

Article

Genomics-Based Exploration of Virulence Determinants and Host-Specific Adaptations of *Pseudomonas syringae* Strains Isolated from Grasses

Alexey Dudnik and Robert Dudler *

Institute of Plant Biology, University of Zurich, Zollikerstrasse 107, 8008 Zurich, Switzerland;
E-Mail: alexey.dudnik@botinst.uzh.ch

* Author to whom correspondence should be addressed; E-Mail: rdudler@botinst.uzh.ch;
Tel.: + 41-44-634-82-52; Fax: + 41-44-634-82-04.

Received: 22 November 2013; in revised form: 20 January 2014 / Accepted: 22 January 2014 /
Published: 28 January 2014

Abstract: The *Pseudomonas syringae* species complex has recently been named the number one plant pathogen, due to its economic and environmental impacts, as well as for its role in scientific research. The bacterium has been repeatedly reported to cause outbreaks on bean, cucumber, stone fruit, kiwi and olive tree, as well as on other crop and non-crop plants. It also serves as a model organism for research on the Type III secretion system (T3SS) and plant-pathogen interactions. While most of the current work on this pathogen is either carried out on one of three model strains found on dicot plants with completely sequenced genomes or on isolates obtained from recent outbreaks, not much is known about strains isolated from grasses (Poaceae). Here, we use comparative genomics in order to identify putative virulence-associated genes and other Poaceae-specific adaptations in several newly available genome sequences of strains isolated from grass species. All strains possess only a small number of known Type III effectors, therefore pointing to the importance of non-Type III secreted virulence factors. The implications of this finding are discussed.

Keywords: *Pseudomonas syringae*; Poaceae; host-pathogen interactions; genomics

1. Introduction

Strains of the *Pseudomonas syringae* group are causal agents of a variety of plant diseases worldwide. Strains of this group have been reported to infect nearly 200 different plant species [1], including both grassy and woody hosts. A number of those are agriculturally important plants, and as a consequence, *P. syringae* is one of the best-studied plant pathogens. While some diseases have been known to recur in the form of outbreaks for a long time [2–5], others have only recently emerged [6,7]. The taxonomy of *P. syringae* is still under discussion. According to the commonly used version, the species is sub-divided into over 50 groups called “pathovars”. This division takes into account host-specificity, disease type, as well as the biochemical characteristics of a strain [8,9]. In contrast, a number of phylogenetic studies have positioned *P. syringae* as a species complex, and depending on the approach used, it has been split either into five phylogenetic groups by multi-locus sequencing [10,11], or into nine so-called genomospecies based on DNA hybridization profiles [12]. In order to remain coherent with the majority of publications, in this work, we are going to regard *P. syringae* as a single species.

The observed large genetic diversity among different pathovars is a direct consequence of the wide host range. Moreover, heterogeneity is also found among strains belonging to the same pathovar [6,8,13]. The highest degree of variation is seen within the complement of virulence factors, which is the key element determining host range and the overall degree of virulence [14]. In addition to that, some strains possess host-specific metabolic pathways, such as enzymes for lignin degradation in pathogens of woody hosts [6,7,15]. For successful survival and reproduction, both epiphytic and endophytic strains of *P. syringae* deploy different sets of Type III and Type VI secreted effector proteins, exopolymeric substances, phytohormones, phytotoxins and other types of secreted molecules [14,16–23]. Among those, the major pathogenicity factor is the Type III secretion system (T3SS). Other notable virulence factors are the necrosis-inducing phytotoxins, syringomycin and syringopeptin, which are presumed to create pores in plant cells by imbedding into the plasma membrane, and the anti-metabolite toxins, tabtoxin, mangotoxin and phaseolotoxin, which inhibit glutamine synthetase, ornithine acetyltransferase and ornithine carbamoyl transferase, respectively [19,24]. Another group of phytotoxins includes coronatine and syringolin, both of which are involved in the inhibition of the salicylic acid-dependent immune response by mimicking jasmonic acid or by irreversibly inhibiting the proteasome, respectively [18,19].

The T3SS is a protein delivery machinery, which uses a structure, called injectosome, for the delivery of effector proteins directly into host cells by puncturing the cell membrane [25,26]. This machinery is essential for the pathogenesis of *P. syringae*, and knocking it out renders the bacterium avirulent. [14,17] The majority of Type III effectors (T3Es) are assumed to be involved in the suppression of plant defense, including pathogen-associated molecular pattern (PAMP)-triggered immunity and the hypersensitive response (effector-triggered immunity) [27,28]. In addition, some effectors were shown to have a cytotoxic effect [29,30]. Currently, there are 58 verified effector families recognized [31]. With a few exceptions, the exact mechanism of action of T3Es remains unknown. Among the well-characterized T3Es is AvrPtoB, which targets the flagellin recognition receptor, FLS2, and marks it for degradation [32], HopU1, which ADP-ribosylates several RNA-binding proteins, thus preventing the association with their target mRNAs [33], HopN1, which

targets photosystem II in order to inhibit reactive oxygen species production [34], AvrRps4, which targets a regulator of plant basal defenses [35], and HopZ1a, which interferes with plant microtubule network formation and jasmonic acid signaling [36,37]. Most *P. syringae* strains produce around two to three dozen T3Es [6,8,31]. However, not all of them are essential for full virulence, due to functional redundancy [38]. As a consequence, strains pathogenic to the same host often have divergent sets of effectors [6,8,13,39]. It is notable that effector repertoires are under heavy evolutionary pressure [27] and, thus, are being continuously remodeled. The remodeling might also result in a change of host specificity. The field of effector biology remains a hot topic, and currently, a lot of research is aimed at improving our knowledge about the molecular biology of interactions between plants and their pathogens.

The extensive research has also led to the accumulation of a large amount of available genome sequence data. Currently, GenBank contains records of the three completely sequenced *P. syringae* model strains DC3000 (pathovar *tomato*, *Pto*; pathogenic to tomato, *Arabidopsis thaliana* and *Nicotiana benthamiana* [40,41]), B728a (pathovar *syringae*, *Psy*; the causal agent of brown spot disease of bean [42]) and 1448A (pathovar *phaseolicola*, *Pph*; causes halo blight on bean [43]). The strains represent phylogenetic Clades one, two and three, respectively [11,31]. In addition, a number of incomplete genome sequences of various qualities are available for a variety of strains [6–8,15,31]. The majority of the sequenced strains were originally isolated from dicot plants; thus, the topic of diversity and adaptations of strains pathogenic to monocot plants remains largely unexplored. Currently, nine genome sequences of different qualities of strains isolated from true grasses (graminoids, family Poaceae) are available. These encompass four wheat (*Triticum aestivum*) isolates (*P. syringae* pv. *syringae* (*Psy*) strains SM and B64, *P. syringae* pv. *atrofaciens* DSM50255 (*Paf*) and *P. syringae* BRIP39023 [44–46]), three strains pathogenic to barley (*Hordeum vulgare*) (*P. syringae* BRIP34876 and BRIP34881 [46] and pathovar *japonica* strain M301072 (*Pja*) [31]), one strain collected from rice (*Oryza sativa*) (pathovar *oryzae* 1_6 (*Por*) [47]) and an isolate pathogenic to proso millet (*Panicum miliaceum*) (pathovar *panici* strain LMG2367 (*Ppa*) [48]). Here, we analyze these nine genomes using comparative genomics tools with the aim of identifying possible adaptations of *P. syringae* strains to life in graminoid host species. In addition, we also compared these strains with a group of unspecialized pathogenic strains belonging to *P. cannabina* pv. *alisalensis* (*Pcal*), which have been shown to colonize several dicot hosts, as well as oat (*Avena sativa*) and great brome (*Bromus diandrus*) plants, both of which belong to the Poaceae [49].

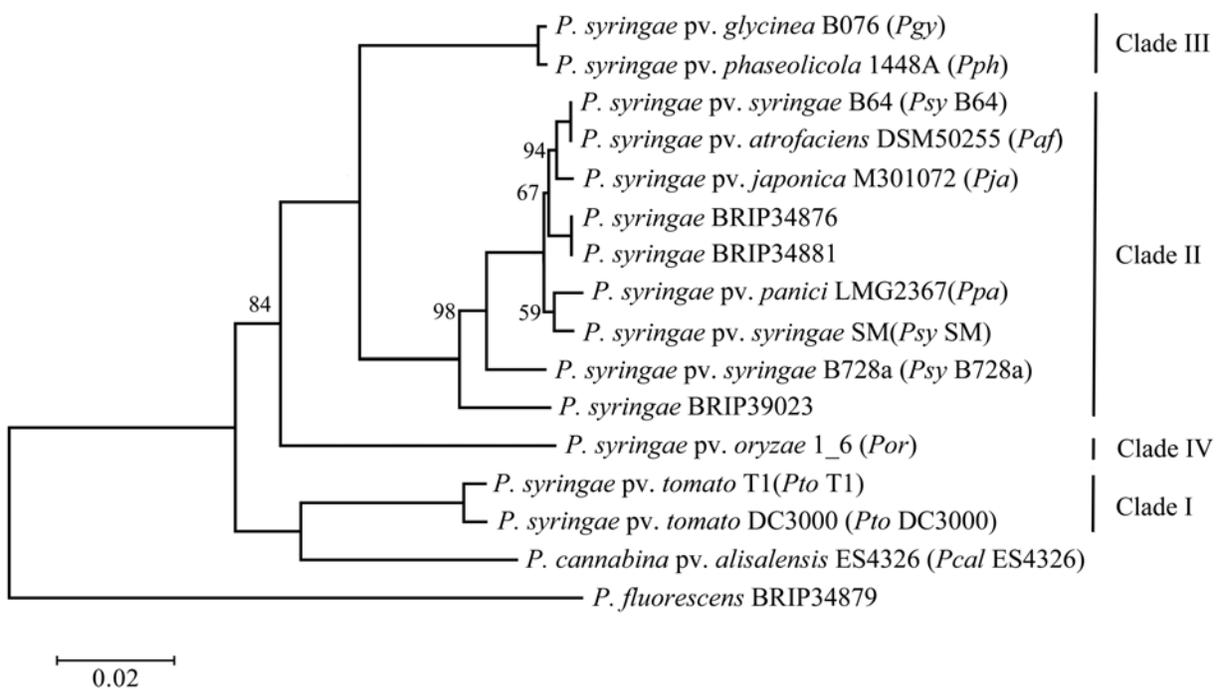
2. Results and Discussion

2.1. Phylogenetic Assessment of the Strains

Since most of the analyzed genomes have been sequenced within the past year, there is no published record of their phylogenetic characterization. Thus, the first step undertaken was to identify their relatedness by constructing a maximum likelihood tree using MLST (Multi-locus sequence typing) loci previously deployed in other studies [10,31]. Several other strains belonging to the three major phylogenetic clades [11] have been included as a reference. *P. fluorescens* BRIP34879, a strain which was also isolated from barley [46], was used as an outgroup. *Pcal* ES4326 (previously known as

P. syringae pv. *maculicola* ES4326) was included to represent the phylogenetic relatedness of the *P. cannabina* pv. *alisalensis* group. The resulting diagram is presented in Figure 1. Even though whole-genome-based phylogenies were recently demonstrated to be more accurate than MLST [50], we still decided to use the MLST approach, due to the poor assembly quality of some of the analyzed genomes, which resulted in a large number of partial and split genes. The presence of such sequences would introduce a bias to the analysis using the whole genome/proteome sequences.

Figure 1. Maximum likelihood phylogenetic tree of the analyzed *P. syringae* strains. *Pgy*, *P. syringae* pv. *glycinea*; *Pph*, *P. syringae* pv. *phaseolicola*; *Psy*, *P. syringae* pv. *syringae*; *Paf*, *P. syringae* pv. *atofaciens*; *Pja*, *P. syringae* pv. *japonica*; *Ppa*, *P. syringae* pv. *panici*; *Por*, *P. syringae* pv. *oryzae*; *Pto*, *P. syringae* pv. *tomato*; *Pcal*, *P. cannabina* pv. *alisalensis* (*Pcal*).



The tree was constructed based on the nucleotide sequences of several housekeeping genes. The scale bar indicates the number of amino acid changes per site. Bootstrapping-based probabilities are displayed only at nodes where the values are less than 100.

While most of the *P. syringae* strains that are the focus of this article cluster together within Clade II, *Por* 1_6 clearly forms a separate branch. This is in line with previously published data, where this strain was placed into Clade IV (also classified as *Pseudomonas coronafaciens* by Gardan and colleagues [12]). Interestingly, Clade IV contains several other strains isolated from cereal crops, including oat [11]. This suggests that a host-shift towards Poaceae has occurred at least twice in the evolution of *P. syringae*. Within Clade II, strain BRIP39023, a non-pathogenic wheat isolate, forms a separate branch and, therefore, is also less related to the rest of the isolates. The remaining strains form a cluster and, therefore, must be of the same ancestral lineage. It is notable that pathovars *syringae*, *atofaciens*, *panici* and *japonica* do not form distinct branches. This shows that discrimination into pathovars is not always descriptive, since particular strains often are pathogenic to several hosts, as shown, for example, for strains BRIP34876, BRIP34881 and *Ppa* [46,48]. This also suggests that a

pathogen can relatively easily change its specificity from one host to another among related plant species. Finally, it is not clear to which pathovar BRIP34876 and BRIP34881 should be assigned. It should be noted that there are several other *P. syringae* pathovars belonging to Clade II (*lapsa*) and Clade IV isolated from Poaceae [12,51], which were not included in this study, due to the lack of available genomic sequence data. *Pcal* ES4326 forms a distinct branch from *Pto* strains and, together with several related strains, was classified as Clade V by Hwang and colleagues [11]. The resulting tree is in agreement with previously published data [11,31].

2.2. Genome Comparison and Identification of Poaceae-Specific Genes

During the process of adaptation to a new host, there is a strong evolutionary pressure on a pathogen. As a consequence, a pathogen loses some genes that reduce the virulence and overall fitness in the new host. In addition, novel genes are acquired by horizontal gene transfer (HGT), in particular from species already living on or inside the new host. Examples include genes for the degradation of lignin, pectin and aromatic compounds found in *P. syringae* strains pathogenic to woody hosts [15,52]. Therefore, the strains isolated from grasses could also exhibit certain adaptations not found in other strains. In order to identify candidate adaptation genes, the genomes of the nine sequenced Poaceae isolates were first compared among themselves and, then, also, with a group of twelve other *P. syringae* strains isolated from a variety of hosts (See Table S1 for a complete list of strains). The comparison was performed using an 80% identity cut off in order to avoid the detection of paralogs and other genes with only partial homology. An outline of the results is shown in Table 1, and a list of the identified unique genes, which are shared by at least four strains, can be found in Table S2.

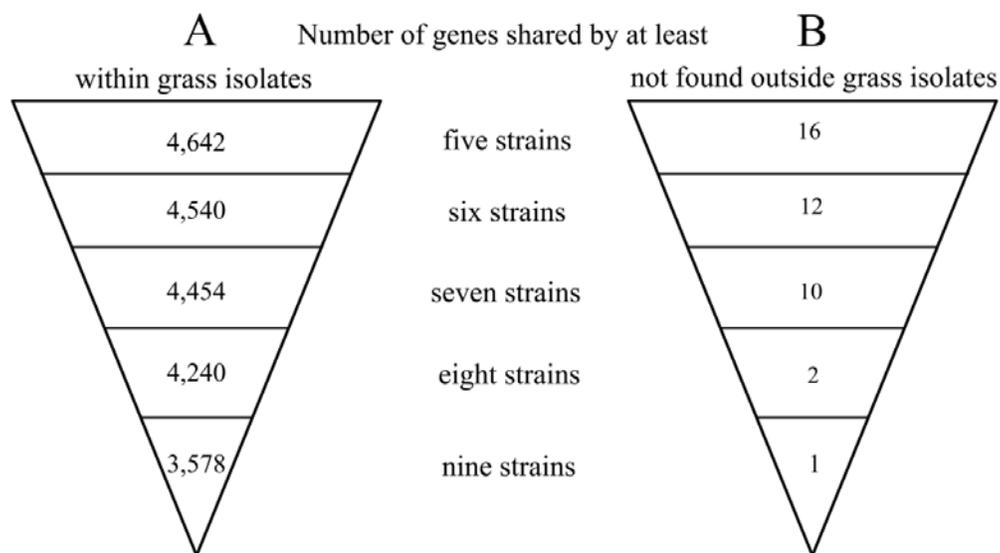
Table 1. Genome comparison of the nine analyzed Poaceae isolates: unique genes.

Ortholog clusters	A (within grass isolates)	B (not found outside grass isolates)
Unique to <i>Por</i>	2,333	1,566
Unique to BRIP34876	11	8
Unique to BRIP34881	13	7
Unique to BRIP39023	349	121
Unique to <i>Pja</i>	3,657	2,563
Unique to <i>Psy</i> SM	178	75
Unique to <i>Psy</i> B64	37	9
Unique to <i>Ppa</i>	484	322
Unique to <i>Paf</i>	216	188

Column **A** represents the numbers of ortholog clusters obtained by comparing the nine genomes of Poaceae isolates among themselves. Column **B** contains an overview on the distribution of ortholog clusters found exclusively in these nine isolates, but not outside. An ortholog cluster is a group of genes from at least one strain in which all members have an identity percentage equal to or above the set cut off. Thus, a single cluster might contain more than one gene per strain, due to the presence of recent gene duplications, which are still more than 80% identical to one another. Such genes are regarded as the same entity by the software, and as a consequence, the actual numbers of shared homologs are slightly different for each strain.

When the Poaceae isolates were compared to one another, the total number of identified ortholog clusters was 13,319. The calculated core genome, *i.e.*, the set of genes shared by all strains, has a size of 3,578 ortholog clusters (Figure 2A), which is comparable to the previously published data [31]. Outside the core genome, 2,471 clusters are shared by at least two strains. An overview of shared genes is depicted in Figure 2A. The remaining ortholog clusters have only one member, and respective genes are therefore unique to one of the strains. The number of unique genes is highly variable, ranging from as little as eleven for BRIP34876 to almost 3,657 for *Pja* (Table 1, Column A). The large values observed for *Pja* and *Por*, however, originate from the low quality of the respective assemblies (see Table S1). Because of this, there are many proteins that are found on several contigs. For example, the *Pja* genome contains seven entries for the chromosomal replication initiation protein, DnaA, which is normally present as a single copy next to the origin of replication. It is highly likely that many partial sequences did not pass our strict identity criterion and therefore ended up as individual ortholog clusters. Moreover, the majority of the genomes did not undergo manual curation, and thus, their annotation data contains a number of very small hypothetical, as well as truncated proteins.

Figure 2. Shared genes in the genomes of the nine analyzed Poaceae isolates.



Column **A** represents the numbers of ortholog clusters shared by five or more of the nine Poaceae isolates when compared among themselves. Column **B** contains an overview of the distribution of ortholog clusters found exclusively in these isolates, *i.e.*, ortholog clusters that were not detected in any other analyzed *P. syringae* strain.

Surprisingly, however, there is only a single gene that is found in all nine Poaceae isolates, but is absent in any of the other *P. syringae* genomes included in the comparison (Figure 2B and Table S2; see Section 2.4.2). Furthermore, the additional gene shared by eight strains (all except BRIP39023) encodes a methyltransferase, which has no homologs in any other sequenced genome. Moreover, there are only 36 additional genes that are found in at least four of the Poaceae-colonizing strains (Figure 2B and Table S2). Out of these 36, only four are present in the *Por* genome, which is most likely due to the different phylogenetic lineage of this strain. Interestingly, most of these genes are found within four genomic regions that are flanked by sequences with different G + C profiles and transposable

elements (as marked in Table S2). Regions 1 and 2 are present in all of the Clade II strains, except BRIP39023, which makes it very tempting to speculate that the asymptomatic phenotype of this strain could be, at least partially, due to the absence of some of these genes. This is not completely unjustifiable, as these genomic regions encode several proteins that could potentially be involved in the protection from oxidative stress and anti-microbial compounds. These include a catalase-like protein, cytochrome b561 and a small multidrug resistance protein. Additionally, there are also three transcriptional regulators that might be needed for the fine-tuning of gene regulation. The other two genomic regions are present in only four strains, but aside from hypothetical proteins, they encode a putative formaldehyde dehydrogenase, a putative general stress protein, a sensor histidine kinase and an MFS (Major Facilitator Superfamily) transporter. However, this conclusion is purely speculative, and the characterization of deletion mutants of these regions is required to elucidate their function.

2.3. The Type III Secretion System and Effector Repertoire

Based on genome sequences, the analyzed strains possess the canonical *hrp/hrc*-type type III secretion system [26,53], which is found in most of the sequenced *P. syringae* isolates. The effector repertoires for each strain are presented in Figure 3. Based on the number of T3Es, the strains form two distinct groups: the first one consists of *Por*, which has 27 full-length T3Es, while the remaining eight strains only contain around a dozen effectors. The distribution also correlates with the phylogenetic relatedness of the strains (see Figure 1), further supporting the assumption that the host shift has occurred independently for the two groups. Moreover, *Por* has only five effectors shared with *Ppa*, which was also reported to infect rice [48], therefore supporting the idea that the same host can be colonized by strains with very little overlap in effector composition. Interestingly, both strains lack AvrE1. Unfortunately, however, there are no reliable data on T3Es from other rice isolates to further validate these findings.

Apart from *Por*, the remaining eight *P. syringae* strains contain a clear core effector set consisting of AvrE1, HopII, HopAA1, HopM1 and HopBA1, of which the first four are found in intact form in most sequenced *P. syringae* strains [6,31]. The phylogenetic relationship of the core T3Es is shown in Figure S1. The presented individual phylogenetic trees in general reflect the one obtained using the MLST approach (Figure 1). Interestingly, HopBA1 is identical in nearly all strains, suggesting a strong evolutionary pressure to maintain the sequence. Moreover, the remaining effectors appear to have different substitution rates, as seen from the corresponding scale bars. The other most common effectors are HopZ3, HopAZ1 and the HopAG1-HopAH1-HopAI1 cluster. The effector repertoire also correlates with phylogenetic data, as, for example, *Ppa* differs from its closest relative, *Psy* SM, only by the HopAG1-HopAH1-HopAI1 locus. Interestingly, the avirulent strain, BRIP39023, has the largest number of effectors, implying that the number of effectors does not correspond to disease severity and overall virulence. However, it is also possible that one of the effectors acts as an avirulence (Avr) protein, triggering a local hypersensitive response and, therefore, suppressing endophytic growth.

All five currently sequenced *Pcal* strains contain a relatively large number of T3Es, which includes 28 to 32 full-length, as well as five to eleven truncated or disrupted effectors [49]. Out of those, 16 full-length T3Es are shared with *Por*: AvrPto1, HopAA1, HopAB3, HopAD1, HopAO1, HopAQ1, HopAS1, HopBF1, HopD1, HopG1, HopM1, HopQ1, HopR1, HopV1, HopBI1 and HopX1. So far, all

P. syringae strains isolated from oat that were used in scientific studies belong to Clade IV [11,54] and, thus, are likely to have comparable effector repertoires. Assuming that this is true and applies also to the *P. syringae* pv. *coronafaciens* (Clade IV) strains, one might expect that they might be able to infect both monocots and dicots, too. However, these strains were reported to cause disease symptoms on brome, rye and oat, but not on crucifers or tomato [55]. Interestingly, neither *Pcal* nor pathovar *coronafaciens* strains were able to cause disease in wheat [56]. In contrast to *Por*, *Pcal* strains have very little overlap with the Clade II *P. syringae* strains with regard to their T3E repertoire. In addition to the core effectors, AvrE1, HopAA1, HopM1 and HopI1, only HopAL1, HopAZ1 and HopBF1 are found in some *Pcal* strains.

Figure 3. Type III effector repertoires.

	avrE1	hopI1	hopAA1	hopM1	hopBA1	hopAG1	hopAH1	hopAI1	hopAZ1	hopZ3	hopBC1	hopA	hopBF1	hopAL1	hopAE1	hopAX1	hopCI	Total
BRIP39023	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	13
BRIP34876	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	11
BRIP34881	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	11
<i>Psy</i> B64	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	10
<i>Paf</i>	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	10
<i>Ppa</i>	Red	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	9
<i>Pja</i>	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	8
<i>Psy</i> SM	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	7
<i>Por</i> *	Red	Red	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	27
<i>Pcal</i> **	Green	Green	Green	Green	Green	Red	Red	Green	Yellow	Green	Green	Green	Yellow	Green	Red	Green	Green	28-32

Green squares indicate the presence and blank squares indicate the absence of the respective T3Es, and red squares indicate a truncated or disrupted effector. Yellow indicates that an effector is not present in all strains assigned to this species. The total number of effectors is indicated at the end of each row.

* The genome of *Por* contains an additional 20 full-length effector-coding genes not found in any of the eight other analyzed *P. syringae* strains (*hopAQ1*, *avrB1*, *avrB3*, *avrPto1*, *avrPto4*, *hopAD1*, *hopAS1*, *hopX1*, *hopAO1*, *hopD1*, *hopG1*, *hopR1*, *hopV1*, *hopY1*, *hopAB3*, *hopH2*, *hopQ1-1*, *hopQ1-2*, *hopBH1* and *hopB11*).

** Likewise, for the *Pcal* strains, only effectors that are also present in Clade II strains are shown. For a detailed T3E composition of the individual *Pcal* strains, see [49]. It should also be note that some of the genomes of these strains encode a number of the so-called discontinued T3Es, which were never experimentally shown to be translocated in a T3SS-dependent manner: HopAN1, HopAC1, HopJ1, HopL1 and HopAH2.

Having relatively few T3Es is not uncommon for *P. syringae* and seems to be a feature of Clade II strains [31]. However, the strains isolated from Poaceae species are rather extreme examples, and at present, it remains unclear why this is the case. Some T3Es form so-called Redundant Effector Groups (REGs) [38], where members of a REG target different steps in a particular defense signaling pathway. Thus, effectors within one REG are often interchangeable, and the possession of one member of each of the required REGs should be sufficient for full virulence. In support of this notion, the minimal effector repertoire for *Pto* DC3000 on *A. thaliana* has been identified to consist of eight out of

31 T3Es [38], a value which is similar to what is seen in the Poaceae-colonizing strains. Therefore, it is possible that in these hosts, there is a strong selection pressure to maintain just one effector for each REG. On the other hand, there is always the possibility that these genomes contain a number of unrecognized novel T3Es. However, based on results by Baltrus and colleagues [31], there should be very few undiscovered effector families, and it is therefore unlikely that our search missed a considerable number of unknown T3Es.

In addition to the *hrp/hrc*-type T3SS, several *P. syringae* strains were described to possess a second rhizobial-like T3SS, belonging to the *Rhc* family subgroup II [57]. This T3SS-2 was originally identified in *Pph*, *Pta* and *Por*. In *Por*, the gene cluster encompasses 31 genes (POR16_18228–POR16_18383) and appears to contain all essential structural components [57]. Among the other Poaceae isolates, we have identified another T3SS-2 gene cluster in BRIP39023 (A988_16153–A988_16283, accession no. KB316298). Structure-wise, the gene cluster rather resembles its homolog in *Pta* than that of *Por* or *Pph*. It is not, however, known whether this T3SS is active and plays any role in interactions with the host. Moreover, the genomes of both BRIP39023 and *Por* do not contain any known rhizobial effectors, based on sequences from *Rhizobium*, *Sinorhizobium* and *Bradyrhizobium* deposited on the T3DB (Type-III-Secretion-System Related Database) website [58].

2.4. Other Virulence Factors

2.4.1. Phytotoxins and Other Small Secreted Molecules

While the major pathogenicity determinant is the T3SS, many *P. syringae* strains also rely on other secreted molecules for entry and colonization of their hosts. The most important and best-studied group of such compounds are the phytotoxins. There seems to be no general requirement for phytotoxin production, and while some strains are capable of producing several different compounds, others lack any known phytotoxins [19,31]. Moreover, even strains belonging to the same pathovar have differences in phytotoxin gene composition [8,59,60]. Nevertheless, it is a well-established fact that their presence enhances disease progression and symptom development [19,61,62]. In terms of overall distribution, syringolin, syringomycin, syringopeptin and mangotoxin are found almost exclusively among Clade II strains, while all other phytotoxins are found outside of this phylogenetic group [31,63]. So far, the only described exception is *P. syringae* pv. *syringae* CFBP3388, a Clade II strain that was identified to produce both phaseolotoxin and mangotoxin [60,63].

The composition of phytotoxin biosynthetic gene clusters among the Poaceae isolates (Figure 4) fits into the above-mentioned phylogenetic clade-dependent distribution scheme. The genome of *Por*, which belongs to Clade IV, contains gene clusters for the biosynthesis of coronatine and tabtoxin, whereas all other strains have genes for the production of mangotoxin, syringopeptin and syringomycin. However, *Psy* SM seems unable to produce syringomycin, as its *syrE* gene is truncated. It should be noted that the biosynthetic gene clusters for syringopeptin and syringomycin encode very large non-ribosomal peptide synthetases (NRPS) containing repetitive sequences. Because these clusters are often found on several contigs, their complete intactness cannot be assessed with certainty. Complete and intact syringolin synthetase gene clusters, and, thus, the ability for syringolin production, are present in all Clade II strains analyzed, except in *Psy* SM, *Ppa* and BRIP39023. The

first two do not have a syringolin synthetase gene cluster, whereas the latter strain contains a frameshift mutation in the *sylA* gene. Genomes of the *Pcal* strains only contain coronatine biosynthesis genes. None of the analyzed strains is capable of producing phaseolotoxin. In conclusion, there appears to be no specific requirement in terms of toxin composition for being able to colonize Poaceae species.

Some *P. syringae* strains are known to produce phytohormones, such as the auxin indole 3-acetic acid (IAA), cytokines and ethylene [64–66]. While cytokine biosynthesis appears to be specific for *P. syringae* pv. *savastanoi* and ethylene production was detected only in pathovars *pisi*, *cannabina*, *glycinea*, *phaseolicola* and *sesame*, IAA biosynthetic genes appear to be widespread among all major phylogroups [31,66]. Interestingly, none of the Poaceae isolates have homologs of the characterized tryptophan 2-monooxygenase (*iaaM*) and indoleacetamide hydrolase (*iaaH*) genes from *P. syringae* pv. *savastanoi* (*Psv*) NCPPB3335 (accession M11035). However, there appears to be an alternative pathway for IAA biosynthesis, as there are a number of strains where IAA production was demonstrated experimentally, but which lack homologs of the *iaaM^{Psv}/iaaH^{Psv}* genes [67]. The majority of *P. syringae* genomes, including the nine Poaceae isolates, encode another gene annotated as tryptophan 2-monooxygenase at a different genomic location (see accession AAY39694 as a reference). There is also an R-amidase-like protein encoded 15 bp downstream of this gene, and most likely, these two genes constitute an operon. Considering the fact that IaaH also belongs to the amidase family, it is possible that this putative operon is responsible for IAA production in strains lacking the *iaaM^{Psv}/iaaH^{Psv}* homologs. A similar situation is observed for the *Pcal* strains. Apart from that, N-ε-(indole-3-acetyl)-L-lysine synthetase (*iaaL*), which modifies IAA, is present in both the *Por* and *Pcal* strains [68]. The ethylene biosynthesis gene, *efe*, is present only in the genomes of the *Pcal* strains, but not in the Poaceae isolates.

Figure 4. Distribution of phytotoxins among the analyzed strains.

Phytotoxin	BRIP39023	BRIP34876	BRIP34881	<i>P_{sp}</i> B64	<i>P_{ia}</i>	<i>P_{pu}</i>	<i>P_{sp}</i> SM	<i>P_{or}</i>	<i>P_{af}</i>	<i>P_{cal}</i>
Coronatine										
Tabtoxin										
Syringomycin							T			
Syringopeptin										
Syringolin	F									
Mangotoxin										

Green squares indicate the presence, while blank squares indicate the absence of the respective biosynthetic genes. “F” indicates a frameshift in at least one gene, “T” indicates that at least one gene is truncated.

2.4.2. Quorum Sensing

Bacterial cell-to-cell communication is an important part of bacterial physiology that influences many processes, such as growth, differentiation and virulence [69]. One type of such intercellular communication is dependent on cell density and referred to as quorum sensing (QS). The first

described QS system is regulating luminescence in *Vibrio fischeri* [70], and ever since its discovery, QS research continues to be a hot topic, due to its large spectrum of influence. A typical Gram-negative type QS system consists of two proteins encoded by adjacent genes: a LuxI-family synthetase, which produces a signaling compound (in the vast majority of cases, these are acyl-homoserine lactones, AHL) and a LuxR-family transcriptional regulator. Upon perception of the respective AHL, the LuxR protein activates transcription of *luxI*, as well as of other genes that contain a *lux*-type box in their promoters [69]. QS has been shown to be important for a number of plant pathogens, including *Agrobacterium tumefaciens*, *Pantoea stewartii*, *Xanthomonas campestris* and *Erwinia carotovora*, where it regulates processes, such as Ti plasmid conjugation, secondary metabolite production, enzyme and exopolysaccharide (EPS) secretion, etc. [71]. Many *P. syringae* strains also possess a QS system encoded by the *ahlR-ahlI* locus [72], which was demonstrated to be involved in the regulation of epiphytic fitness [71].

Interestingly, none of the analyzed Poaceae-colonizing strains, as well as none of the currently sequenced *Pcal* strains contain *ahlR-ahlI* homologs or any other known quorum sensing system found in Gram-negative bacteria. The trend seems to be general for strains isolated from Poaceae and also extends to pathovars *atrofacies* and *lapsa* [51]. Moreover, strain *Ppa* CFBP 2345 used in that study and *Ppa* LMG2367 are identical [73], thus validating our genomics-based findings. Because of the demonstrated importance of the QS for a number of pathogens, its loss might be expected to have negative consequences on leaf colonization. However, since the lack of the canonical QS system is relatively abundant among *P. syringae* group strains [51], it is likely that the AHL-mediated gene regulation has no significant effect on the T3SS in these strains. As to why the lack of QS appears to be a general pattern for Poaceae isolates, this could be connected to the fact that some groups of AHLs are actually recognized by plant cells and induce changes in gene expression, including defense genes [74]. Moreover, exposure to AHL was shown to be an inducer of systemic acquired resistance in barley [75]. The AHL type used in that study was different from the one produced by *P. syringae*. However, it is likely that the recognition spectrum is broader, thus making AHL producers less virulent and giving a selective advantage to those which have lost the respective genes.

Bacterial genomes often encode orphan LuxR-type transcriptional regulators, which are not associated with a cognate LuxI partner. Some were shown to bind signaling molecules originating not only from cells of the same species, but also coming from other bacteria and even from eukaryotic organisms [74,76,77]. There also exists a distinct clade of the orphan LuxR proteins identified in several groups of plant-associated bacteria, which were reported to have an affinity towards plant-borne low-molecular weight compounds [78]. For example, the LuxR family member OryR from *X. oryzae* pv. *oryzae* was identified to be a global regulator and a virulence factor that responds to the presence of homogenized plant material, but not to AHLs [79,80]. With the exception of BRIP39023, all other Clade II Poaceae isolates have a gene encoding an orphan LuxR family protein that is not found in any other *P. syringae* genome (see Table S2). This gives room for speculation that this protein could be involved in the detection of plant signals that are associated with Poaceae. However, unlike other such sensor proteins, it is not encoded adjacent to a proline immunopeptidase *pip* [74]. The protein shares 78% sequence similarity with the syringomycin biosynthesis regulator, SyrG, of *Psy* B301D [81]. Syringomycin production is known to be enhanced by p-arbutin [82], which is a glycosylated hydroquinone found in plant tissues and which also appears to be relatively abundant in wheat [83]. It is

not known whether SyrG actually binds this compound; however, it seems possible that these proteins are able to interact with plant-borne phenolic compounds.

Additionally, the only identified unique gene that is shared among all nine genomes belongs to the LysE/RhtB family of efflux proteins, which are hypothesized to be involved in the export of AHLs [84]. One proposed function for this family of exporters is protection from interference by AHLs produced from other bacterial species. Interestingly, two recent studies showed the production of AHL-mimicking substances in rice and *Medicago truncatula* [85,86], and therefore, it is possible that such a protein could be involved in the protection of the bacteria from these compounds.

2.4.3. Exopolysaccharides

P. syringae strains are known to produce two types of EPS: levan and alginate [87]. While levan is a branched β -polyfructan [87], alginate consists of blocks of non-repeating α -L-glucuronate and β -D-mannuronate, with the latter residue being randomly acetylated [88]. Levan is produced from sucrose by the action of a single enzyme, called levansucrase [89]. *P. syringae* genomes usually contain more than one levansucrase gene located at different genomic sites [89]. In contrast, alginate biosynthesis is a multi-step process, which involves several enzymes [88]. Respective genes are organized in a gene cluster and have an identical architecture in both *P. syringae* and *P. aeruginosa* [90]. One of the major roles of EPS is thought to be protection from desiccation. Furthermore, while alginate was shown to be involved in epiphytic fitness and osmotolerance [22,90], levan is presumed to be a storage polymer [91]. In addition, most *P. syringae* strains have homologs of genes from *P. aeruginosa* involved in the biosynthesis of another EPS called Psl. Psl, a complex polysaccharide with repeating pentamer units of D-glucose, D-mannose and L-rhamnose [92], was demonstrated to be required for biofilm formation in *P. aeruginosa* [93].

When compared with one another, all Poaceae isolates possess the alginate biosynthesis cluster (*algD*-*algA*), as well as all other genes involved in its biosynthesis and regulation (*algP*, *algQ*, *algR*, *algZ*, *algC*, *algB*, *kinB*, *algH*, *algU*, *mucA*, *mucB* and *mucD*). With respect to levan production, all strains were identified to have the *lscA* gene, which, however, was shown not to contribute to levan production [89]. With the exception of *Por*, all other genomes contained a copy of the *lscC* gene. In *Paf*, both *lscA* and *lscC* are located on two separate contigs, and thus, it is not possible to tell whether the two genes are completely intact. None of the strains had a plasmid-born *lscB* gene. Thus, among the analyzed strains, only *Por* appears to be levan-negative. Last, with respect to *psl* gene content, *P. syringae* is slightly different from *P. aeruginosa* and lacks a homolog of the acyltransferase *pslL*. In addition, the *pslC* gene is located outside of the main gene cluster. However, all nine genomes contain an additional gene between the *pslJ* and *pslK* homologs, which encodes a putative maltose O-acyltransferase family protein. Since the *psl* gene cluster was shown to be transcriptionally active [94], it is possible that *P. syringae* is producing a slightly different version of this EPS. The analyzed *Pcal* strains appear to contain genes for the biosynthesis of both alginate and Psl. A single levansucrase-coding gene is present in all of these strains, with the exception of ES4326, which did not give any hits in our BLAST (Basic Local Alignment Search Tool) search.

2.4.4. Type VI Secretion System

The Type VI secretion system (T6SS) is a multipurpose protein delivery machinery encoded as a single gene cluster [95]. It is found in the genomes of Gram-negative bacteria, both pathogenic and non-pathogenic, with the highest abundance among Proteobacteria species [95,96]. It is not uncommon that a single genome contains more than one T6SS gene cluster [96,97]. However, different T6SS are found within one genome function independently, and there is no overlap in terms of the translocated effector repertoire [98,99]. In several species, the T6SS was identified to be a pathogenicity factor or at least a virulence enhancer. However, its function is not limited to that, and it was also shown to be involved in inter-species competition, protection from predation and biofilm formation [96]. The first identified effectors for this secretion system were Hcp and VgrG family proteins, which are also structural components of the injection apparatus and are encoded within and outside the T6SS gene cluster [96]. A recent study on *Pto* DC3000 T6SS gene clusters (or HSI, Hcp secretion islands) showed that the Hcp-2 protein encoded in the HSI-2 gene cluster plays a role in competition with other bacterial species, but not in virulence [16].

All nine genomes of Poaceae-colonizing strains contain two individual T6SS gene clusters, which belong to the Group 1 and Group 4B loci based on the homology of the respective ClpV proteins, as classified by Barret and colleagues [97]. However, the *tssB* gene in the T6SS-2 gene cluster (a Group 4B member) of *Pja* contains a potential frameshift. Both types of T6SS are found in most other sequenced *P. syringae* strains [97] and, at least in the case of Poaceae isolates, are located within the same genomic regions. Thus, it appears that these T6SS gene clusters were acquired long ago. A detailed analysis of the T6SS gene clusters from plant-pathogenic bacteria has been performed by Sarris and colleagues [100,101], which also depicts the phylogenetic position of both T6SS-1 and T6SS-2 from *Por* and, thus, of their respective homologs from the other analyzed Poaceae isolates. It is notable that none of the genomes encodes a homolog of the above-mentioned Hcp-2 from *Pto* DC3000.

While the HSIs remain conserved, the number of putative orphan effectors varies between the strains. For example, the genome of *Psy* SM encodes five orphan Hcp and three orphan VgrG proteins; *Psy* B64 has only two such Hcp and three such VgrG-coding genes; whereas the genome of *Ppa* encodes for three members of the Hcp family and five VgrG family members located outside the HSIs. The orphan VgrG family proteins sometimes carry an additional domain at their C-terminus and are referred to as “evolved” VgrG proteins. Some of these domains have been described to have a catalytic function, such as actin crosslinking [96,102]. Among the analyzed genomes, two VgrG proteins could potentially belong to this group: VgrG-3 of *Psy* SM (PssSM_4133) has an additional domain of unknown function (COG4253), while the C-terminus of VgrG-4 (PssSM_4495) has a weak homology to the anti-sigma factor FecI-like domain. Both proteins are well conserved among *P. syringae* isolates, and in addition, the first one is also found in *P. putida* and *X. oryzae*, while the second one is present in some *P. aeruginosa* strains.

Another group of known T6SS effector encompasses the so-called Type VI lipase effectors (Tle), which are often encoded downstream of VgrG family proteins [103]. Among the analyzed genomes, there are six that contain genes homologous to PSPTO_5055, which is a member of the Tle3 family: A988_02933 from BRIP39023, POR16_12266 from *Por*, PssB64_01780 from *Psy* B64,

PSYJA_00260/PSYJA_00265 from *Pja*, as well as unannotated genes located on contig 283 of the *Paf* genome (accession no. AWUI01000283) and on contig 6 of the *Ppa* genome (accession no. ALAC01000006). All three genes have an overlapping VgrG coding sequence upstream, as well as another ORF (Open Reading Frame) immediately downstream, which could be an immunity protein (based on the architecture of the other Tle3 gene clusters) [103]. Notably, A988_02933 contains the conventional GxSxG motif described for the family (residues 237–241), while in all other proteins, the motif is AxSxG (residues 238–242). In addition to that, the genomes of *Por* and BRIP39023 contain putative members of the Tle5 family, POR16_25160 and A988_18422, which are, however, only ~67% homologous to the described family member, PLA107_28855, from *P. syringae* pv. *lachrymans* M301315 [103]. Nevertheless, both proteins are found within the same genomic region and have a VgrG family protein encoded nearby (POR16_25135 and A988_18442, respectively). Moreover, both of them have a perfect dual HxKxxxxD motif characteristic of phospholipase D and Tle5 family members. In addition, the genome of *Ppa* contains a gene coding for a less related putative Tle5 family protein, located adjacent to one of the T6SS gene clusters (contig 27, accession no. ALAC01000027). Tle family proteins were recently shown to have antibacterial activity, and thus, they could be used by these strains to compete with epiphytic bacteria. No homologs of Tse1, Tse2 or Tse3 from *P. aeruginosa* [104] were detected in any of the Poaceae isolates.

2.5. Mobile and Extrachromosomal Elements

P. syringae genomes contain a relatively large number of mobile genetic elements (MGE) belonging to various families. The specific composition and the number of MGEs tend to differ from strain to strain [43]. More importantly, MGEs are often associated with pathogenicity islands, thus allowing their spread from one cell to another [105]. It is also known that strains have overcome race-specific resistance of their hosts after the respective avirulence gene was inactivated by an MGE insertion [106]. The heterogeneity in MGE composition is also observed for the Poaceae isolates: for example, the genome of *Psy* SM contains at least 21 transposases belonging to nine different families: ISPsy9, ISPpu10, ISPsy6, ISPpu14, ISPsy1, ISRso10, ISPsy5, ISPsy24 and IS200. In comparison, the *Psy* B64 genome encodes at least 22 transposases. The majority of these belong to the ISPsy9 and ISPpu10 families, whereas the remaining ones, with the exception of IS200, represent families not found in *Psy* SM: ISPsy7, ISPsy26, ISPsy5 and ISPsy22. In addition, the genomes encode a number of phage-type integrases/recombinases, the majority of which are truncated. A large portion of the transposable elements have at least one of their ORFs truncated, thus rendering them inactive. However, a number of transposons are present as partial sequences found on contig ends, which makes it impossible to judge their integrity. Moreover, the presence of sequence gaps only allows giving an estimate of their copy number.

Aside from simple insertional elements, the genomes also contain several prophages of variable size. While some of the prophages are very small and contain about 20 ORFs (such as prophage PSSSM-04, PssSM_4692–4712), others are over 57 kb-long (for example, prophage PSSSM-02, PssSM_2181–2256). Interestingly, prophage PSSSM-04 is the only prophage that is common to all Poaceae isolates. Prophages often serve as preferred sites of integration for other MGEs, as well as for genes acquired by horizontal transfer. One such example is prophage PSSB64-02, which has an

additional ORF (PssB64_01273) present in comparison to its homolog, PSSSM-04. A more extreme example is prophage PSSSM-03, which contains a number of regions with a differing G + C profile, which is a signature of horizontal gene transfer (PssSM_4287–4288, 4292–4296, 4304–4308, 4312–4313 and 4316–4318). None of the prophages seems to be complete. However, since some components could be shared between different prophages, it is not possible to exclude that at least one is capable of forming functional viral particles. It is noteworthy that the prophage PSSSM-02 contains *recT* and *recE* genes (PssSM_2197 and PssSM_2196, respectively), which encode homologs of the lambda Red Exo/Beta proteins. The RecTE system has been successfully deployed for the recombineering of linear DNA fragments into the *P. syringae* genome [107].

Another interesting type of mobile element present in bacteria are the so-called integrative and conjugative elements (ICEs). These elements are always localized within a set of specific sites within a genome (the so-called *att* sites) and are flanked by direct repeats. They contain genes required for integration and excision and are presumably self-transmittable [108]. One of the first discovered ICEs was the *clc* element from *Pseudomonas sp.* B13, which encodes enzymes required for 3-chlorobenzoate degradation [109]. The ICEs were demonstrated to serve as vectors for the spreading of virulence factors, drug and metal resistance genes and other beneficial metabolic traits [110,111].

Several examples of ICEs were described for *P. syringae*, as well. The first one characterized was PPHGI-1 from *P. syringae* pv. *phaseolicola* 1302A [112]. This genomic region contains a gene encoding the Type III effector, *avrPhpB* (*hopARI*), and was shown to excise itself upon infiltration into a resistant cultivar of bean. In addition to being located between two tRNA genes and having a recombinase gene required for its excision, it also contains a gene cluster for the biogenesis of conjugative pili. Interestingly, it also contains a copy of the two syringomycin biosynthesis regulator genes, *sala* and *syrF* (*salC*). A similar ICE was also identified in the genome of *Psy* B728a, where a part of it was replaced by an arsenic and copper resistance gene cluster, and in *Pto* DC3000, where only a part of the ICE is present [112]. A similar ICE was recently described for a number of *P. syringae* pv. *actinidiae* strains [6].

The genomes of all nine Poaceae isolates have two well-conserved *att* sites, each within a distinct tRNA-Lys gene. However, only seven of the genomes contain an integrated ICE (Table S3A). The exceptions are *Psy* B64 and *Paf*, where the genomic island likely has excised itself at some point after they diverged from *Pja*. Interestingly, several different types of ICEs were identified, which are located in either of the two potential integration sites: in *Ppa* and *Psy* SM, the ICE was adjacent to the *queC* gene, while in all other strains, it was located next to the *clpB* gene. The second observation was that most of the shared Poaceae isolate-specific genes were found adjacent to both *att* sequences, but not associated with the ICE. In general, among the different groups of ICEs, all structural genes were highly conserved, and the observed differences were mainly found within a variable middle part between the *pil* gene cluster and the topoisomerase III gene. The majority of the variable regions encoded heavy metal resistance genes: the ICEs of BRIP34876 and BRIP34881 contain mercury and arsenate resistance operons; *Pja* and *Ppa* have copper and arsenate resistance gene clusters, while in the case of *Psy* SM, chromate, nickel/cobalt and bacteriocin/lantibiotic efflux transporters were found. The ICE of *Por* did not contain any heavy metal resistance genes and was rather similar to the one found in pv. *actinidiae* strain NZ V-13 [6]. However, it must be noted that the respective *Por* scaffold has a number of gaps and appears to be rather misassembled, and thus, it is possible that some genes

were missed. The ICE of BRIP39023, in contrast to the other six ICEs, is relatively similar to PPHGI-1. However, it contains a number of regions with no homology to PPHGI-1, namely a unique region with a number of sensor histidine kinases and transcriptional regulators (A988_21582–A988_21632), as well as another locus surrounding a poly(β -D-mannuronate)-C5-epimerase (A988_21477). Notably, the first locus encodes the T3E HopBF1 (A988_21602), while the second locus contains a truncated copy of another T3E, HopBA1-2 (A988_21482). In addition, it appears that the levansucrase C locus, which is normally located downstream of the respective *att* site, has jumped into the *syrF* homolog, resulting in its duplication. From this comparison, it appears that either there is a “population” of ICEs with distinct variable regions circulating among *P. syringae* strains, or this region is a recombination hotspot, which would also account for the observed diversity.

The last type of mobile genetic elements discussed here are plasmids. *P. syringae* strains often have one or more plasmid, some of which also encode T3Es [40,43,113], or other virulence factors, such as coronatine [8,114] or phytohormone biosynthesis enzymes [115]. Of the nine Poaceae isolates, four were identified to contain a plasmid based on the presence of a replicase gene adjacent to UV-resistance genes *rulA/rulB* (Table S3B). All four identified plasmids belong to the pPT32A family, which is frequently found in *P. syringae* strains [113,115]. In addition, *Psy* SM appears to have remnants of an integrated plasmid in its genome, which encodes a non-homologous replicase with a 78% similarity to the RepA protein of pPT14-32 from *P. syringae* PT14.

The identified plasmids have a complete or nearly complete Type IV-A secretion system (*virB1–virB11* and *virD4* genes [116]). No homologs of *tra/trb* genes from pDC3000B (accession AE016854) or *mob* genes were detected. Type IV secretion systems are frequently found on *P. syringae* plasmids; however, whether they play any role in virulence or are only involved in the spreading of the plasmids is yet to be determined [113]. Based on the current annotation, most of the proteins encoded on the detected plasmids are hypothetical, and no virulence genes were detected.

2.6. Other Notable Genome Components

2.6.1. Defense Mechanisms against Foreign DNA

The CRISPR/Cas (clustered regularly interspaced short palindromic repeats/CRISPR-associated genes) system provides a wide-spread defense mechanism against bacteriophages. CRISPR repeat sequences contain fragments of phage DNA (the so-called spacers), which, upon transcription and processing, are involved in guiding Cas proteins, which, in turn, recognize and cleave the genome of an invading phage [117]. The system is adaptive, and new spacers can be acquired upon encountering novel invaders [118]. Among the genomes of Poaceae-colonizing strains, only *Psy* SM was identified to have two putative CRISPRs with more than one spacer sequence (CRISPR 1: 1,828,342–1,828,605; CRISPR 2: 1,828,678–1,829,372; both are inside the *inaZ* gene). However, no putative Cas proteins, which are expected to be encoded in the vicinity [119], were detected. It should be noted that *Por* and *Pja* were excluded from the search, due to a high level of genome fragmentation, which would not allow the detection of direct repeats. In addition, no PIWI-domain proteins, the analogs of the RNA interference complex protein, Argonaute [120], were identified in any of the genomes.

2.6.2. Bacteriocins

Bacteriocins are ribosomally synthesized peptides, which might or might not be modified, that show bactericidal activity, usually against a narrow range of species. The producing bacterial species is unaffected by its own bacteriocins, either due to the presence of an immunity protein or by other mechanisms [121]. The most common targets of bacteriocins are the cell envelope [122], but they can also interfere with DNA, RNA and protein metabolism [121]. Among the target genomes, several bacteriocin-coding genes were identified (Table 2). From those, the product of only one gene is currently characterized. This bacteriocin is called syringacin M and was first described in *Pto* DC3000. It was shown to induce the death of susceptible bacteria by the inhibition of peptidoglycan biosynthesis through degradation of lipid II [123].

Table 2. Putative bacteriocin genes.

Strain	Bacteriocin	Location	Strain	Bacteriocin	Location
BRIP34876	Syringacin M	A979_09507	<i>Psy</i> SM	S-type pyocin	PssSM_0863
BRIP34881	Syringacin M	A987_14290		Colicin-DNAse	PssSM_0293 ¹
BRIP39023	S-type pyocin	A988_14544		Bacteriocin	PssSM_4261
	Colicin-DNAse	A988_07204 ¹	<i>Ppa</i>	Syringacin M	ALAC01000012
	Cyclized peptide	A988_02638 ²		S-type pyocin	ALAC01000003
<i>Pja</i>	Syringacin M	PSYJA_15707	<i>Psy</i> B64	Syringacin M	PssB64_01273
	Bacteriocin	PSYJA_00994		S-type pyocin	PssB64_04860
	Lasso peptide	PSYJA_17421	<i>Por</i>	Colicin-DNAse	POR16_04054 ¹
	Colicin-DNAse	AEA01000987 ^{1,3}		Bacteriocin	POR16_27431
<i>Paf</i>	Syringacin M	AWUI01000310		Lasso peptide	DS996947 ⁴
	S-type pyocin	AWUI01000345			

¹ Immunity protein-coding gene located nearby. ² ORF (Open Reading Frame) located inside the specified gene on the opposite strand. ³ The gene contains a potential frameshift. ⁴ ORF is not annotated, located between POR16_00075 and POR16_00080.

2.6.3. Non-Ribosomal Peptide Synthetases/Polyketide Synthetases

Non-ribosomal peptide synthetases (NRPS) and polyketide synthases (PKS) are large multi-modular proteins with enzymatic function found in bacteria and fungi. These proteins are involved in the biosynthesis of secondary metabolites, such as antibiotics, siderophores, toxins and other bioactive compounds [124]. In *P. syringae*, with the exception of phaseolotoxin and tabtoxin, all other phytotoxins, including syringolin, are produced by NRPS/PKS-type enzymes [19,125,126]. Other NRPS/PKS gene clusters, which were found in all nine isolates, were genes for the biosynthesis of the siderophore, pyoverdine, and the surfactant, syringofactin. No yersiniabactin biosynthetic cluster was identified in any of the strains. Among NRPS/PKS genes, which were shown to be upregulated *in planta* [127], only homologs of *Psy*_3722 are found in all strains, whereas a homolog of the PKS gene cluster, *Psy*_4311–4315, is exclusively found in *Ppa*. As the products of these genes are unknown, it is currently not possible to draw a conclusion regarding their impact on virulence.

3. Experimental Section

Genomic sequences used in this study were downloaded from the NCBI (National Center for Biotechnology Information) FTP (File Transfer Protocol) server [128]. Accession numbers of the used genomic sequences are listed in Table S1. The GenBank entries for *Ppa* and *Paf* only contain nucleotide sequence data. Therefore, the genomes were re-annotated using the RAST (Rapid Annotation using Subsystem Technology) server [129] and Prokka [130]. In order to avoid any conflicts, locus_tag qualifiers were set to “panici” and “Paf”, respectively. The phylogenetic tree was generated with MEGA (Molecular Evolutionary Genetics Analysis) version 5.2.2 [131] with 1,000 bootstraps using nucleotide sequences of seven conserved genes: *gyrB*, *gap-1*, *fruK*, *pgi*, *rpoD*, *acnB* and *gltA*. Phylogenetic trees of core Type III effectors were also generated in MEGA 5.2.2 using respective protein sequences and the neighbor-joining method with 1,000 bootstraps. Ortholog cluster searches were performed with the Pan-genome analysis pipeline [132] using the MultiParanoid method with the cut off value set to 80%, an e-value of 10^{-9} and, otherwise, default parameters. Visualization of annotated genomic sequences was done using Artemis [133]. Homology searches on a gene-to-gene basis were performed using tBLASTN and BLASTP [134,135] against the non-redundant protein sequence database (nr) and other custom databases. The Type III effector sequence database was downloaded from the *Pseudomonas syringae* Genome Resources website [136]. The T3E and phytotoxin compositions of *Por* and *Pja* are based on Baltrus *et al.* [31] and Mucyn *et al.* [137], while the corresponding data for the *Pcal* strains were obtained from Sarris *et al.* [49]. Additional analysis of the *Pcal* strains was performed using IMG/ER (Integrated Microbial Genomes/Expert Review) [138]. Protein domain analysis was done using the NCBI Conserved Domain Database [139]. Bacteriocin prediction was done using BAGEL3 software [140]. The CRISPR search was performed using the CRISPRFinder tool [141].

4. Conclusions

Comparative genomics is a powerful tool for the discovery of virulence traits and potential host-specific adaptations. Here, we identified that *P. syringae* strains isolated from Poaceae belong to two different phylogenetic lineages, with each group evolving independently. All of these strains possess the canonical Type III secretion system, but in contrast to the majority of *P. syringae* strains, most Poaceae isolates have a relatively small repertoire of T3Es. Two strains from different phylogenetic lineages were identified to have a second T3SS similar to that of *Rhizobia*. Moreover, in line with previous studies, we have shown that strains pathogenic to the same host could have almost no overlap in T3E composition. In addition, several other putative adaptations were identified, such as a lack of quorum sensing and a number of genes that could be involved in the removal of plant anti-microbial compounds. Moreover, several traits important for leaf surface colonization, such as Type VI secretion systems and bacteriocins, are described. Lastly, the strains were compared with a non-specialized pathogen *P. cannabina* pv. *alisalensis*, which is able to grow on both dicot and monocot plants. The analyses provide a foundation for further experimental validation.

Acknowledgments

This project was supported by the Swiss National Science Foundation grant 31003A-134936 and the Foundation for Research in Science and the Humanities at the University of Zurich to R.D. The support of the Functional Genomics Center, Zurich, is acknowledged.

Author Contributions

Alexey Dudnik collected and analysed the data. Alexey Dudnik and Robert Dudler conceived and designed the study and wrote the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Bradbury, J.F. *Pseudomonas syringae* pv. *syringae*. In *Guide to Plant Pathogenic Bacteria*; CAB International Mycological Institute: Kew, England, 1986; pp. 175–177.
2. Gironde, S.; Manceau, C. Housekeeping gene sequencing and multilocus variable-number tandem-repeat analysis to identify subpopulations within *Pseudomonas syringae* pv. *maculicola* and *Pseudomonas syringae* pv. *tomato* that correlate with host specificity. *Appl. Environ. Microbiol.* **2012**, *78*, 3266–3279.
3. Martín-Sanz, A.; de la Vega, M.P.; Murillo, J.; Caminero, C. Strains of *Pseudomonas syringae* pv. *syringae* from pea are phylogenetically and pathogenically diverse. *Phytopathology* **2013**, *103*, 673–681.
4. Ramos, C.; Matas, I.M.; Bardaji, L.; Aragón, I.M.; Murillo, J. *Pseudomonas savastanoi* pv. *savastanoi*: some like it knot. *Mol. Plant Pathol.* **2012**, *13*, 998–1009.
5. Arnold, D.L.; Lovell, H.C.; Jackson, R.W.; Mansfield, J.W. *Pseudomonas syringae* pv. *phaseolicola*: from “has bean” to supermodel. *Mol. Plant Pathol.* **2011**, *12*, 617–627.
6. McCann, H.C.; Rikkerink, E.H.A.; Bertels, F.; Fiers, M.; Lu, A.; Rees-George, J.; Andersen, M.T.; Gleave, A.P.; Haubold, B.; Wohlers, M.W.; et al. Genomic Analysis of the Kiwifruit Pathogen *Pseudomonas syringae* pv. *actinidiae* Provides Insight into the Origins of an Emergent Plant Disease. *PLoS Pathog.* **2013**, *9*, e1003503.
7. Green, S.; Studholme, D.J.; Laue, B.E.; Dorati, F.; Lovell, H.; Arnold, D.; Cottrell, J.E.; Bridgett, S.; Blaxter, M.; Huitema, E.; et al. Comparative genome analysis provides insights into the evolution and adaptation of *Pseudomonas syringae* pv. *aesculi* on *Aesculus hippocastanum*. *PLoS One* **2010**, *5*, e10224.
8. 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*, e16451.
9. Young, J.M. Taxonomy of *Pseudomonas syringae*. *J. Plant Pathol.* 2010, *92*, S1.5–S1.14.
10. Sarkar, S.F.; Guttman, D.S. Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. *Appl. Environ. Microbiol.* **2004**, *70*, 1999–2012.

11. Hwang, M.S.H.; Morgan, R.L.; Sarkar, S.F.; Wang, P.W.; Guttman, D.S. Phylogenetic characterization of virulence and resistance phenotypes of *Pseudomonas syringae*. *Appl. Environ. Microbiol.* **2005**, *71*, 5182–5191.
12. Gardan, L.; Shafik, H.; Belouin, S.; Broch, R.; Grimont, F.; Grimont, P.A.D. DNA relatedness among the pathovars of *Pseudomonas syringae* and description of *Pseudomonas tremiae* sp. nov. and *Pseudomonas cannabina* sp. nov. (ex Sutic and Dowson 1959). *Int. J. Syst. Bacteriol.* **1999**, *49*, 469–478.
13. O'Brien, H.E.; Thakur, S.; Gong, Y.; Fung, P.; Zhang, J.; Yuan, L.; Wang, P.W.; Yong, C.; Scortichini, M.; Guttman, D.S. Extensive remodeling of the *Pseudomonas syringae* pv. *avellanae* type III secretome associated with two independent host shifts onto hazelnut. *BMC Microbiol.* **2012**, *12*, 141.
14. Lindeberg, M.; Cunnac, S.; Collmer, A. The evolution of *Pseudomonas syringae* host specificity and type III effector repertoires. *Mol. Plant Pathol.* **2009**, *10*, 767–775.
15. Rodríguez-Palenzuela, P.; Matas, I.M.; Murillo, J.; López-Solanilla, E.; Bardaji, L.; Pérez-Martínez, I.; Rodríguez-Mosquera, M.E.; Penyalver, R.; López, 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.
16. Haapalainen, M.; Mosorin, H.; Dorati, F.; Wu, R.-F.; Roine, E.; Taira, S.; Nissinen, R.; Mattinen, L.; Jackson, R.; Pirhonen, M.; *et al.* Hcp2, a secreted protein of the phytopathogen *Pseudomonas syringae* pv. *tomato* DC3000, is required for fitness for competition against bacteria and yeasts. *J. Bacteriol.* **2012**, *194*, 4810–4822.
17. Lindeberg, M.; Cunnac, S.; Collmer, A. *Pseudomonas syringae* type III effector repertoires: Last words in endless arguments. *Trends Microbiol.* **2012**, *20*, 199–208.
18. Schellenberg, B.; Ramel, C.; Dudler, R. *Pseudomonas syringae* Virulence Factor Syringolin A Counteracts Stomatal Immunity by Proteasome Inhibition. *Mol. Plant-Microbe Interact.* **2010**, *23*, 1287–1293.
19. Bender, C.L.; Alarcón-Chaidez, F.; Gross, D.C. *Pseudomonas syringae* phytotoxins: Mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. *Microbiol. Mol. Biol. Rev. MMBR* **1999**, *63*, 266–292.
20. Denny, T.P. Involvement of bacterial polysaccharides in plant pathogenesis. *Annu. Rev. Phytopathol.* **1995**, *33*, 173–197.
21. Misas-Villamil, J.C.; Kolodziejek, I.; Crabill, E.; Kaschani, F.; Niessen, S.; Shindo, T.; Kaiser, M.; Alfano, J.R.; van der Hoorn, R.A.L. *Pseudomonas syringae* pv. *syringae* uses proteasome inhibitor syringolin A to colonize from wound infection sites. *PLoS Pathog.* **2013**, *9*, e1003281.
22. Yu, J.; Penaloza-Vazquez, A.; Chakrabarty, A.M.; Bender, C.L. Involvement of the exopolysaccharide alginate in the virulence and epiphytic fitness of *Pseudomonas syringae* pv. *syringae*. *Mol. Microbiol.* **1999**, *33*, 712–720.
23. Rodríguez-Moreno, L.; Barceló-Muñoz, A.; Ramos, C. In vitro analysis of the interaction of *Pseudomonas savastanoi* pvs. *savastanoi* and *nerii* with micropropagated olive plants. *Phytopathology* **2008**, *98*, 815–822.

24. Arrebola, E.; Cazorla, F.M.; Perez-García, A.; de Vicente, A. Chemical and metabolic aspects of antimetabolite toxins produced by *Pseudomonas syringae* pathovars. *Toxins (Basel)*. **2011**, *3*, 1089–1110.
25. Galán, J.E.; Wolf-Watz, H. Protein delivery into eukaryotic cells by type III secretion machines. *Nature* **2006**, *444*, 567–573.
26. Büttner, D. Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria. *Microbiol. Mol. Biol. Rev.* **2012**, *76*, 262–310.
27. Jones, J.D.G.; Dangl, J.L. The plant immune system. *Nature* **2006**, *444*, 323–329.
28. Cunnac, S.; Lindeberg, M.; Collmer, A. *Pseudomonas syringae* type III secretion system effectors: Repertoires in search of functions. *Curr. Opin. Microbiol.* **2009**, *12*, 53–60.
29. Munkvold, K.R.; Martin, M.E.; Bronstein, P.A.; Collmer, A. A survey of the *Pseudomonas syringae* pv. *tomato* DC3000 type III secretion system effector repertoire reveals several effectors that are deleterious when expressed in *Saccharomyces cerevisiae*. *Mol. Plant-Microbe Interact.* **2008**, *21*, 490–502.
30. Salomon, D.; Bosis, E.; Dar, D.; Nachman, I.; Sessa, G. Expression of *Pseudomonas syringae* type III effectors in yeast under stress conditions reveals that HopX1 attenuates activation of the high osmolarity glycerol MAP kinase pathway. *Microbiology* **2012**, *158*, 2859–2869.
31. Baltrus, D.A.; Nishimura, M.T.; Romanchuk, A.; Chang, J.H.; Mukhtar, M.S.; Cherkis, K.; Roach, J.; Grant, S.R.; Jones, C.D.; *et al.* Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathog.* **2011**, *7*, e1002132.
32. Göhre, V.; Spallek, T.; Häweker, H.; Mersmann, S.; Mentzel, T.; Boller, T.; de Torres, M.; Mansfield, J.W.; Robatzek, S. Plant pattern-recognition receptor FLS2 is directed for degradation by the bacterial ubiquitin ligase AvrPtoB. *Curr. Biol.* **2008**, *18*, 1824–1832.
33. Nicaise, V.; Joe, A.; Jeong, B.-R.; Korneli, C.; Boutrot, F.; Westedt, I.; Staiger, D.; Alfano, J.R.; Zipfel, C. *Pseudomonas* HopU1 modulates plant immune receptor levels by blocking the interaction of their mRNAs with GRP7. *EMBO J.* **2013**, *32*, 701–712.
34. Rodríguez-Herva, J.J.; González-Melendi, P.; Cuartas-Lanza, R.; Antúnez-Lamas, M.; Río-Alvarez, I.; Li, Z.; López-Torrejón, G.; Díaz, I.; del Pozo, J.C.; Chakravarthy, S.; *et al.* A bacterial cysteine protease effector protein interferes with photosynthesis to suppress plant innate immune responses. *Cell. Microbiol.* **2012**, *14*, 669–681.
35. Bhattacharjee, S.; Halane, M.K.; Kim, S.H.; Gassmann, W. Pathogen effectors target *Arabidopsis* EDS1 and alter its interactions with immune regulators. *Science* **2011**, *334*, 1405–1408.
36. Lee, A.H.-Y.; Hurley, B.; Felsensteiner, C.; Yea, C.; Ckurshumova, W.; Bartetzko, V.; Wang, P.W.; Quach, V.; Lewis, J.D.; Liu, Y.C.; *et al.* A bacterial acetyltransferase destroys plant microtubule networks and blocks secretion. *PLoS Pathog.* **2012**, *8*, e1002523.
37. Jiang, S.; Yao, J.; Ma, K.-W.; Zhou, H.; Song, J.; He, S.Y.; Ma, W. Bacterial Effector Activates Jasmonate Signaling by Directly Targeting JAZ Transcriptional Repressors. *PLoS Pathog.* **2013**, *9*, e1003715.

38. Cunnac, S.; Chakravarthy, S.; Kvitko, B.H.; Russell, A.B.; Martin, G.B.; Collmer, A. Genetic disassembly and combinatorial reassembly identify a minimal functional repertoire of type III effectors in *Pseudomonas syringae*. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 2975–2980.
39. Almeida, N.F.; Yan, S.; Lindeberg, M.; Studholme, D.J.; Schneider, D.J.; Condon, B.; Liu, H.; Viana, C.J.; Warren, A.; Evans, C.; *et al.* A draft genome sequence of *Pseudomonas syringae* pv. *tomato* T1 reveals a type III effector repertoire significantly divergent from that of *Pseudomonas syringae* pv. *tomato* DC3000. *Mol. Plant Microbe Interact.* **2009**, *22*, 52–62.
40. 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.; *et al.* The complete genome sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. *tomato* DC3000. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 10181–10186.
41. Wei, C.-F.; Kvitko, B.H.; Shimizu, R.; Crabill, E.; Alfano, J.R.; Lin, N.-C.; Martin, G.B.; Huang, H.-C.; Collmer, A. A *Pseudomonas syringae* pv. *tomato* DC3000 mutant lacking the type III effector HopQ1-1 is able to cause disease in the model plant *Nicotiana benthamiana*. *Plant J.* **2007**, *51*, 32–46.
42. 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.
43. 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.
44. Dudnik, A.; Dudler, R. High-Quality Draft Genome Sequence of *Pseudomonas syringae* pv. *Syringae* Strain SM, Isolated from Wheat. *Genome Announc.* **2013**, *1*, e00610–e00613.
45. Dudnik, A.; Dudler, R. Non contiguous-finished genome sequence of *Pseudomonas syringae* pathovar *syringae* strain B64 isolated from wheat. *Stand. Genomic Sci.* **2013**, *8*, 420–429.
46. Gardiner, D.M.; Stiller, J.; Covarelli, L.; Lindeberg, M.; Shivas, R.G.; Manners, J.M. Genome Sequences of *Pseudomonas* spp. Isolated from Cereal Crops. *Genome Announc.* **2013**, *1*, e00209–e00213.
47. 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.
48. Liu, H.; Qiu, H.; Zhao, W.; Cui, Z.; Ibrahim, M.; Jin, G.; Li, B.; Zhu, B.; Xie, G.L. Genome sequence of the plant pathogen *Pseudomonas syringae* pv. *panici* LMG 2367. *J. Bacteriol.* **2012**, *194*, 5693–5694.
49. Sarris, P.F.; Trantas, E.A.; Baltrus, D.A.; Bull, C.T.; Wechter, W.P.; Yan, S.; Ververidis, F.; Almeida, N.F.; Jones, C.D.; Dangl, J.L.; *et al.* Comparative genomics of multiple strains of *Pseudomonas cannabina* pv. *alisalensis*, a potential model pathogen of both monocots and dicots. *PLoS One* **2013**, *8*, e59366.

50. Baltrus, D.A.; Dougherty, K.; Beckstrom-Sternberg, S.M.; Beckstrom-Sternberg, J.S.; Foster, J.T. Incongruence between multi-locus sequence analysis (MLSA) and whole-genome-based phylogenies: *Pseudomonas syringae* pathovar *pisi* as a cautionary tale. *Mol. Plant Pathol.* **2014**, doi: 10.1111/mpp.12103.
51. Elasm, M.; Delorme, S.; Lemanceau, P.; Stewart, G.; Laue, B.; Glickmann, E.; Oger, P.M.; Dessaux, Y. Acyl-homoserine lactone production is more common among plant-associated *Pseudomonas* spp. than among soilborne *Pseudomonas* spp. *Appl. Environ. Microbiol.* **2001**, *67*, 1198–1209.
52. Marcelletti, S.; Ferrante, P.; Petriccione, M.; Firrao, G.; Scortichini, M. *Pseudomonas syringae* pv. *actinidiae* draft genomes comparison reveal strain-specific features involved in adaptation and virulence to *Actinidia* species. *PLoS One* **2011**, *6*, e27297.
53. Deng, W.-L.; Preston, G.; Collmer, A.; Chang, C.-J.; Huang, H.-C. Characterization of the *hrpC* and *hrpRS* operons of *Pseudomonas syringae* pathovars *syringae*, *tomato*, and *glycinea* and analysis of the ability of *hrpF*, *hrpG*, *hrcC*, *hrpT*, and *hrpV* mutants to elicit the hypersensitive response and disease in plants. *J. Bacteriol.* **1998**, *180*, 4523–4531.
54. Sarkar, S.F.; Gordon, J.S.; Martin, G.B.; Guttman, D.S. Comparative genomics of host-specific virulence in *Pseudomonas syringae*. *Genetics* **2006**, *174*, 1041–1056.
55. Cintas, N.A.; Koike, S.T.; Bull, C.T. A New pathovar, *Pseudomonas syringae* pv. *alisalensis* pv. nov., proposed for the causal agent of bacterial blight of broccoli and broccoli raab. *Plant Dis.* **2002**, *86*, 992–998.
56. Ishiyama, Y.; Yamagishi, N.; Ogiso, H.; Fujinaga, M.; Takikawa, Y. Bacterial brown spot on *Avena storigosa* Schreb. caused by *Pseudomonas syringae* pv. *alisalensis*. *J. Gen. Plant Pathol.* **2013**, *79*, 155–157.
57. Gazi, A.D.; Sarris, P.F.; Fadoulglou, V.E.; Charova, S.N.; Mathioudakis, N.; Panopoulos, N.J.; Kokkinidis, M. Phylogenetic analysis of a gene cluster encoding an additional, rhizobial-like type III secretion system that is narrowly distributed among *Pseudomonas syringae* strains. *BMC Microbiol.* **2012**, *12*, 188.
58. Wang, Y.; Huang, H.; Sun, M.; Zhang, Q.; Guo, D. T3DB: An integrated database for bacterial type III secretion system. *BMC Bioinform.* **2012**, *13*, 66.
59. Cai, R.; Lewis, J.; Yan, S.; Liu, H.; Clarke, C.R.; Campanile, F.; Almeida, N.F.; Studholme, D.J.; Lindeberg, M.; Schneider, D.; et al. The plant pathogen *Pseudomonas syringae* pv. *tomato* is genetically monomorphic and under strong selection to evade tomato immunity. *PLoS Pathog.* **2011**, *7*, e1002130.
60. Murillo, J.; Bardaji, L.; Navarro de la Fuente, L.; Führer, M.E.; Aguilera, S.; Álvarez-Morales, A. Variation in conservation of the cluster for biosynthesis of the phytotoxin phaseolotoxin in *Pseudomonas syringae* suggests at least two events of horizontal acquisition. *Res. Microbiol.* **2011**, *162*, 253–261.
61. Arrebola, E.; Cazorla, F.M.; Codina, J.C.; Gutiérrez-Barranquero, J.A.; Pérez-García, A.; de Vicente, A. Contribution of mangotoxin to the virulence and epiphytic fitness of *Pseudomonas syringae* pv. *syringae*. *Int. Microbiol.* **2009**, *12*, 87–95.

62. Groll, M.; Schellenberg, B.; Bachmann, A.S.; Archer, C.R.; Huber, R.; Powell, T.K.; Lindow, S.; Kaiser, M.; Dudler, R. A plant pathogen virulence factor inhibits the eukaryotic proteasome by a novel mechanism. *Nature* **2008**, *452*, 755–758.
63. Carrión, V.J.; Gutiérrez-Barranquero, J.A.; Arrebola, E.; Bardaji, L.; Codina, J.C.; de Vicente, A.; Cazorla, F.M.; Murillo, J. The mangotoxin biosynthetic operon (mbo) is specifically distributed within *Pseudomonas syringae* genomospecies 1 and was acquired only once during evolution. *Appl. Environ. Microbiol.* **2013**, *79*, 756–767.
64. Costacurta, A.; Vanderleyden, J. Synthesis of Phytohormones by Plant-Associated Bacteria. *Crit. Rev. Microbiol.* **2008**, *21*, 1–18.
65. Akiyoshi, D.E.; Regier, D.A.; Gordon, M.P. Cytokinin production by *Agrobacterium* and *Pseudomonas* spp. *J. Bacteriol.* **1987**, *169*, 4242–4248.
66. Weingart, H.; Völksch, B.; Ullrich, M.S. Comparison of Ethylene Production by *Pseudomonas syringae* and *Ralstonia solanacearum*. *Phytopathology* **1999**, *89*, 360–365.
67. Glickmann, E.; Gardan, L.; Jacquet, S.; Hussain, S.; Elasri, M.; Petit, A.; Dessaux, Y. Auxin production is a common feature of most pathovars of *Pseudomonas syringae*. *Mol. Plant. Microbe. Interact.* **1998**, *11*, 156–162.
68. Roberto, F.F.; Klee, H.; White, F.; Nordeen, R.; Kosuge, T. Expression and fine structure of the gene encoding N epsilon-(indole-3-acetyl)-L-lysine synthetase from *Pseudomonas savastanoi*. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 5797–5801.
69. Miller, M.B.; Bassler, B.L. Quorum sensing in bacteria. *Annu. Rev. Microbiol.* **2001**, *55*, 165–199.
70. Nealson, K.H. Autoinduction of bacterial luciferase. *Arch. Microbiol.* **1977**, *112*, 73–79.
71. Von Bodman, S.B.; Bauer, W.D.; Coplin, D.L. Quorum sensing in plant-pathogenic bacteria. *Annu. Rev. Phytopathol.* **2003**, *41*, 455–482.
72. Dumenyo, C.K.; Mukherjee, A.; Chun, W.; Chatterjee, A.K. Genetic and physiological evidence for the production of N-acyl homoserine lactones by *Pseudomonas syringae* pv. *syringae* and other fluorescent plant pathogenic *Pseudomonas* species. *Eur. J. Plant Pathol.* **1998**, *104*, 569–582.
73. BCCM/LMG bacteria catalogue-Strain details LMG2367. Available online: http://bccm.belspo.be/db/lmg_strain_details.php?NUM=2367 (accessed on 7 November 2013).
74. Venturi, V.; Fuqua, C. Chemical signaling between plants and plant-pathogenic bacteria. *Annu. Rev. Phytopathol.* **2013**, *51*, 17–37.
75. Schikora, A.; Schenk, S.T.; Stein, E.; Molitor, A.; Zuccaro, A.; Kogel, K.-H. N-acyl-homoserine lactone confers resistance toward biotrophic and hemibiotrophic pathogens via altered activation of AtMPK6. *Plant Physiol.* **2011**, *157*, 1407–1418.
76. Brencic, A.; Winans, S.C. Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 155–194.
77. Subramoni, S.; Venturi, V. LuxR-family “solos”: Bachelor sensors/regulators of signalling molecules. *Microbiology* **2009**, *155*, 1377–1385.
78. Subramoni, S.; Gonzalez, J.F.; Johnson, A.; Péchy-Tarr, M.; Rochat, L.; Paulsen, I.; Loper, J.E.; Keel, C.; Venturi, V. Bacterial subfamily of LuxR regulators that respond to plant compounds. *Appl. Environ. Microbiol.* **2011**, *77*, 4579–4588.
79. Ferluga, S.; Venturi, V. OryR is a LuxR-family protein involved in interkingdom signaling between pathogenic *Xanthomonas oryzae* pv. *oryzae* and rice. *J. Bacteriol.* **2009**, *191*, 890–897.

80. González, J.F.; Myers, M.P.; Venturi, V. The inter-kingdom solo OryR regulator of *Xanthomonas oryzae* is important for motility. *Mol. Plant Pathol.* **2013**, *14*, 211–221.
81. Lu, S.-E.; Scholz-Schroeder, B.K.; Gross, D.C. Characterization of the *salA*, *syrF*, and *syrG* regulatory genes located at the right border of the syringomycin gene cluster of *Pseudomonas syringae* pv. *syringae*. *Mol. Plant-Microbe Interact.* **2002**, *15*, 43–53.
82. Mo, Y.Y.; Gross, D.C. Plant signal molecules activate the *syrB* gene, which is required for syringomycin production by *Pseudomonas syringae* pv. *syringae*. *J. Bacteriol.* **1991**, *173*, 5784–5792.
83. Deisinger, P.J.; Hill, T.S.; English, J.C. Human exposure to naturally occurring hydroquinone. *J. Toxicol. Environ. Health* **1996**, *47*, 31–46.
84. Zakataeva, N.P.; Kutukova, E.A.; Gronskiy, S.V.; Troshin, P.V.; Livshits, V.A.; Aleshin, V.V. Export of metabolites by the proteins of the DMT and RhtB families and its possible role in intercellular communication. *Microbiology* **2006**, *75*, 438–448.
85. Degrassi, G.; Devescovi, G.; Solis, R.; Steindler, L.; Venturi, V. *Oryza sativa* rice plants contain molecules that activate different quorum-sensing N-acyl homoserine lactone biosensors and are sensitive to the specific AiiA lactonase. *FEMS Microbiol. Lett.* **2007**, *269*, 213–220.
86. Gao, M.; Teplitski, M.; Robinson, J.B.; Bauer, W.D. Production of Substances by *Medicago truncatula* that Affect Bacterial Quorum Sensing. *Mol. Plant-Microbe Interact.* **2003**, *16*, 827–834.
87. Osman, S.F.; Fett, W.F.; Fishman, M.L. Exopolysaccharides of the phytopathogen *Pseudomonas syringae* pv. *glycinea*. *J. Bacteriol.* **1986**, *166*, 66–71.
88. Franklin, M.J.; Nivens, D.E.; Weadge, J.T.; Howell, P.L. Biosynthesis of the *Pseudomonas aeruginosa* Extracellular Polysaccharides, Alginate, Pel, and Psl. *Front. Microbiol.* **2011**, *2*, 167.
89. Li, H.; Ullrich, M.S. Characterization and mutational analysis of three allelic *lsc* genes encoding levansucrase in *Pseudomonas syringae*. *J. Bacteriol.* **2001**, *183*, 3282–3292.
90. Penalzoza-Vazquez, A.; Kidambi, S.; Chakrabarty, A.; Bender, C. Characterization of the alginate biosynthetic gene cluster in *Pseudomonas syringae* pv. *syringae*. *J. Bacteriol.* **1997**, *179*, 4464–4472.
91. Laue, H.; Schenk, A.; Li, H.; Lambertsen, L.; Neu, T.R.; Molin, S.; Ullrich, M.S. Contribution of alginate and levan production to biofilm formation by *Pseudomonas syringae*. *Microbiology* **2006**, *152*, 2909–2918.
92. Byrd, M.S.; Sadovszkaya, I.; Vinogradov, E.; Lu, H.; Sprinkle, A.B.; Richardson, S.H.; Ma, L.; Ralston, B.; Parsek, M.R.; Anderson, E.M.; *et al.* Genetic and biochemical analyses of the *Pseudomonas aeruginosa* Psl exopolysaccharide reveal overlapping roles for polysaccharide synthesis enzymes in Psl and LPS production. *Mol. Microbiol.* **2009**, *73*, 622–638.
93. Jackson, K.D.; Starkey, M.; Kremer, S.; Parsek, M.R.; Wozniak, D.J. Identification of *psl*, a locus encoding a potential exopolysaccharide that is essential for *Pseudomonas aeruginosa* PAO1 biofilm formation. *J. Bacteriol.* **2004**, *186*, 4466–4475.
94. Hockett, K.L.; Burch, A.Y.; Lindow, S.E. Thermo-regulation of genes mediating motility and plant interactions in *Pseudomonas syringae*. *PLoS One* **2013**, *8*, e59850.
95. Shrivastava, S.; Mande, S.S. Identification and functional characterization of gene components of Type VI Secretion system in bacterial genomes. *PLoS One* **2008**, *3*, e2955.
96. Kapitein, N.; Mogk, A. Deadly syringes: type VI secretion system activities in pathogenicity and interbacterial competition. *Curr. Opin. Microbiol.* **2013**, *16*, 52–58.

97. Barret, M.; Egan, F.; Fargier, E.; Morrissey, J.P.; O’Gara, F. Genomic analysis of the type VI secretion systems in *Pseudomonas* spp.: Novel clusters and putative effectors uncovered. *Microbiology* **2011**, *157*, 1726–1739.
98. Hachani, A.; Lossi, N.S.; Hamilton, A.; Jones, C.; Bleves, S.; Albesa-Jové, D.; Filloux, A. Type VI secretion system in *Pseudomonas aeruginosa*: Secretion and multimerization of VgrG proteins. *J. Biol. Chem.* **2011**, *286*, 12317–12327.
99. Burtnick, M.N.; Brett, P.J.; Harding, S.V.; Ngugi, S.A.; Ribot, W.J.; Chantratita, N.; Scorpio, A.; Milne, T.S.; Dean, R.E.; Fritz, D.L.; *et al.* The cluster 1 type VI secretion system is a major virulence determinant in *Burkholderia pseudomallei*. *Infect. Immun.* **2011**, *79*, 1512–1525.
100. Sarris, P.F.; Skandalis, N.; Kokkinidis, M.; Panopoulos, N.J. *In silico* analysis reveals multiple putative type VI secretion systems and effector proteins in *Pseudomonas syringae* pathovars. *Mol. Plant Pathol.* **2010**, *11*, 795–804.
101. Sarris, P.F.; Trantas, E.A.; Skandalis, N.; Tampakaki, A.P.; Kapanidou, M.; Kokkinidis, M.; Panopoulos, N.J. Phytobacterial type VI secretion system—Gene distribution, phylogeny, structure and biological functions. In *InTech Plant Pathology*; Cumagun, C.J.R., Ed.; InTech: Rijeca, Croatia, 2012; pp. 53–84.
102. Pukatzki, S.; Ma, A.T.; Revel, A.T.; Sturtevant, D.; Mekalanos, J.J. Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 15508–15513.
103. Russell, A.B.; LeRoux, M.; Hathazi, K.; Agnello, D.M.; Ishikawa, T.; Wiggins, P.A.; Wai, S.N.; Mougous, J.D. Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors. *Nature* **2013**, *496*, 508–512.
104. Hood, R.D.; Singh, P.; Hsu, F.; Güvener, T.; Carl, M.A.; Trinidad, R.R.S.; Silverman, J.M.; Ohlson, B.B.; Hicks, K.G.; Plemel, R.L.; *et al.* A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. *Cell Host Microbe* **2010**, *7*, 25–37.
105. 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.
106. Greenberg, J.T.; Vinatzer, B.A. Identifying type III effectors of plant pathogens and analyzing their interaction with plant cells. *Curr. Opin. Microbiol.* **2003**, *6*, 20–28.
107. Swingle, B.; Bao, Z.; Markel, E.; Chambers, A.; Cartinhour, S. Recombineering using RecTE from *Pseudomonas syringae*. *Appl. Environ. Microbiol.* **2010**, *76*, 4960–4968.
108. Juhas, M.; van der Meer, J.R.; Gaillard, M.; Harding, R.M.; Hood, D.W.; Crook, D.W. Genomic islands: Tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol. Rev.* **2009**, *33*, 376–393.
109. Gaillard, M.; Vallaeys, T.; Vorhölter, F.J.; Minoia, M.; Werlen, C.; Sentchilo, V.; Pühler, A.; van der Meer, J.R. The *clc* element of *Pseudomonas* sp. strain B13, a genomic island with various catabolic properties. *J. Bacteriol.* **2006**, *188*, 1999–2013.
110. Wozniak, R.A.F.; Waldor, M.K. Integrative and conjugative elements: Mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat. Rev. Microbiol.* **2010**, *8*, 552–563.

111. Rodríguez-Blanco, A.; Lemos, M.L.; Osorio, C.R. Integrating conjugative elements as vectors of antibiotic, mercury, and quaternary ammonium compound resistance in marine aquaculture environments. *Antimicrob. Agents Chemother.* **2012**, *56*, 2619–2626.
112. Pitman, A.R.; Jackson, R.W.; Mansfield, J.W.; Kaitell, V.; Thwaites, R.; Arnold, D.L. Exposure to host resistance mechanisms drives evolution of bacterial virulence in plants. *Curr. Biol.* **2005**, *15*, 2230–2235.
113. Zhao, Y.; Ma, Z.; Sundin, G.W. Comparative genomic analysis of the pPT23A plasmid family of *Pseudomonas syringae*. *J. Bacteriol.* **2005**, *187*, 2113–2126.
114. Bender, C.; Palmer, D.; Peñalosa-Vázquez, A.; Rangaswamy, V.; Ullrich, M. Biosynthesis of coronatine, a thermoregulated phytotoxin produced by the phytopathogen *Pseudomonas syringae*. *Arch. Microbiol.* **1996**, *166*, 71–75.
115. Pérez-Martínez, I.; Zhao, Y.; Murillo, J.; Sundin, G.W.; Ramos, C. Global genomic analysis of *Pseudomonas savastanoi* pv. *savastanoi* plasmids. *J. Bacteriol.* **2008**, *190*, 625–635.
116. Juhas, M.; Crook, D.W.; Hood, D.W. Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence. *Cell. Microbiol.* **2008**, *10*, 2377–2386.
117. Makarova, K.S.; Wolf, Y.I.; Koonin, E.V. Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Res.* **2013**, *41*, 4360–4377.
118. Toussaint, A.; Prangishvili, D.; Molineux, I.; Fineran, P.C.; Charpentier, E. Memory of viral infections by CRISPR-Cas adaptive immune systems: Acquisition of new information. *Virology* **2012**, *434*, 202–209.
119. Haft, D.H.; Selengut, J.; Mongodin, E.F.; Nelson, K.E. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* **2005**, *1*, e60.
120. Makarova, K.S.; Wolf, Y.I.; van der Oost, J.; Koonin, E. V Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biol. Direct* **2009**, *4*, 29.
121. Cotter, P.D.; Ross, R.P.; Hill, C. Bacteriocins—A viable alternative to antibiotics? *Nat. Rev. Microbiol.* **2013**, *11*, 95–105.
122. Hassan, M.; Kjos, M.; Nes, I.F.; Diep, D.B.; Lotfipour, F. Natural antimicrobial peptides from bacteria: Characteristics and potential applications to fight against antibiotic resistance. *J. Appl. Microbiol.* **2012**, *113*, 723–736.
123. Grinter, R.; Roszak, A.W.; Cogdell, R.J.; Milner, J.J.; Walker, D. The crystal structure of the lipid II-degrading bacteriocin syringacin M suggests unexpected evolutionary relationships between colicin M-like bacteriocins. *J. Biol. Chem.* **2012**, *287*, 38876–38888.
124. Nikolouli, K.; Mossialos, D. Bioactive compounds synthesized by non-ribosomal peptide synthetases and type-I polyketide synthases discovered through genome-mining and metagenomics. *Biotechnol. Lett.* **2012**, *34*, 1393–1403.
125. Ramel, C.; Tobler, M.; Meyer, M.; Bigler, L.; Ebert, M.-O.; Schellenberg, B.; Dudler, R. Biosynthesis of the proteasome inhibitor syringolin A: The ureido group joining two amino acids originates from bicarbonate. *BMC Biochem.* **2009**, *10*, 26.
126. Arrebola, E.; Carrión, V.J.; Cazorla, F.M.; Pérez-García, A.; Murillo, J.; de Vicente, A. Characterisation of the mgo operon in *Pseudomonas syringae* pv. *syringae* UMAF0158 that is required for mangotoxin production. *BMC Microbiol.* **2012**, *12*, 10.

127. Yu, X.; Lund, S.P.; Scott, R.A.; Greenwald, J.W.; Records, A.H.; Nettleton, D.; Lindow, S.E.; Gross, D.C.; Beattie, G.A. Transcriptional responses of *Pseudomonas syringae* to growth in epiphytic versus apoplastic leaf sites. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E425–E434.
128. NCBI FTP Server. Available online: <ftp://ftp.ncbi.nlm.nih.gov/genomes/> (accessed on 4 September 2013).
129. Aziz, R.K.; Bartels, D.; Best, A.A.; DeJongh, M.; Disz, T.; Edwards, R.A.; Formsma, K.; Gerdes, S.; Glass, E.M.; Kubal, M.; *et al.* The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **2008**, *9*, 75.
130. Seemann, T. Prokka: Prokaryotic Genome Annotation System. Available online: <http://vicbioinformatics.com/> (Accessed on 10 September 2013).
131. Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **2011**, *28*, 2731–2739.
132. Zhao, Y.; Wu, J.; Yang, J.; Sun, S.; Xiao, J.; Yu, J. PGAP: Pan-genomes analysis pipeline. *Bioinformatics* **2012**, *28*, 416–418.
133. Carver, T.; Harris, S.R.; Berriman, M.; Parkhill, J.; McQuillan, J.A. Artemis: An integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* **2012**, *28*, 464–469.
134. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410.
135. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinform.* **2009**, *10*, 421.
136. *Pseudomonas syringae* Genome Resources Home Page. Available online: <http://www.pseudomonas-syringae.org/> (accessed on 20 January 2014).
137. Mucyn, T.S.; Yourstone, S.; Lind, A.L.; Biswas, S.; Nishimura, M.T.; Baltrus, D.A.; Cumbie, J.S.; Chang, J.H.; Jones, C.D.; Dangl, J.L.; *et al.* Variable suites of non-effector genes are co-regulated in the type III secretion virulence regulon across the *Pseudomonas syringae* phylogeny. *PLoS Pathog.* **2014**, *10*, e1003807.
138. Markowitz, V.M.; Mavromatis, K.; Ivanova, N.N.; Chen, I.-M.A.; Chu, K.; Kyrpides, N.C. IMG ER: A system for microbial genome annotation expert review and curation. *Bioinformatics* **2009**, *25*, 2271–2278.
139. Marchler-Bauer, A.; Lu, S.; Anderson, J.B.; Chitsaz, F.; Derbyshire, M.K.; DeWeese-Scott, C.; Fong, J.H.; Geer, L.Y.; Geer, R.C.; Gonzales, N.R.; *et al.* CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* **2011**, *39*, D225–D229.
140. De Jong, A.; van Heel, A.J.; Kok, J.; Kuipers, O.P. BAGEL2: Mining for bacteriocins in genomic data. *Nucleic Acids Res.* **2010**, *38*, W647–W651.
141. Grissa, I.; Vergnaud, G.; Pourcel, C. CRISPRFinder: A web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* **2007**, *35*, W52–W57.