

Review

Information Management of Genome Enabled Data Streams for *Pseudomonas syringae* on the Pseudomonas-Plant Interaction (PPI) Website

Magdalen Lindeberg

Department of Plant Pathology and Plant Microbe Biology, 302 Plant Science Building,
Cornell University, Ithaca NY 14853, USA; E-Mail: ML16@cornell.edu; Tel.: +1-607-254-7297;
Fax: +1-607-255-4471

Received: 2 September 2011; in revised form: 12 October 2011 / Accepted: 13 October 2011 /
Published: 2 November 2011

Abstract: Genome enabled research has led to a large and ever-growing body of data on *Pseudomonas syringae* genome variation and characteristics, though systematic capture of this information to maximize access by the research community remains a significant challenge. Major *P. syringae* data streams include genome sequence data, newly identified type III effectors, biological characterization data for type III effectors, and regulatory feature characterization. To maximize data access, the Pseudomonas-Plant Interaction (PPI) website [1] is primarily focused on categorization of type III effectors and curation of effector functional data represented in the Hop database and Pseudomonas-Plant Interaction Resource, respectively. The PPI website further serves as a conduit for incorporation of new genome characterization data into the annotation records at NCBI and other data repositories, and clearinghouse for additional data sets and updates in response to the evolving needs of the research community.

Keywords: *Pseudomonas syringae*; genome annotation; PPI; type III effector

1. Introduction

The bacterial phytopathogen, *Pseudomonas syringae*, is composed of a large number of strains having diverse and host-specific interactions with host plants and as such, has long been recognized as a valuable model system for dissecting the molecular interactions governing pathogenesis [2]. Complete genome sequences for three diverse *P. syringae* strains [3-5] together with the more recent

wave of draft sequences [6-11] has provided new opportunities for characterization of *P. syringae*-host interactions; however, the systematic extraction and communication of meaningful information from the large volume of data being generated remains a significant hurdle. Indeed, *P. syringae* exemplifies the promises and challenges of using genome sequence data and other high throughput analyses to address complex molecular dynamics in host-pathogen interactions.

The *P. syringae* pan-genome has been shaped by a long-standing evolutionary arms race between the various strains and their plant hosts. Bacterial features playing a prominent role in determining the outcome of the host-pathogen interaction include conserved features of the bacterium that function as triggers for innate immunity in the host, effector proteins secreted by the type III secretion system and interfering with detection of the bacterium by the plant, toxins modulating virulence through a variety of mechanisms, and various metabolic capabilities implicated in the exploitation of the available nutrients in different plant hosts [2,12,13]. Genome sequencing has revealed the extent of strain-to-strain differences in these features, particularly among the Type III effector and toxin repertoires. Surprising levels of variation are evident even among closely related strains [8,11], generally defying a simple explanation for differences in host specificity.

Fundamental outstanding questions include some of the following. What is the repertoire of type III effectors and other factors collectively required for pathogenesis on different hosts? What is the inventory of host genes interacting with effectors and involved in defense? How are *P. syringae* isolates evolving and outcompeting one another in response to inter-strain competition and selection pressure from host plants carrying specific resistance gene profiles? Ultimately, what elements of host defense could potentially be manipulated to create more durable resistance in economically important plants? Genome-wide analyses coupled with targeted characterization of individual system components of greatest interest offer the best hope for comprehensively addressing these questions. However, the scale of data currently being generated by the “next generation” sequencing technologies in particular, presents a real challenge in how best to extract and disseminate the information contained within these datasets. This review will address some of the major types of data being generated by the *P. syringae* research community as well as strategies for maximizing accessibility to the user community.

2. *P. syringae* Data Streams

Data being generated by the *P. syringae* research community falls into several categories, each presenting its own particular challenges. Specific approaches in use to process these data at the Pseudomonas-Plant Interaction (PPI) website will be addressed in Section 3.

2.1. Genome Sequences

Genome sequencing of *P. syringae* strains continues to provide an important source of new data on the potential bases of strain to strain variation in the interactions with host plants. Complete sequences were published for *P. syringae* pv. *tomato* DC3000 (*Pto* DC3000), *P. syringae* pv. *syringae* B728a [4], and *P. syringae* pv. *phaseolicola* 1448A [5] between 2003 and 2005, with *Pto* DC3000 receiving the most comprehensive experimental characterization. Since then, advances in sequencing technology have made possible the rapid generation of draft genomes of which 13 have been released of the over 30 *P. syringae* genome projects currently registered at NCBI. These sequences represent a wealth of

critical raw material for a wide range of analyses including identification of predicted virulence factors and, through genome comparison, both identification of genes linked to particular phenotypes and characterization of the evolving *P. syringae* pan-genome [14].

2.2. Type III Effectors—Identification in New Genomes

Of the *P. syringae* genetic components, type III effectors have been subject to the most intensive characterization owing to their central role in pathogenicity and host range. Indeed, one of the chief motivations for sequencing *Pto* DC3000 and subsequent strains has been the comprehensive identification of type III effector repertoires. Sequence features associated with the type III effector or *hop* genes (*hrp* outer proteins) include a binding site for the HrpL regulator upstream of the gene start, and conserved residues in the N-terminal regions of predicted Hop proteins associated with translocation by the type III secretion system. Methods for effector-mining, exploiting these features and developed using the closed genomes, have been applied to subsequent sequences [8,9].

Hop gene identification in the newly sequenced genomes adds to the expanding picture of effector repertoires in several respects, the most significant being identification of novel type III effector gene families. A recent study characterizing the effector repertoires of 14 newly sequenced *P. syringae* genomes revealed the existence of nine new *hop* gene families [11], raising the tantalizing possibility that the full range of effector-host interactions has yet to be identified. The vast majority of type III effectors encoded by the newly sequenced genomes are similar to ones previously identified, though as demonstrated for the HopZ and AvrPto families, relatively small changes in sequence can dramatically alter interactions with plant hosts [15-17]. Indeed, phenotypic differences observed for these naturally occurring variants have contributed significantly to our understanding of effector actions. Finally, genome sequencing has led to identification of large numbers of *hop* gene fragments and pseudogenes. Though not encoding functional proteins, these gene fragments provide important clues as to how effector repertoires have changed over time due to selection pressure from the host [14].

2.3. Type III Effectors—Biological Characterization

Identification of new type III effector families is paralleled by experimental characterization of their interactions with host plants. The generally accepted model of plant defense involves recognition of conserved elements of the bacterial cell known as pathogen associated molecular patterns or PAMPS by pattern recognition receptors, resulting in signal transduction and ultimately induction of various defense processes [18]. This process, also known as PAMP-triggered immunity or PTI, can be disrupted or suppressed at various stages by the Type III effectors. In the ongoing evolutionary arms race between hosts and pathogens, plants have countered effector suppression of PTI by evolving the capacity to detect the effectors, inducing effector triggered immunity or ETI [19]. Characterization of the interaction between individual effectors and specific plant proteins and processes is ongoing in several model plant systems including *Arabidopsis thaliana*, tomato, and *Nicotiana benthamiana*. A variety of primary and secondary effects of type III effectors on plant biological processes have been documented using experimental approaches ranging from pull-down assays confirming specific protein-protein interactions to microarray data showing large scale effector-induced changes in plant gene expression [20,21]. Host interactions for other bacterial features including the various PAMPS

and toxins such as coronatine are also being characterized, resulting in a large and complex body of literature that describes the intricacies of host-pathogen interaction at the molecular level [22,23].

2.4. Regulatory Features and Transcript Identification

Although the type III effectors are a major focus of *P. syringae* research, other genome scale data streams are simultaneously being generated with significant potential to impact our understanding of the larger picture of *P. syringae* metabolism [24-26]. Most significant among these has been the application of transcriptome and proteome sequencing to evaluate gene expression in *Pto* DC3000. The transcriptome mapping has also revealed locations of small RNAs which are linked to gene regulation in numerous biological systems [24]. To date, transcriptome sequencing has focused primarily on transcript profiles of bacteria grown in an iron-limiting media analogous to plant apoplastic fluid. As additional conditions are tested, an increasingly nuanced picture of bacterial gene expression is expected to emerge.

3. Challenges and Opportunities in the Analysis and Dissemination of *P. syringae* Data Streams

A major challenge presented by this wealth of data is how to best manage the different data streams for the purpose of maximizing accessibility to users, reducing duplication among data producers, and more efficiently identifying emergent properties of the system. Achievement of these goals further requires that data consolidation be conducted according to consistent standards, that it maintain fidelity and transparency to the information source so that users can rapidly access primary information as needed, and that it keep pace with ongoing research.

Online data repositories presently in existence come in a variety of forms, the best known of which are the large-scale, stably funded data clearinghouses at NCBI, EMBL, and DDBJ. Journal-driven mandates that published sequences be deposited at these sites historically have provided a durable link between sequence data and related publications. However, this system has struggled to accommodate the scale of data generated in the post-sequencing era with a major casualty being the link between ongoing experimental characterization and deposited sequence data. Even the ongoing curation of select genome sequences by Refseq [27], created in 2002 to provide comprehensive and consistently annotated sequence records, remains highly general and rarely incorporates species-specific, experimentally-documented findings. As experimental studies rely increasingly on previously deposited sequence data, and published output often fails to reference the source sequence or produce updates to the annotation record, the link between experimental findings and gene annotation becomes even more tenuous.

In response to the limitations inherent to large generic databases, many model organism research communities maintain organism-specific databases where sequence data can be downloaded, new sequence characterizations added, and tools for genome manipulation made available. Examples include the *Pseudomonas* Genome Database [28] as well as the SOL Genomics Network (SGN) [29] and The Arabidopsis Information Resource (TAIR) [30], devoted to solanaceous plants and *Arabidopsis*, respectively. These sites are typically maintained by a team of personnel and require significant computational resources, thus being dependent on dedicated funding sources. The model adopted by the *Pseudomonas*-Plant Interaction (PPI) genome resources website [1] is less resource

intensive, attempting to address data consolidation needs specific to the *P. syringae* research community through targeted curation of features of particular interest, leveraging existing resources when possible, and minimizing infrastructure.

In practice, the PPI website serves to keep users informed of the many categories of data relevant to *P. syringae* as well as the range of tools available for data analysis. For example, the site maintains a running tally of complete and draft *P. syringae* genome sequences on its main page. Additionally it serves as a repository for datasets that are either too extensive for publication or have been previously published as supplemental files. Examples of these can be found under links to “Comparative genomics and genome organization”, “Plant defense gene characterization”, and “Systems biology and gene expression modeling”, the last of which contains files pertaining to characterization of regulatory features. To facilitate visualization of genome sequences and sequence features, critical to understanding overall genome structure, tutorials have also been provided for use of the Artemis genome viewer and Artemis Comparison Tool (ACT) [31].

3.1. Type III Effector Sequence Curation on the PPI Website

A category of data receiving particular attention on the PPI website are the *hop* gene sequences themselves, information on which is summarized in the PPI Hop Database. The rapid burst of *hop* gene characterization following release of the *Pto* DC3000 genome sequence made apparent the need for systematic curation of these genes with emphasis on the need to develop community standards regarding the criteria for gene identification and name assignment. The resulting nomenclature system and criteria on which it is based have proven highly durable in the face of ongoing genome sequencing [32]. Newly characterized *hop* sequences can be submitted to the PPI website directly [1], or characterization and nomenclature assignments performed in collaboration with website personnel. More commonly, website personnel generate independent annotations of draft genomes, from which *hop* gene sequences are identified on the basis of BLAST hits and locations of predicted HrpL binding sites. At present, over 650 individual type III effector genes have been sequenced, composing 59 distinct families. Only full-length *hop* genes from draft genomes are incorporated into the type III effector database except in cases where re-sequencing has confirmed the validity of truncations arising from frameshifts or internal stop codons.

3.2. Type III Effector Functional Curation on the PPI Website

As described in section 2.3, biological characterization of the impact of type III effectors on host plants has been conducted over many years using various host plants, a variety of experimental methods, and is described in the literature with inconsistent terminology which itself has evolved as our understanding of the nature of the plant interaction has changed. Subsequent to development of the Hop Database, the *Pseudomonas*-Plant Interaction Resource was added to the PPI website with the goal of systematically capturing the varied data streams relevant to ongoing biological characterization of pathogen gene products interacting with the host. A major question regarding the design of this resource concerned how best to translate and condense the functional data into a more accessible format. While its complexity defies easy comprehension, particularly for newcomers in the field, it is too incomplete to merit diagrammatic representation akin to that used by Kyoto Encyclopedia of Genes

and Genomes (KEGG) to represent metabolic pathways [33]. It was ultimately decided that Gene Ontology Annotation (GOA), involving assignment of systematic descriptive terminology to the individual gene products and attached to evidence codes, source publications, and host taxonomy, represented the best strategy for capturing these data [34].

The decision to use Gene Ontology Annotation was reinforced by the availability of new terms developed by the Plant-Associated Microbe Gene Ontology working group to specifically describe host-symbiont interactions [35]. GO term assignments to *P. syringae* Type III effectors form the basis of the *P. syringae*-Plant Interaction Resource, with terms describing the cellular components or locations to which the effectors are targeted, their molecular functions, and the biological processes in which they engage. A “Host Interaction” field is also included to provide additional descriptive notes and hyperlinks to interacting host proteins. At present, over 600 GO annotation assignments are listed for 50 type III effectors and 8 helper proteins. Gene Ontology annotations generated for type III effector HopX1 are shown in Figure 1. The data can be accessed in a sortable, non-formatted spreadsheet and in formatted tables hyperlinked to interacting proteins, source publications, and the GO database itself. Relevant annotation fields from the *P. syringae*-Plant Interaction Resource are forwarded on to the GO database [36].

3.3. PPI Integration with Outside Resources

Consistent with the goal of leveraging existing resources when possible, the PPI site does not host independent annotation records or a dedicated genome viewer but rather functions as a portal for integrating new information on genome features into other, durable supported databases. Data flow typically begins with direct communication of annotation updates from PPI to Genbank, with updates automatically reflected in RefSeq, EMBL and uploaded into the annotation records at the *Pseudomonas* Genome database [28] on a semi-annual basis. At this time, over 3400 annotation changes to *P. syringae* genomes have been forwarded to NCBI via the PPI website. Given that Genbank, Refseq, and EMBL annotation records form the basis for most secondary genome analyses, this approach has the added benefit of maximizing access to the most current research finding, in contrast to many model organism databases where in-house annotation changes are not reliably reflected by the large generic databases.

Although corrected *hop* gene nomenclature and mapping of HrpL binding sites were an early focus of annotation updates, proteome and transcriptome sequencing have more recently led to updated gene coordinates, addition of locus qualifiers describing the bioinformatic and experimental support for individual gene calls, and annotation of small RNAs. In light of the constant challenge presented by the large numbers of hypothetical and conserved hypothetical genes in bacterial genomes, incorporation of these qualifiers into the *Pto* DC3000 genome provides invaluable information for those using the *Pto* DC3000 annotation as a basis for annotation of related genomes. An outline of the overall data handling strategy employed for a variety of data types at the PPI website is shown in Figure 2.

Figure 1. Gene ontology (GO) annotation term assignments for type III effector HopX1 from *P. syringae* pv. *tomato*. GO terms used fall into three separate ontologies corresponding to cellular component (C), molecular function (F), and biological process (P). First author and publication year of the source of the data are shown together with a hyperlinked Pubmed ID. A key to evidence codes can be found at [37].

hopX1 (PSPTO_A0012) (RefSeq link)

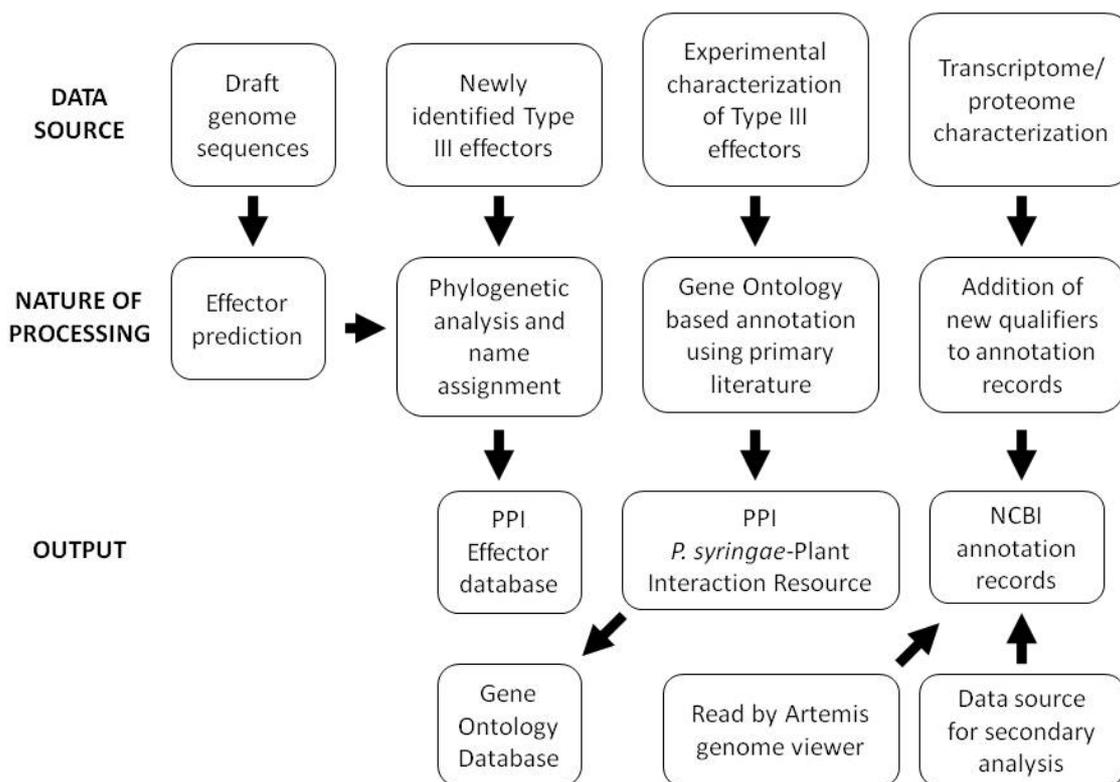
[return to top](#)

GO term	GO ID	Ontology	Reference	PubMed ID	with	Evidence Code	Host ID	Host interaction
host cell cytoplasm	GO:0030430	C	Nimchuk 2007	17427805	-	IDA	A thal Ws-0	-
interaction with host via protein secreted by type III secretion system	GO:0052049	P	Schechter 2004	14702323	-	IDA	N benth	-
cysteine-type peptidase activity	GO:0030234	F	Nimchuk 2007	17427805	-	IMP	-	-
positive regulation by symbiont of host resistance gene-dependent defense response	GO:0052527	P	Nimchuk 2007	17427805	-	IDA	P vulgaris	positive regulation of R2-dependent defense-related PCD
pathogenesis	GO:0009405	P	Jamir 2004	14756767	-	IDA	A thal Ws-0	negative regulation by symbiont of host defense-related PCD
			Jamir 2004	14756767	-	IDA	N tab Xanthi	negative regulation by symbiont of host defense-related PCD
negative regulation by symbiont of host defense-related PCD	GO:0034254	P	Jamir 2004	14756767	-	IDA	N tab Xanthi	-
			Jamir 2004	14756767	-	IDA	A thal Ws-0	-
negative regulation by symbiont of host defense-related protein level	GO:0033363	P	Jamir 2004	14756767	-	IDA	N tab Xanthi	reduces upregulation of PR1a expression by HopA1
positive regulation by symbiont of host defense-related protein level	GO:0033364	P	He 2004	14756769	-	IMP	A thal Col-0	reduced upregulation of RAP2.6 in hopX1 mutant
negative regulation by symbiont of host resistance gene-dependent defense response	GO:0033660	P	Guo 2009	19656042	-	IDA	A thal Col-0	reduces HR elicited by AvrRpm1-RPM1 interaction
			Jamir 2004	14756767	-	IDA	N tab xanthi	reduces HR elicited by HopA1- RPS6 interaction
negative regulation by symbiont of defense-related host cell wall callose deposition	GO:0052202	P	Guo 2009	19656042	-	IDA	A thal	negative regulation of callose deposition induced by AvrRpm1
			Guo 2009	19656042	-	IDA	A thal	negative regulation of callose deposition induced by <i>Pseudomonas fluorescens</i>
positive regulation by symbiont of host defense-related PCD	GO:0034255	P	Nimchuk 2007	17427805	-	IDA	A thal Col-0	-
			Nimchuk 2007	17427805	-	IDA	P vulgaris	R2-dependent defense-related PCD
			Wei 2007	17559511	-	IDA	N benth	-
negative regulation by symbiont of host apoptosis	GO:0033668	P	Jamir 2004	14756767	-	IDA	S cerevisiae	-
			Jamir 2004	14756767	-	IDA	N benth	-

3.4. PPI Website—Design Philosophy

The PPI website is designed around a simple model for data delivery, depending primarily on flat file formats such as Excel spreadsheets rather than relational databases. While this approach may ultimately shift as the nature of data being produced by the *P. syringae* research community changes, the current strategy has proven easy to maintain with the limited resources and may in fact represent a more flexible model for the volume of data in question. As Matthews et al have pointed out, online resources that permit data access only in response to specific queries can be overly limiting while maintaining a significant degree of browsability permits “the possibility of higher order perusal of data to allow the potential for recognizing unexpected associations” [38]. The durability of this model is further borne out by that fact that the PPI website and component elements have been regularly updated for over nine years, during a period when the average lifetime of biological databases is approximately 18 months [39].

Figure 2. Strategy for data management used at the Pseudomonas-Plant Interaction (PPI) genome resources website [1].



4. Conclusions

The larger mission of the PPI website is to orient users to a wide range of *P. syringae* data with particular emphasis on type III effector identification and nomenclature via the Hop Database, and Gene Ontology-based documentation of functional properties as represented in the *P. syringae*-Plant Interaction Resource. The latter is currently being expanded to include annotations for the bacterial features (PAMPs) implicated in triggering innate immunity. A large scale analysis of tomato kinases interacting with PAMPs and individual effectors has been recently initiated and is expected to result in

significant expansion of known biological interactions. The ultimate goal of an expanded *P. syringae*-Plant Interaction Resource is the generation of a data network where correlations can be made between repertoires of effectors, functional roles, and host range.

These analyses depend not only on effective data consolidation but also on sufficiently detailed analysis of draft genome sequences. Drafts have the potential to yield important insights on effector repertoires from phenotypically diverse strains; however, re-sequencing of gene fragments to distinguish true effector mutations from assembly errors is critical to the utility of these data.

Acknowledgements

This work is supported by the National Science Foundation Plant Genome Research Program DBI-0605059, National Research Initiative of the USDA Cooperative State Research, Education and Extension Service, grant number 2005-35600-16370 and by the U.S. National Science Foundation, grant number EF-0523736

References

1. *Pseudomonas*-Plant Interaction Website. Available online: <http://pseudomonas-syringae.org/> (accessed on 5 October 2011).
2. Mansfield, J.W. 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.* **2009**, *10*, 721–734.
3. Buell, C.R.; Joardar, V.; Lindeberg, M.; Selengut, J.; Paulsen, I.T.; Gwinn, M.L.; Dodson, R.J.; Deboy, R.T.; Durkin, A.S.; Kolonay, J.F. The complete genome sequence of the Arabidopsis and tomato pathogen *Pseudomonas syringae* pv. tomato DC3000. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 10181–10186.
4. Feil, H.; Feil, W.S.; Chain, P.; Larimer, F.; DiBartolo, G.; Copeland, A.; Lykidis, A.; Trong, S.; Nolan, M.; Goltsman, E.; *et al.* Comparison of the complete genome sequences of *Pseudomonas syringae* pv. *syringae* B728a and pv. tomato DC3000. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 11064–11069.
5. Joardar, V.; Lindeberg, M.; Jackson, R.W.; Selengut, J.; Dodson, R.; Brinkac, L.M.; Daugherty, S.C.; Deboy, R.; Durkin, A.S.; Giglio, M.G.; *et al.* Whole-genome sequence analysis of *Pseudomonas syringae* pv. *phaseolicola* 1448A reveals divergence among pathovars in genes involved in virulence and transposition. *J. Bacteriol.* **2005**, *187*, 6488–6498.
6. Reinhardt, J.A.; Baltrus, D.A.; Nishimura, M.T.; Jeck, W.R.; Jones, C.D.; Dangl, J.L. De novo assembly using low-coverage short read sequence data from the rice pathogen *Pseudomonas syringae* pv. *oryzae*. *Genome Res.* **2009**, *19*, 294–305.
7. Green, S.; Laue, B.; Fossdal, C.G.; Hara, S.W.A.; Cottrell, J.E. Infection of horse chestnut (*Aesculus hippocastanum*) by *Pseudomonas syringae* pv. *aesculi* and its detection by quantitative real-time PCR. *Plant Pathol.* **2009**, *58*, 731–744.

8. Almeida, N.; Yan, S.; Lindeberg, M.; Studholme, D.; Condon, B.; Liu, H.; Viana, C.; Warren, A.; Evans, C.; Kemen, E.; *et al.* A draft genome sequence of *Pseudomonas syringae* pv. tomato strain T1 reveals a repertoire of type III related genes significantly divergent from that of *P. syringae* pv. tomato strain DC3000. *Mol. Plant Microbe Interact.* **2009**, *22*, 52–62.
9. Rodriguez-Palenzuela, P.; Matas, I.M.; Murillo, J.; Lopez-Solanilla, E.; Bardaji, L.; Perez-Martinez, I.; Rodriguez-Mosquera, M.E.; Penyalver, R.; Lopez, M.M.; Quesada, J.M.; *et al.* Annotation and overview of the *Pseudomonas savastanoi* pv. *savastanoi* NCPPB 3335 draft genome reveals the virulence gene complement of a tumour-inducing pathogen of woody hosts. *Environ. Microbiol.* **2010**, *12*, 1604–1620.
10. Qi, M.; Wang, D.; Bradley, C.A.; Zhao, Y. Genome sequence analyses of *Pseudomonas savastanoi* pv. *glycinea* and subtractive hybridization-based comparative genomics with nine pseudomonads. *PLoS One* **2011**, *6*, doi:10.1371/journal.pone.0016451.
11. Baltrus, D.A.; Nishimura, M.T.; Romanchuk, A.; Chang, J.H.; Mukhtar, M.S.; Cherkis, K.; Roach, J.; Grant, S.R.; Jones, C.D.; Dangl, J.L. Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathog.* **2011**, *7*, doi:10.1371/journal.ppat.1002132.
12. Rico, A.; Preston, G.M. *Pseudomonas syringae* pv. tomato DC3000 uses constitutive and apoplast-induced nutrient assimilation pathways to catabolize nutrients that are abundant in the tomato apoplast. *Mol. Plant Microbe Interact.* **2008**, *21*, 269–282.
13. Bender, C.L.; Alarcon-Chaidez, F.; Gross, D.C. *Pseudomonas syringae* phytotoxins: Mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. *Microbiol. Mol. Biol. Rev.* **1999**, *63*, 266–292.
14. Lindeberg, M.; Myers, C.R.; Collmer, A.; Schneider, D.J. Roadmap to new virulence determinants in *Pseudomonas syringae*: Insights from comparative genomics and genome organization. *Mol. Plant Microbe Interact.* **2008**, *21*, 685–700.
15. Morgan, R.L.; Zhou, H.; Lehto, E.; Nguyen, N.; Bains, A.; Wang, X.; Ma, W. Catalytic domain of the diversified *Pseudomonas syringae* type III effector HopZ1 determines the allelic specificity in plant hosts. *Mol. Microbiol.* **2010**, *76*, 437–455.
16. Zhou, H.; Morgan, R.L.; Guttman, D.S.; Ma, W. Allelic Variants of the *Pseudomonas syringae* Type III Effector HopZ1 Are Differentially Recognized by Plant Resistance Systems. *Mol. Plant Microbe Interact.* **2009**, *22*, 176–189.
17. Kunkeaw, S.; Tan, S.; Coaker, G. Molecular and evolutionary analyses of *Pseudomonas syringae* pv. tomato race 1. *Mol. Plant Microbe Interact.* **2010**, *23*, 415–424.
18. Segonzac, C.; Zipfel, C. Activation of plant pattern-recognition receptors by bacteria. *Curr. Opin. Microbiol.* **2011**, *14*, 54–61.
19. Jones, J.D.; Dangl, J.L. The plant immune system. *Nature* **2006**, *444*, 323–329.
20. Truman, W.; de Zabala, M.T.; Grant, M. Type III effectors orchestrate a complex interplay between transcriptional networks to modify basal defence responses during pathogenesis and resistance. *Plant J.* **2006**, *46*, 14–33.
21. Xiang, T.; Zong, N.; Zou, Y.; Wu, Y.; Zhang, J.; Xing, W.; Li, Y.; Tang, X.; Zhu, L.; Chai, J.; *et al.* *Pseudomonas syringae* effector AvrPto blocks innate immunity by targeting receptor kinases. *Curr. Biol.* **2008**, *18*, 74–80.

22. Block, A.; Alfano, J.R. Plant targets for *Pseudomonas syringae* type III effectors: Virulence targets or guarded decoys? *Curr. Opin. Microbiol.* **2011**, *14*, 39–46.
23. Lindeberg, M.; Collmer, A. Gene Ontology for type III effectors: Capturing processes at the host-pathogen interface. *Trends Microbiol.* **2009**, *17*, 304–311.
24. Filiatrault, M.J.; Stodghill, P.V.; Bronstein, P.A.; Moll, S.; Lindeberg, M.; Grills, G.; Schweitzer, P.; Wang, W.; Schroth, G.P.; Luo, S.; *et al.* Transcriptome analysis of *Pseudomonas syringae* identifies new genes, noncoding RNAs, and antisense activity. *J. Bacteriol.* **2010**, *192*, 2359–2372.
25. Butcher, B.G.; Bronstein, P.A.; Myers, C.R.; Stodghill, P.V.; Bolton, J.J.; Markel, E.J.; Filiatrault, M.J.; Swingle, B.; Gaballa, A.; Helmann, J.D.; *et al.* Characterization of the Fur Regulon in *Pseudomonas syringae* pv. tomato DC3000. *J. Bacteriol.* **2011**, *193*, 4598–4611.
26. Swingle, B.; Thete, D.; Moll, M.; Myers, C.R.; Schneider, D.J.; Cartinhour, S. Characterization of the PvdS-regulated promoter motif in *Pseudomonas syringae* pv. tomato DC3000 reveals regulon members and insights regarding PvdS function in other pseudomonads. *Mol. Microbiol.* **2008**, *68*, 871–889.
27. Pruitt, K.D.; Tatusova, T.; Klimke, W.; Maglott, D.R. NCBI reference sequences: Current status, policy and new initiatives. *Nucl. Acids Res.* **2009**, *37*, D32–D36.
28. *Pseudomonas* Genome Database. Available online: <http://pseudomonas.com/> (accessed on 5 October 2011).
29. SOL Genomics Network. Available online: <http://solgenomics.net/> (accessed on 5 October 2011).
30. The *Arabidopsis* Information Resource. Available online: <http://www.arabidopsis.org/> (accessed on 5 October 2011).
31. Carver, T.; Berriman, M.; Tivey, A.; Patel, C.; Böhme, U.; Barrell, B.G.; Parkhill, J.; Rajandream, M.A. Artemis and ACT: Viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics* **2008**, *24*, 2672–2676.
32. Lindeberg, M.; Stavrinides, J.; Chang, J.H.; Alfano, J.R.; Collmer, A.; Dangl, J.L.; Greenberg, J.T.; Mansfield, J.W.; Guttman, D.S. Proposed guidelines for a unified nomenclature and phylogenetic analysis of type III Hop effector proteins in the plant pathogen *Pseudomonas syringae*. *Mol. Plant Microbe Interact.* **2005**, *18*, 275–282.33.
33. KEGG PATHWAY Database. Available online: <http://www.kegg.com/kegg/pathway.html> (accessed on 25 October 2011).
34. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; *et al.* Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **2000**, *25*, 25–29.
35. Torto-Alalibo, T.; Collmer, C.W.; Gwinn-Giglio, M.; Lindeberg, M.; Meng, S.; Chibucos, M.C.; Tseng, T.T.; Lomax, J.; Biehl, B.; Ireland, A.; *et al.* Unifying themes in microbial associations with animal and plant hosts described using the gene ontology. *Microbiol. Mol. Biol. Rev.* **2010**, *74*, 479–503.
36. The Gene Ontology. Available online: <http://www.geneontology.org/> (accessed on 5 October 2011).
37. Guide to GO Evidence Codes. Available online: <http://www.geneontology.org/GO.evidence.shtml> (accessed on 5 October 2011).

38. Matthews, D.E.; Lazo, G.R.; Anderson, O.D. Plant and crop databases. *Methods Mol. Biol.* **2009**, *513*, 243–262.
39. Rhee, S.Y.; Crosby, B. Biological databases for plant research. *Plant Physiol.* **2005**, *138*, 1–3.

© 2011 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).