prediction of type-III secretion system effectors. Bioinformatics 38:2341-2343.

Wang, Q., Jiang, C., Wang, C., Chen, C., Xu, J. R., and Liu, H. 2017. Characterization of the two-speed subgenomes of *Fusarium graminearum* reveals the fast-speed subgenome specialized for adaption and infection. Front. Plant Sci. 8:140.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. 2018. SWISS-MODEL: Homology modelling of protein structures and complexes. Nucleic Acids Res. 46:W296-W303.

Webb, B., and Sali, A. 2016. Comparative protein structure modeling using MODELLER. Curr. Protoc. Bioinform. 54:5.6.1-5.6.37.

Wilson, C. J., Choy, W. Y., and Karttunen, M. 2022. AlphaFold2: A role for disordered protein/region prediction? Int. J. Mol. Sci. 23:4591.

Win, J., Krasileva, K. V., Kamoun, S., Shirasu, K., Staskawicz, B. J., and Banfield, M. J. 2012. Sequence divergent RXLR effectors share a structural fold conserved across plant pathogenic oomycete species. PLoS Pathog. 8:e1002400.

Win, J., Morgan, W., Bos, J., Krasileva, K. V., Cano, L. M., Chaparro-Garcia, A., Ammar, R., Staskawicz, B. J., and Kamoun, S. 2007. Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. Plant Cell 19:2349-2369.

Wood, K. J., Nur, M., Gil, J., Fletcher, K., Lakeman, K., Gann, D., Gothberg, A., Khuu, T., Kopetzky, J., Naqvi, S., Pandya, A., Zhang, C., Maisonneuve, B., Pel, M., and Michelmore, R. 2020. Effector prediction and characterization in the oomycete pathogen Bremia lactucae reveal host-recognized WY domain proteins that lack the canonical RXLR motif. PLoS Pathog. 16:e1009012.

Xin, X. F., Kvitko, B., and He, S. Y. 2018. *Pseudomonas syringae:* What it takes to be a pathogen. Nat. Rev. Microbiol. 16:316-328.

Yan, X., Tang, B., Ryder, L. S., MacLean, D., Were, V. M., Eseola, A. B., Cruz-Mireles, N., Foster, A. J., Osés-Ruiz, M., and Talbot, N. J. 2022. The transcriptional landscape of plant infection by the rice blast fungus *Magnaporthe oryzae* reveals distinct families of temporally co-regulated and structurally conserved effectors. bioRxiv 2022.07.18.500532.

Yang, Y., Zhao, J., Morgan, R. L., Ma, W., and Jiang, T. 2010. Computational prediction of type III secreted proteins from gram-negative bacteria. BMC Bioinf. 11:S47.

Zhang, Y., Jalan, N., Zhou, X., Goss, E., Jones, J. B., Setubal, J. C., Deng, X., and Wang, N. 2015. Positive selection is the main driving force for evolution of citrus canker-causing *Xanthomonas*. ISME J. 9:2128-2138.

Zhang, Z. M., Ma, K. W., Gao, L., Hu, Z., Schwizer, S., Ma, W., and Song, J. 2017. Mechanism of host substrate acetylation by a YopJ family effector. Nat. Plants 3:17115.

Zhang, Z. M., Ma, K. W., Yuan, S., Luo, Y., Jiang, S., Hawara, E., Pan, S., Ma, W., and Song, J. 2016. Structure of a pathogen effector reveals the enzymatic mechanism of a novel acetyltransferase family. Nat. Struct. Mol. Biol. 23: 847-852.

Zhou, H., Morgan, R. M., Guttman, S., and Ma, W. 2009. Allelic variants of the *Pseudomonas syringae* Type III effector HopZ1 are differentially recognized by plant resistance systems. Mol. Plant-Microbe Interact. 22:176-189.