



Comparative Genomics of *Ralstonia solanacearum* Identifies Candidate Genes Associated with Cool Virulence

Ana M. Bocsanczy¹, Jose C. Huguet-Tapia² and David J. Norman^{1*}

¹ Mid-Florida Research and Education Center, Department of Plant Pathology, University of Florida, Apopka, FL, United States, ² Department of Plant Pathology, University of Florida, Gainesville, FL, United States

OPEN ACCESS

Edited by:

Phillippe Prior,
Institut National de la Recherche
Agronomique, France

Reviewed by:

Fanhong Meng,
Texas A&M University, United States
David John Studholme,
University of Exeter, United Kingdom

*Correspondence:

David J. Norman
djn@ufl.edu

Specialty section:

This article was submitted to
Plant Microbe Interactions,
a section of the journal
Frontiers in Plant Science

Received: 29 March 2017

Accepted: 28 August 2017

Published: 13 September 2017

Citation:

Bocsanczy AM, Huguet-Tapia JC
and Norman DJ (2017) Comparative
Genomics of *Ralstonia solanacearum*
Identifies Candidate Genes
Associated with Cool Virulence.
Front. Plant Sci. 8:1565.
doi: 10.3389/fpls.2017.01565

Strains of the *Ralstonia solanacearum* species complex in the phylotype IIB group are capable of causing Bacterial Wilt disease in potato and tomato at temperatures lower than 24°C. The capability of these strains to survive and to incite infection at temperatures colder than their normally tropical boundaries represents a threat to United States agriculture in temperate regions. In this work, we used a comparative genomics approach to identify orthologous genes linked to the lower temperature virulence phenotype. Six *R. solanacearum* cool virulent (CV) strains were compared to six strains non-pathogenic at low temperature (NPLT). CV strains can cause Bacterial Wilt symptoms at temperatures below 24°C, while NPLT cannot. Four *R. solanacearum* strains were sequenced for this work in order to complete the comparison. An orthologous genes comparison identified 44 genes present only in CV strains and 19 genes present only in NPLT strains. Gene annotation revealed a high percentage of genes compared with whole genomes in the transcriptional regulator and transport categories. A single nucleotide polymorphism (SNP) analysis identified 265 genes containing conserved non-synonymous SNPs in CV strains. Ten genes in the pathogenicity category were identified in this group. Comparisons of type 3 secretion system, type 6 secretion system (T6SS) clusters, and associated effectors did not indicate a correlation with the CV phenotype except for one T6SS VGR effector potentially associated with the CV phenotype. This is the first *R. solanacearum* genomic comparative analysis of multiple strains with different temperature related virulence. The candidate genes identified by this comparison are potential factors involved in virulence at low temperatures that need to be investigated. The high percentage of transcriptional regulators among the genes present only in CV strains supports the hypothesis that temperature dependent regulation of virulence genes explains the differential virulence phenotype at low temperatures. This comparison contributes to find new possible connections of temperature dependent virulence to the previously described complex regulatory system involving quorum-sensing, phenotype conversion (phcA), acyl-HSL production and responses to SA. It also added novel candidate T6SS effectors and useful detailed information about the T6SS in *R. solanacearum*.

Keywords: transcriptional regulators, temperature, quorum-sensing, T3SS, T6SS, effectors

INTRODUCTION

Ralstonia solanacearum has been listed among the top 10 plant pathogenic bacteria because of its broad host range, genetic diversity, geographical spread, and the destructive symptoms it produces (Mansfield et al., 2012). This bacterial species can infect more than 200 plant species, causing different diseases including Brown Rot in potatoes, Bacterial Wilt of solanaceous plants and ornamentals, and Moko Disease of musaceae family (Hayward, 1964; Elphinstone, 2005). The bacterial species as a whole is heterogeneous and has been difficult to classify. Currently the most used classification divides the species into four main phylotypes, correlated with geographical locations in tropical, subtropical, and temperate regions of the world (Fegan and Prior, 2005; Remenant et al., 2010). *R. solanacearum* can survive in wet soil for years, from there it can infect plant roots, penetrating through wounds or natural openings (Buddenhagen and Kelman, 1964). Once in the plant, it moves to the xylem, where it multiplies to large population and produces the wilting symptoms. The disease is devastating. Its overall economic impact is difficult to quantify due to the geographical span of the disease and the degree and variety of symptoms it causes on the different crops. In the United States, the most significant economic impact is on potato, tomato, and tobacco production in southern states, caused by endemic strains. A particular subgroup of phylotype II populations, known as R3B2 or phylotype IIB-sequevar 1, can cause disease in potato and to a lesser degree in tomato at temperatures below 24°C (Ciampi and Sequeira, 1980; Milling et al., 2009). This monophyletic group of strains is not present in the United States, although it has been inadvertently introduced a number of times through imported geraniums, without establishment. Because of the threat this group represents for United States agriculture, it has been designated select agent under the United States Agricultural Bioterrorism Protection Act of 2002 (Lambert, 2002). Although R3B2 strains are not present in the United States, CV strains that belong to phylotype IIB but not in the R3B2 group and are pathogenic in tomato and potato plants at low temperatures are present in the United States and were isolated from pothos and other ornamentals (Norman and Yuen, 1997, 1999; Bocsanczy et al., 2012). P673, a representative of this group that was previously identified as Race 1 biovar 1, belongs to phylotype II-sequevar 4NPB (Bocsanczy et al., 2013) and has been recently identified in the Caribbean as an emergent group closely related to the Moko disease strains (Deberdt et al., 2014). The danger of introduction of R3B2 strains and the presence in the United States of other groups of *R. solanacearum* strains that are capable of causing disease at low temperatures highlights the need to understand the sources and mechanisms of pathogenicity at low temperatures. Advances in DNA sequencing technology, along with lower sequencing costs and computational tools

Abbreviations: CDS, coding sequences; COG, Clusters of Orthologous Genes; CV, cool virulent; EMBL, European Molecular Biology Laboratory; EPS, exopolysaccharides; NCBI, National Center for Biotechnology Information; NPLT, non-pathogenic at low temperatures; PGAP, prokaryotic genome annotation pipeline; R3B2, race 3 biovar 2; SNP, single nucleotide polymorphisms; T3SS, type 3 secretion system; T6SS, type 6 secretion system.

development, have enabled the inference of phenotypic traits and functional genomics analysis through genomic comparisons (Jim et al., 2004; Feldbauer et al., 2015). Protein-coding genes in microbial genome databases are annotated by sequence similarity and conserved domains using Clusters of Orthologous Groups (COGs), protein families (Pfam), and other orthologous conserved domains databases across genomes (Sonnhammer et al., 1998; Tatusov et al., 2000; Punta et al., 2012; Tatusova et al., 2015). Protein functions can be inferred by phylogenetic profiling based on the absence or presence of orthologous genes in microbial genomes and similar profiles are thought to share similar functions (Pellegrini et al., 1999). Thus, the presence or absence of orthologous genes in microbial genomes is very informative, when combined with phenotype profiles, which group microbial genomes by the desired phenotype (Jim et al., 2004; Goh et al., 2006; Tamura and D'Haeseleer, 2008). It allows the prediction of metabolic functions (Kanehisa et al., 2014), protein-protein interactions (Szklarczyk et al., 2015), and particular traits, such as virulence at low temperature in our case. Based on the hypothesis that proteins can lose or alter functions when there are differences in amino acids, a genome-wide search for conserved alleles in particular genome groups associated with a particular trait can identify a group of genes that might be associated with the trait (Read and Massey, 2014). In addition, comparative examination of virulence associated systems such as T3SS and T6SS and associated sets of effectors (effectorome) could identify differences associated with the cool virulence trait. In order to simplify terminology, in this work we will refer to the strains capable of causing disease at low temperatures as CV strains and to the ones that are not capable of causing disease at low temperatures as NPLT.

Few comparative works of *R. solanacearum* strains with different virulence at low temperature have been published. We reported differential proteomics profiles of CV and NPLT strains at 30 and 18°C when in contact with plant roots (Bocsanczy et al., 2014a). In that work, we compared protein expression of two CV with two NPLT *R. solanacearum* strains. We showed that 101 proteins at low temperature were downregulated in NPLT strains while their expression was unchanged or upregulated in CV strains. In that case, the proteins were present in all the strains and the difference was in their expression, leading to the hypothesis that the difference in virulence might be partially explained by the differential regulation of virulence associated genes. Another work examined the difference between the transcriptome of one CV strain and one NPLT either in rich media or *in planta* during infection (Meng et al., 2015). In this work the strains had different transcriptional responses to temperature change. At cool temperature the CV strain upregulated a cluster encoding a quorum sensing-dependent and a hypothetical proteins. These genes were not present in the NPLT strain, and were shown to contribute to virulence of CV at low temperature leading to the hypothesis that genomic differences between CV and NPLT strains also may contribute to differential virulence at low temperature.

In the present work, we hypothesize that candidate genes which contribute to the CV phenotype can be identified by genotypic differences either by the absence/presence of genes

or SNP differences correlated with the phenotype. We present the first genomic comparison of 12 *R. solanacearum* genomes with the aim to identify and characterize *in silico* protein-coding genes associated with the cool virulence phenotype. To date many genomes of *R. solanacearum* have been sequenced and are publicly available either complete or as whole genome shotguns (WGS) in the NCBI database, however, there is limited information about pathogenicity at low temperatures. We selected six CV and six NPLT strains for this comparison. We generated draft genomes sequences of four strains reported in this work to complete this comparison. Overall, only 44 orthologous protein-coding genes are present exclusively in CV strains and only 19 are present exclusively in NPLT strains. Annotation distribution analysis of the 44 orthologous genes revealed a high percentage of transcriptional regulators and transport categories, compared with their abundance in complete genomes, and absence of virulence-associated genes. SNP comparison identified 265 non-synonymous SNPs that correlate with the cool virulence phenotype. Comparative analysis of two secretion systems revealed a high degree of core genes conservation among strains, and a high degree of diversity for their associated effectome.

MATERIALS AND METHODS

R. solanacearum Strains

Twelve strains were selected for this study. **Table 1** lists the strains, their source, the host, their phylogenetic classification, the GenBank sequence accession number, and the genome public references.

Pathogenicity Tests

Pathogenicity tests were carried out as described in previous works (Bocsanczy et al., 2012) with small modifications. Briefly, bacteria were grown for 48 h on Nutrient Agar (NA) plates. Bacterial colonies were diluted in saline solution (NaCl at 8.5 g/liter) up to OD₆₀₀ = 0.01 (1×10^7 CFU/ml) using a spectrophotometer (Nanodrop 2000C, Thermo Scientific). Individual plants were inoculated with 50 ml of the dilution, poured into the pot over the soil, and placed in temperature, light, and humidity controlled chambers. Temperature was set at 18°C. At the end of 30 days, wilted plants were counted, and the percentage of wilted plants was calculated. The experiment was repeated two times with at least 10 plants each treatment.

Genome Sequencing

Ralstonia solanacearum strains were grown at 28°C overnight in CPG (Casamino acid-peptone-glucose) medium to OD₆₀₀ = 1.2. Genomic DNA was obtained using Ultraclean Microbial DNA Isolation Kit cat# 12224-50 (Mobio, Carlsbad, CA, United States) following manufacturer's instructions. Libraries and sequencing were performed using the Illumina HiSeq-2500 with pair end 100 cycles with 150X coverage. Reads were processed in house with Geneious 9.0.2¹ (Kearse et al., 2012). Adapter and low

quality sequences at the reads ends were trimmed. Reads shorter than 50 nt or of low quality were discarded. The remaining reads were assembled with Geneious. The contigs were checked for remaining adaptor sequences and low quality nucleotides. Contigs less than 200 nt were discarded.

Genome Assemblies and Annotation

Contigs were assembled against closely related reference genomes, using Geneious software. Then contigs were sorted and numbered in order with the assembly. Contig annotation was performed with the PGAP of NCBI. Manual editing of annotation was performed with Artemis (Carver et al., 2012).

Genome Statistics

Quality Assemble Tests

Assessment of genome assembly completeness was conducted using Benchmarking Universal Single-Copy Orthologs (BUSCO v3) (Simao et al., 2015). Dataset corresponding to betaproteobacteria was used to evaluate the assembly completeness and gene prediction of the assembled contigs for each genome. In addition SNPs were verified by re-mapping raw reads against the assembled contigs of each genome using Bowtie 2 (Langmead and Salzberg, 2012). Mapping files were transformed into binary files using SAMtools (Li et al., 2009) and the output was used to obtain a corrected set of contigs using Pilon tool v1.16 (Walker et al., 2014).

Significance of Comparative Results

A shuffling experiment was performed in order to test the significance of the comparative genomics comparison. Each one of the 12 strains in the comparison was randomly assigned to an arbitrary class (1 or 2) using a random generator function. Then the strains were searched for genes present or absent in each class and for genome-wide SNP linked to each class. This procedure was repeated 10 times.

Genome Comparison

Orthologous Genes and Pan-Genome

Fasta files and GenBank annotated contigs files were concatenated in one pseudomolecule for visualization and comparative analysis. Both molecules of each published complete genome were also concatenated to facilitate comparison.

Compared strains were classified as "CV" and "NPLT" based on their ability to cause disease at 18°C. Pseudomolecules were aligned with MAUVE progressive alignment tool (Darling et al., 2004) using minimum identity at 60% and minimum coverage at 70% as criteria. Orthologous gene list as calculated by MAUVE was used to produce gene distribution, pan-genome, and identify orthologous genes present only in one of the two phenotypes. Presence and absence of genes were confirmed by protein sequences BLASTP searches (Altschul et al., 1997).

Prokaryotic genome annotation pipeline functional annotation of selected groups of genes were updated and confirmed with annotation with Blast2GO software (Conesa et al., 2005).

¹<http://www.geneious.com>

TABLE 1 | Strains of *Ralstonia solanacearum* used for the genome comparison.

| Strain (other names) | Origin | Isolated | Phylotype/sequence | Cool virulence | Genome reference |
|----------------------|--------------------|-----------|--------------------|----------------|-------------------------|
| P742 (GMI1000) | French Guyana | Tomato | I/18 | NPLT | Salanoubat et al., 2002 |
| P597 | United States | Tomato | IIA/38 | NPLT | This work |
| P660 (K60-1) | United States | Tomato | IIA/7 | NPLT | Remenant et al., 2012 |
| P795 (CIP120) | Peru | Potato | IIA/38 | NPLT | This work |
| P796 (CFBP2957) | French West Indies | Tomato | IIA/36 | NPLT | Remenant et al., 2010 |
| P673 | United States | Pothos | IIB/4-NPB | CV | Bocsanczy et al., 2013 |
| P714 (UW551) | Kenya | Geranium | IIB/1 | CV | Gabriel et al., 2006 |
| P797 (CFBP6783) | Martinique | Heliconia | IIB/4-NPB | CV | CIRAD and this work |
| P800 (UW163) | Peru | Plantain | IIB/4 | CV | CIRAD |
| N807 (23-10BR) | Brazil | Potato | IIB/27 | CV | Clarke et al., 2015 |
| NCPPB909 | Egypt | Potato | IIB/1 | CV | Yuan et al., 2015 |
| P799 (CFBP3059) | Burkina Faso | Eggplant | III/48 | CV | This work |

SNP Detection and Comparison

Single nucleotide polymorphisms detection was carried out with kSNP3 v3 software (Gardner et al., 2015). Pseudomolecules of compared genomes were used as input using as a reference the UW551 genome. KSNP3 v3 finds SNP positions in genomes using MUMmer (Delcher et al., 2002) and produces clusters of SNPs by genome groups. Optimal kmer length (21 bp) for SNP detection was selected using the Kchooser script implemented in the KSNP tool. The output is summarized in a variant call format file (VCF) file that was analyzed and filtered for those SNPs that are conserved only in the CV strains. SNPdat software was used to annotate SNPs (Doran and Creevey, 2013). Additional visual inspection of SNPs and analysis of amino acid conservation in selected cluster of orthologs was conducted using Jalview alignment tool (Waterhouse et al., 2009).

PCR Amplification of Transcriptional Regulators

Conserved primers were designed to amplify the complete CDS of the transcriptional regulators found in the comparison. Transcriptional regulators P673_16370 and P673_16375 are adjacent, and they were amplified together with one pair of primers. Primer sequences are listed in Supplementary Table S1A, and specific PCR conditions such as annealing temperatures and extension time in Supplementary Table S1B. Some of the reactions required DMSO 3% since products are high in GC contents. General PCR conditions were: [denaturation: 98°C 30 s; amplification: 25x (98°C 15 s, annealing 30 s, extension time min); final extension: 72°C 10 min]. 0.8% agarose gels were prepared and run at 90 V for 1 h in order to visualize PCR bands.

Clusters and Effectors Comparisons

Individual clusters of T3SS and T6SS were compared, organized, and manually annotated with aid of MAUVE, Artemis, and Geneious software. T3SS effectors were predicted using the RalstoT3E tool (Peeters et al., 2013). T3SS effectors were identified *in silico* using the software tool Secret6 (Li et al., 2015), and extracting viral genome replication (VGR), proline-alanine-alanine-arginine (PAAR), and rearrangement hot spots (Rhs) domains from gene annotations. All predicted sequences were manually revised and checked for frameshifts or truncations.

Supporting Data

Genome sequences of all compared genomes are in the NCBI database under accession numbers: GMI1000 [NC_003295/NC_003296], P597 [JIBY00000000], K60-1 [CAGT01000001], P795 [JXAY00000000], CFBP2957 [NC_014307], P673 [JALO00000000], UW551 [AAKL00000000], P797 [JXAZ00000000], UW163 [CDMB00000000], 23-10BR [JQOI00000000], NCPPB909 [JNGD00000000], and P799 [JXBA00000000]. K60-1 megaplasmid sequence is available at Microscope Genome annotation and Analysis Platform.²

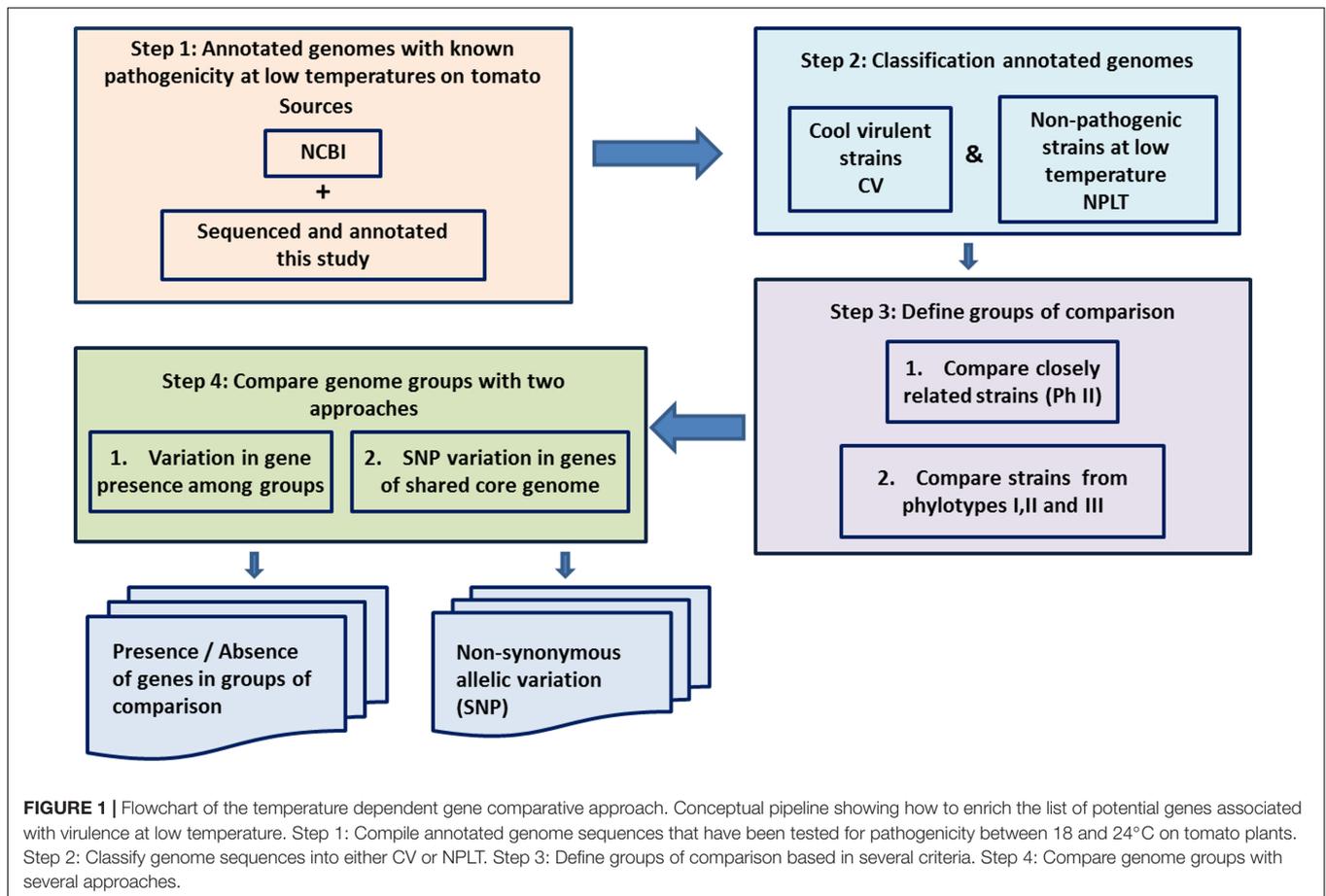
Raw sequence data was deposited in the NCBI Sequence Read Archive (SRA) under accession numbers: P597 [experiment: SRX2850877, RUN: SRR5595754], P799 [experiment: SRX2849979, RUN: SRR5593154], P797 [experiment: SRX2849980, RUN: SRR5593153], P795 [experiment: SRX2849981, RUN: SRR5593152], and P673 [experiments: SRX2850875 (Illumina), SRX2850876 (LS454), RUNS: SRR5595753 (Illumina), SRR5595752 (LS454)].

RESULTS AND DISCUSSION

Selection of *R. solanacearum* Strains for the Comparison

Twelve strains of *R. solanacearum* were selected for the genomic comparison (Table 1). Equal number of CV strains and NPLT strains from the phylotypes I, II and III, with tomato as the common host were compared. The process flowchart is described in Figure 1. The first step was to obtain the annotated genomes for the strains to be compared. Four strains were sequenced to complete the genome database necessary for the comparison. The second step was to confirm their pathogenicity phenotype at low temperature. Since there is little information about cool virulence in the literature, besides the known low temperature pathogenicity of the R3B2 group, candidate strains were tested for low temperature pathogenicity. Strains P673, CFBP6783, UW551, 23-BR10, and UW163 were pathogenic at 18°C, while P597, CIP120, K60-1, CFBP2957, GMI1000, and CFBP3099 were

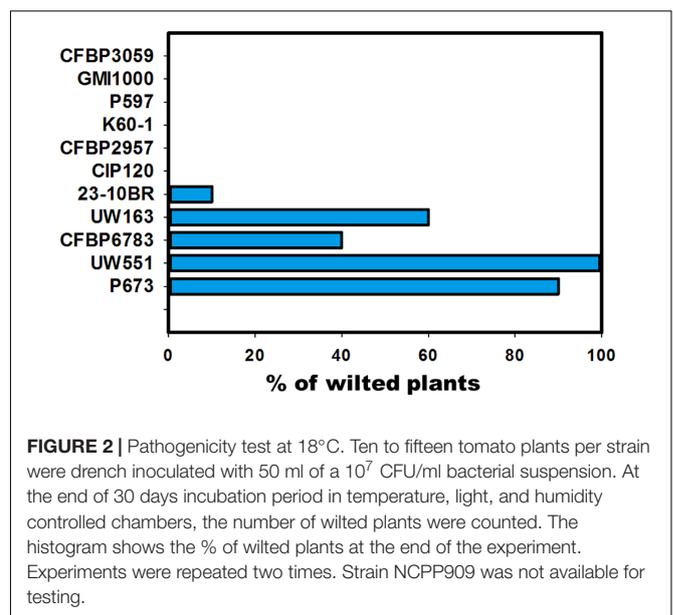
²<https://www.genoscope.cns.fr/agc/microscope/home/index.php>



non-pathogenic (**Figure 2**). We used previous reports of cool virulence for strain NCPPB909 since we did not have the strain available for testing (Clarke et al., 2015; Yuan et al., 2015). Genomes of strains GMI1000, K60-1, CFBP2957, P673, UW551, UW163, 23-BR10, and NCPPB909 were publicly available at the time of this work (Salanoubat et al., 2002; Gabriel et al., 2006; Remenant et al., 2010, 2012; Bocsanczy and Norman, 2013).

General Features of Genome Sequences

The genomes of P597, P795 (CIP120), P797 (CFBP6783), and P799 (CFBP3059) strains were sequenced for this work on Illumina HiSeq platform generating paired end sequences. The reads were assessed, processed, and assembled with Geneious 9.0.2 (Kearse et al., 2012). The genomes were assembled in contigs and aligned to the most similar complete genome publicly available as reference. Contigs were concatenated to form pseudomolecules following the reference order. Assembly statistics are shown in Supplementary Table S2. Reads and assemble quality was measured by several metrics. The Phred score which provides a measurement of the reads quality (Ewing and Green, 1998; Ewing et al., 1998) was approximately 36 for all the strains. A Phred score of 20 or above is considered acceptable, thus 36 corresponds to accurate sequences (Illumina, 2011). The read depth is the number of reads that include a given



nucleotide. It is measured as number of times the length of the genome is covered by the length of reads and gives a sense of the sequencing coverage. Average coverage ranged between 140X and

252X. A coverage of 150X or more is usually considered good coverage with Illumina platform for whole genome sequences (Desai et al., 2013). Another important measure of quality is the N50 length. This metric defines the shortest sequence length of the set of contigs ordered from longest to shortest for which the combined length represents at least 50% of the assembly (Miller et al., 2010). The longer N50, the better the assembly provided high genome coverage is achieved (Desai et al., 2013). P597 and P799 assemblies had a much lower N50 value than P797 and P795. As observed in Supplementary Table S2, the number of contigs mapped for P597 was 530 and for P799 was 399 compared with 145 for P795 and 178 for P797. Assessment of genome assembly completeness indicates that the genomes used in this study contain in average 97% of betaproteobacteria core genome reported in the BUSCO database. The lowest set of completed genes was found in UW551 with 93% of the total set in BUSCO database. Within the group of genomes we assembled in this study, the average increased to 98%. Our results indicate that although some genes are missing we believe that the level of completeness can offer a fair approximation of the gene content of our selected sample (Supplementary Table S3A).

A summary of general genomic features of the four *R. solanacearum* strains sequenced is presented in **Table 2**. The genomes sequenced had similar lengths and were organized in the chromosome and megaplasmid similarly to other *R. solanacearum* strains, confirming that the sequencing covered most of the genomes and did not have long missing sections. Their estimated average size was approximately 5.6 Mb, including a chromosome of approximately 3.5 Mb and a megaplasmid of approximately 2.1 Mb (**Table 2**). Differences between distributions in chromosome and megaplasmid might be due to misplacements of contigs in the draft genomes. The number of proteins predicted ranges between approximately 4500 and 4900. The number of predicted pseudogenes ranges between 141 and 216. Differences in these numbers with other publicly available strains are explained by different bioinformatics methods used to predict open reading frames (ORFs) and differences in quality of the draft genomes.

Genome Comparisons and Genes Distribution

Genome concatenated pseudomolecules were compared using MAUVE software (Darling et al., 2004). In a first step as described in the flowchart (**Figure 1**, Step 3), we included six CV strains all belonging to phylotype IIB and four NPLT strains belonging

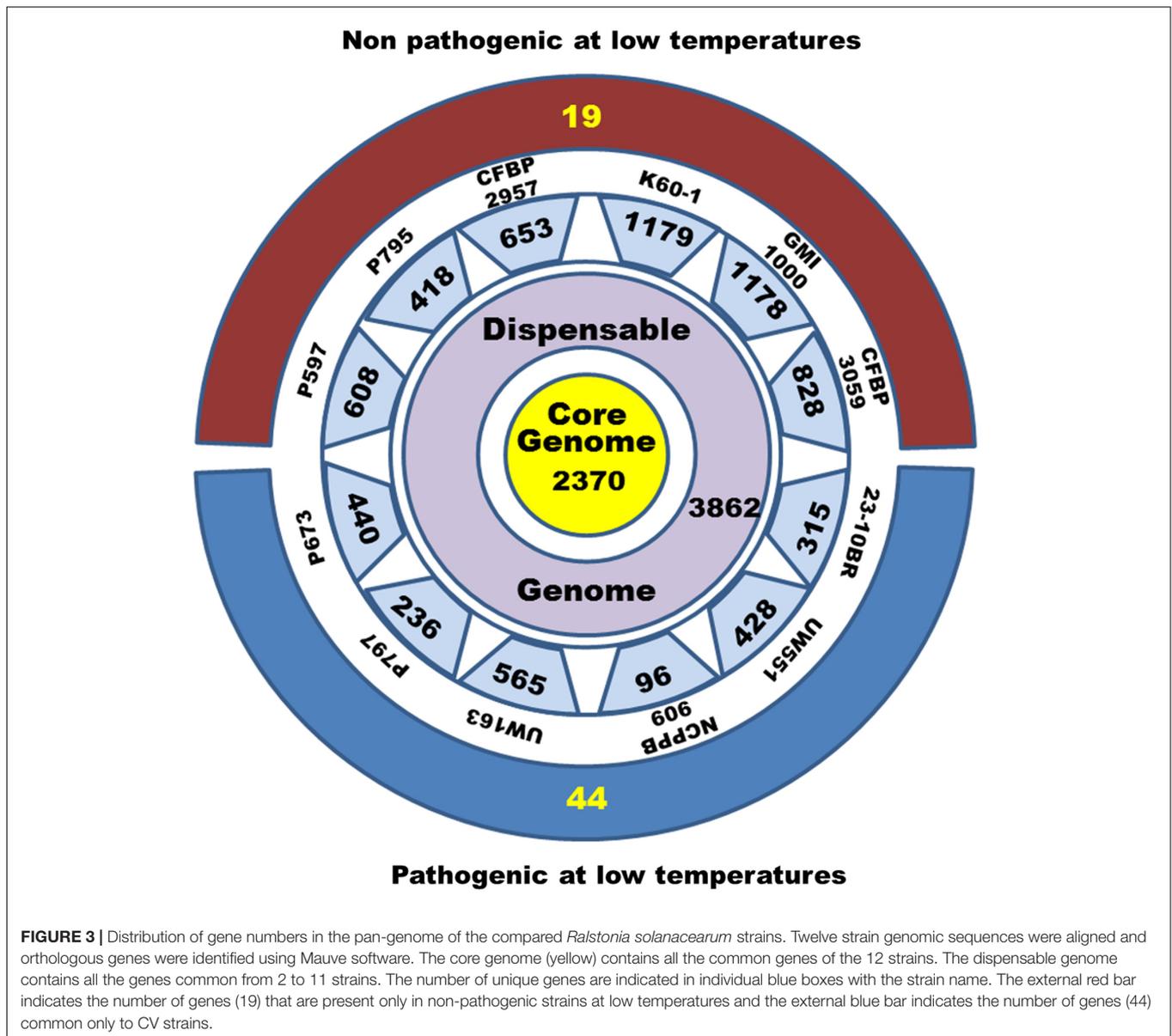
to phylotype IIA (results not shown). All CV strains we have currently identified as CV belong to phylotype IIB sequevars 1, 4, and 27. We included strain UW163, whose main host is Musa (banana). UW163 is also capable of causing disease in tomato at both 30 and 18°C. We also included strain 23-10BR, an atypical R3B2 strain that was classified as sequevar 27 due to its differences with the usually clonal R3B2 strains (Clarke et al., 2015). In order to eliminate orthologous genes that might be present in NPLT strains due to phylogenetic closeness, but not related to the phenotype, for our definitive comparison we added a strain from phylotype I (GMI1000) and a strain from phylotype III (P799-CFBP3059) which are NPLT completing a 12 strain comparison. The phylogenetic relationship of the strains compared in this work are illustrated in Supplementary Figure S1.

Two approaches were used to extract lists of genes correlated with the CV phenotype. (1) Phylogenetic profile grouping genes present or absent in each phenotypic group; and (2) Genome-wide search of non-synonymous SNP associated with each phenotypic group (**Figure 1**, Step 4). The core genome of the 12 strains comparison had 2370 genes (**Figure 3**). The number of genes unique to a strain varied from 96 for NCPPB909 to 1179 for K60-1. The variation in number of unique genes is probably due to differences in annotation systems and the quality of the draft genomes. A better sequence coverage will have more genes that may be missing due to sequencing gaps in other genomes. Annotation is also important. If the CDS prediction and databases searched are different, there might be a decrease in the number of orthologs detected. An example is found with strain K60-1. The sequence was taken from the Microbial Genome Annotation and Analysis Platform (MicroScope) (Vallenet et al., 2013) database because the megaplasmid sequence was missing from the EMBL and NCBI databases. The different prediction methods used for CDS probably caused differences in the number and location of CDS produced. The list of the 44 orthologous genes present only in CV is listed in Supplementary Table S4A, while the 19 present only in NPLT strains is listed in Supplementary Table S4B.

In order to determine if the observed association between genotype and phenotype is different from expected associations by chance to two arbitrary categories, we performed a shuffling experiment. Each strain was assigned randomly to either arbitrary category 1 or 2, and the set of genes present or absent in each category was calculated. We repeated this experiment 10 times. As observed in Supplementary Table S3B only two of the shuffling experiments produced a set of genes. Experiment 5 produced a set of five genes present only in

TABLE 2 | General features of the genomes sequenced for this work.

| Strain | WGS (contigs) | Genome size (Kbp) | Chromosome (Kbp) | Megaplasmid (Kbp) | G+C content (%) | # of CDS | # of RNAs | # of pseudogenes |
|--------|---------------|-------------------|------------------|-------------------|-----------------|----------|-----------|------------------|
| P797 | 178 | 5,557 | 3,568 | 1,989 | 66.64 | 4677 | 57 | 146 |
| P597 | 530 | 5,648 | 3,296 | 2,352 | 66.45 | 4580 | 57 | 206 |
| P795 | 145 | 5,671 | 3,551 | 2,120 | 66.56 | 4776 | 57 | 141 |
| P799 | 399 | 5,396 | 3,391 | 2,005 | 66.90 | 4573 | 58 | 216 |



category 1 and Experiment 8 a set of four genes present only in category 2. The distribution of these genes were completely atypical. We confirmed that the correlation between phenotype and genotype for CV and NPLT strains was not random.

Genes were annotated using Gene Ontology (GO) terminology for biological process categories (Ashburner et al., 2000). The genes distribution by general biological processes is shown in **Figure 4**. In the group of genes present only in CV strains (**Figure 4A**), the highest percentage category was 34% for unknown genes, as expected. Interestingly, the percentages of categories for transport and regulation of transcription were higher than expected in a whole genome distribution, with 18 and 16%, respectively. The transport category is usually in the range of 10–14% and regulation of

transcription is typically in the range of 6% for *R. solanacearum* and other similar bacterial genomes (Madeira and Gabriel, 2007). Particularly, the regulation of transcription category is of interest because its high number supports the hypothesis that differential regulation of virulence genes largely explains the different virulence between CV and NPLT strains (Bocsanczy et al., 2014a). The absence of virulence genes in this list is also indicative that there are no absent virulence genes only in NPLT strains. The distribution of genes present only in NPLT strains was completely atypical (**Figure 4B**). Only three categories were represented where metabolic processes and unknowns were dominating with 63 and 32%, respectively. There were no genes in the categories transcriptional regulators or transport. Almost all genes in this comparison belonged to one cluster mostly containing genes associated with fatty acid and lipid metabolism.

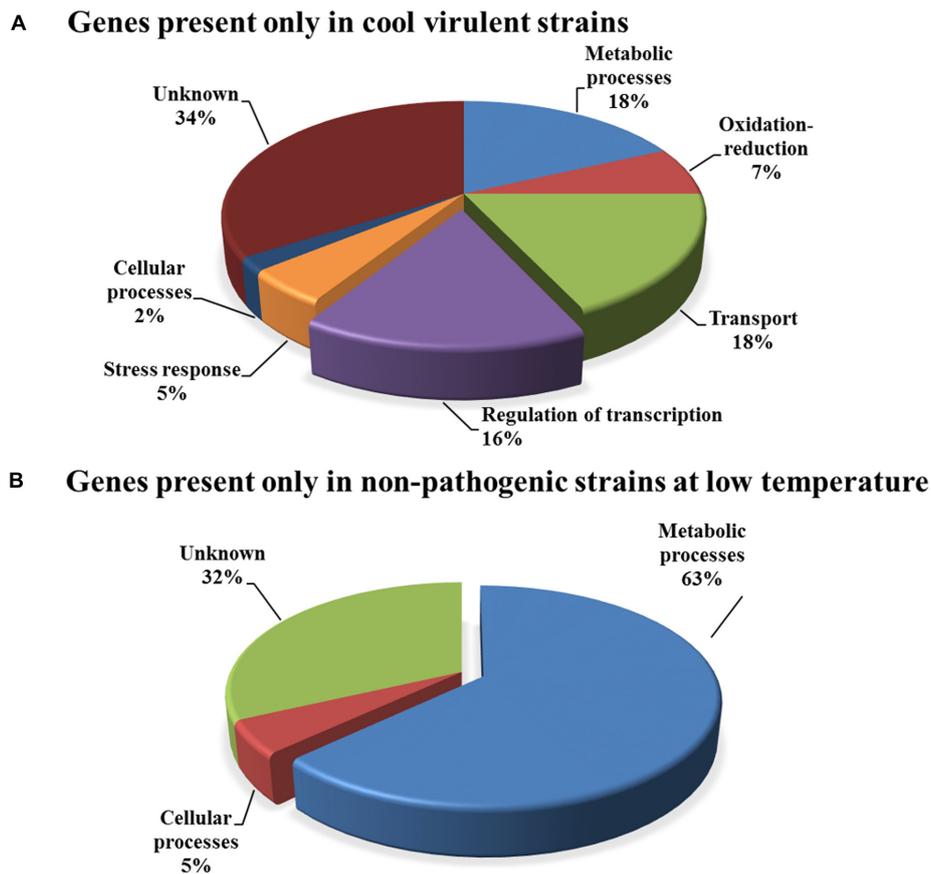


FIGURE 4 | Distribution by biological process categories of orthologous genes associated with the temperature virulence phenotype. Annotated genes were classified according to biological process Gene Ontology (GO) terms. **(A)** Genes present only in CV strains. **(B)** Genes present only in non-pathogenic strains at low temperature.

TABLE 3 | Orthologs of transcriptional regulators present only in CV strains.

| P673 | CFBP6783 | UW551 | NCPBP909 | 23-10BR | UW163 |
|--------------|--------------|------------|------------|------------|---------------|
| RSP673_06495 | RSP797_06395 | RRSL_02465 | CQ06_06630 | KR96_09045 | UW163_RS10960 |
| RSP673_07855 | RSP797_07815 | RRSL_04509 | CQ06_14165 | KR96_20550 | UW163_RS15060 |
| RSP673_09075 | RSP797_08995 | RRSL_00185 | CQ06_01545 | KR96_21300 | UW163_RS13925 |
| RSP673_16370 | RSP797_19820 | RRSL_00747 | CQ06_21525 | KR96_12215 | UW163_RS16390 |
| RSP673_16375 | RSP797_19825 | RRSL_00748 | CQ06_21530 | KR96_12220 | UW163_RS16385 |
| RSP673_21385 | RSP797_22290 | RRSL_00108 | CQ06_07015 | KR96_15245 | UW163_RS19305 |
| RSP673_22445 | RSP797_23380 | RRSL_00831 | CQ06_05885 | KR96_08010 | UW163_RS20305 |

This explains the atypical functional distribution of the gene annotations.

Transcriptional Regulators Present Only in the Compared CV Strains

Seven transcriptional regulators were identified as present only in CV strains. **Table 3** lists the locus tags of the regulators homologs for the six CV strains compared in this work. The list was manually curated in order to verify their presence or absence in CV or NPLT. An *in silico* analysis of these

genes revealed that most were Lys type regulators. The LysR family of transcriptional regulators is the most abundant in bacterial genomes. This family is characterized by a conserved structure with a DNA-binding helix-turn-helix motif in their N-terminus. LysR are conserved global transcriptional regulators that can act as activators or repressors. They have evolved to regulate genes with diverse functions including nitrogen fixation, stress responses, toxin production motility, metabolism, quorum sensing, and virulence (Maddocks and Oyston, 2008). They did not seem to belong to a particular cluster or to occur in close proximity to known virulence factors. Supplementary

Figure S2 illustrates the annotated adjacent genes of these transcriptional regulators. The sequences of genes RSP673_21385 and RSP673_22445 appeared to be truncated; this could be an artifact of the draft genomes assemblies. We designed primers in conserved regions at the N-terminus and C-terminus of each gene to confirm the presence/absence of complete transcriptional regulators in CV/NPLT strains, respectively, and to determine if they were truncated or not. Primers used and recommended annealing temperatures are shown in Supplementary Table S1. PCR amplification of the putative genes carried with the described primers confirmed that the transcriptional regulators are present only in the CV strains compared and are not present in the NPLT strains (Figure 5). Additionally, we confirmed that genes RSP673_21385, RSP673_22445, and their homologs in P797 (CFBP6783) are not truncated as confirmed by the presence of bands of expected length in the PCR products.

In order to strengthen our results we expanded our PCR test to strains of *R. solanacearum* that were not included in our genomic comparison because their genome has not been sequenced. We tested strains P446, P487, P618, and P666 previously shown to be CV (Bocsanczy et al., 2012) and strains P781, P822, P824, and UW757 which we confirmed were non-pathogenic at 18°C (results not shown), As shown in Supplementary Figure S3, the CV strains have all transcriptional regulators while the NPLT do not. Supplementary Figure S3 also show the phylogenetic classification of these strains. These transcriptional regulators are of special interest for further investigation because their presence supports the hypothesis that gene regulation contributes largely to the differences in virulence at low temperatures. This is consistent with the results of previous proteomics works where abundance of virulence and virulence associated proteins was different for CV strains than for NPLT strains at low temperatures (Bocsanczy et al., 2014a).

SNP Comparison

The list of SNPs for all the strains compared using UW551 as reference were produced in a VCF. The SNPs were verified by re-mapping the raw reads available to the contigs with only few differences encountered. kSNP3 was run with different k-mer sizes obtaining the optimum run with a k-mer size of 21 producing a 98.18% of SNPs aligned only one time (non-ambiguous alignment). VCF file was parsed and filtered to select only the SNPs that were conserved in the CV strains. Then the position of the selected SNPs was overlapped to the annotation file and only the SNPs in gene sequences producing changes in the protein sequence (non-synonymous) were considered in the analysis. To further verify the presence of the non-synonymous predicted SNPs, we retrieved and aligned the protein sequence of homologs where the SNPs were detected and confirmed visually the presence of the amino acid change. Non-synonymous SNPs were identified in 256 genes. We assessed the significance of our results performing 10 shuffling experiments (Supplementary Table S3B). We only found two SNPs in Experiment 1 and both were located in genes encoding for metabolic functions. Our test confirm that the SNPs correlation with CV strains is not random.

Gene annotation with non-synonymous SNPs distributed by biological process is presented in Figure 6. The main categories

were unknown with 23% and metabolic processes with 30%. The rest of the category percentages were in the range of typical complete bacterial genomes. The list of genes with non-synonymous conserved SNPs can be found in Supplementary Table S5.

Analysis of previous studies on cool virulence and responses to salicylic acid (SA) indicated significant overlap of the SA regulon with our SNPs results. Thirty and sixteen percent of the genes annotated in the cellular processes and metabolic processes categories, respectively, were also identified previously as downregulated in presence of SA (Lowe-Power et al., 2016). In the latter category, most genes were related to the amino acids metabolic pathways. Thirty percent of the genes annotated in the oxidation-reduction category were also part of the SA regulon. Roughly, half of them were upregulated, including a catalase protein KatE that was previously identified (Bocsanczy et al., 2014a) as temperature dependent, and part of the HrpG regulon (Valls et al., 2006). KatE probably contributes to survival in the presence of an oxidative environment.

All the genes annotated in the motility category belong to the type IV twitching motility system. Twitching motility is associated with attachment and virulence in *R. solanacearum* (Kang et al., 2002). The twitching motility response is temperature dependent *in vitro* and *in planta* during the early stages of infection (Bocsanczy et al., 2012). PilC is a protein part of the structural twitching motility pilus, and PilB is an ATPase thought to power the pilus movement (Burdman et al., 2011), while PilJ has a chemotaxis domain and a homolog in *Pseudomonas aeruginosa* was required for functional pili (Mattick, 2002). SA repressed their gene expression (Lowe-Power et al., 2016). In the stress responses category 20% of the annotated genes were upregulated by SA (Lowe-Power et al., 2016). In addition, a secreted alkaline phosphatase was previously less abundant at 18°C (Bocsanczy et al., 2014a).

In the pathogenicity category, SA repressed 60% of the identified genes (Lowe-Power et al., 2016). We identified a polygalacturonase PehB that is part of the HrpG regulon (Valls et al., 2006). PehB contributes to disease in tomato (Huang and Allen, 2000). Three proteins in the eps operon or related to its function, also regulated by SA, were present in this category. Extracellular polysaccharides (EPS) contribute to virulence and stem colonization (Saile et al., 1997). Tek gene was identified in this category. It encodes a precursor protein which is exported and processed extracellularly to release a 28 kDa from its C-terminus (TeK). This is the single most abundant protein produced by *R. solanacearum* in culture (Denny et al., 1996). This protein abundance was temperature dependent in our proteomics previous study (Bocsanczy et al., 2014a). Although this protein is the most abundantly secreted during pathogenesis, and it is coregulated with EPS, it did not contribute to virulence or affect the levels of EPS during pathogenesis (Denny et al., 1996). A T3 effector AvrPphD also known as RipD was also identified. AvrPphD protein from *Pseudomonas syringae* pv. *phaseolicola* elicits non-host hypersensitive response (HR) on pea (*Pisum sativum*); it is also found in a great variety of *P. syringae* pathovars. As a virulence effector, it is less effective due to redundancy of effector functions (Arnold et al., 2001).

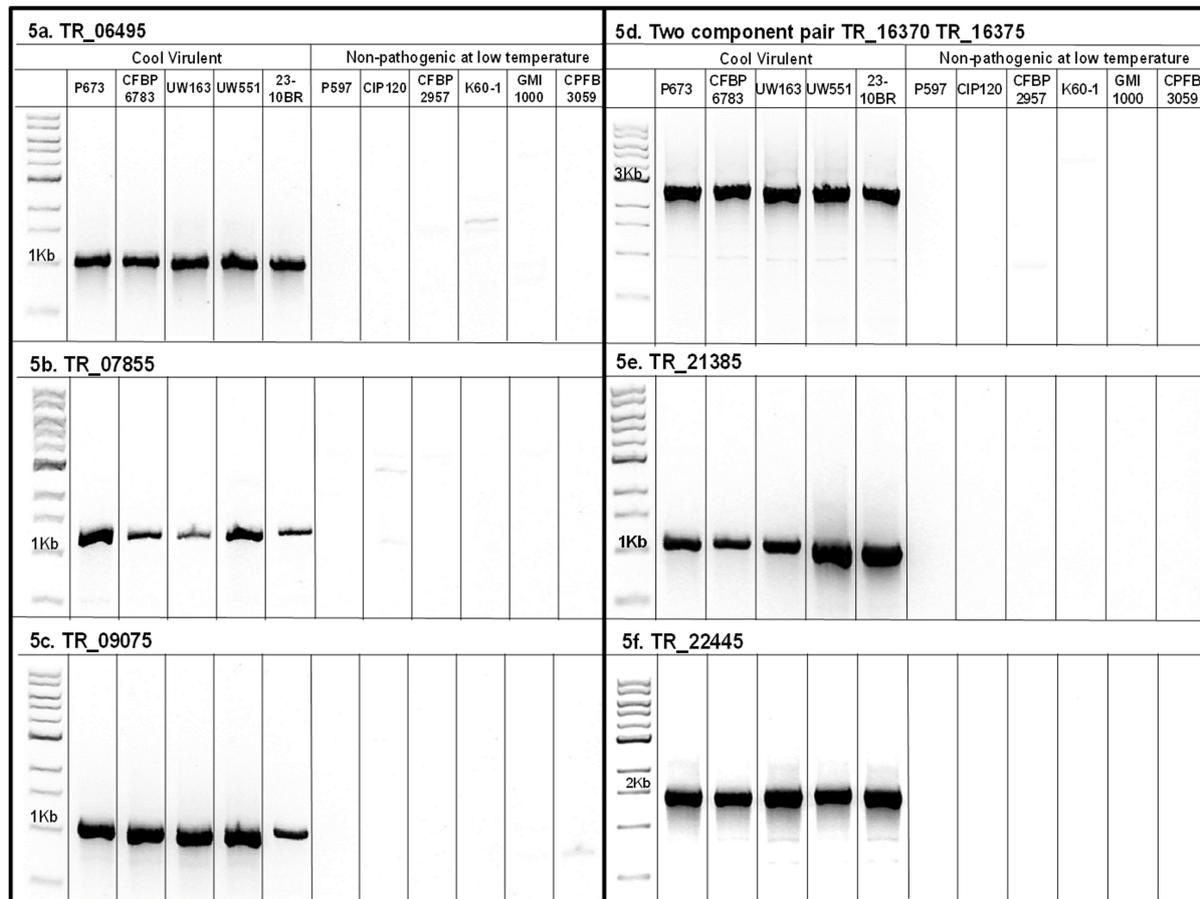


FIGURE 5 | PCR amplification of seven transcriptional regulators found only in the CV strains in the comparison. PCR was run with primers described in Supplementary Table S1 at recommended annealing temperatures. PCR reactions were run in an agarose gel 0.8% 90 V for 1 h. The presence of bands indicates the presence and completeness of the genes in CV strains and their absence in NPLT strains.

In the transcriptional regulators category 12.5% were repressed by SA (Lowe-Power et al., 2016) and 19% were downregulated at low temperatures (Meng et al., 2015). SolR and SolI are LuxR and Lux I homologs that regulate the production of two acyl-homoserine lactones (acyl-HSL). SolI expression is dependent on SolR and the presence of the acyl-HSL. Inactivation of SolI abolishes production of these acyl-HSL but does not affect the expression of virulence genes in culture or pathogenicity in tomato (Flavier et al., 1997). The acyl-HSL regulate the production of *aidA* in strain AW1 which is a phylotype IIA-seq 7 as K60 is, and SolR/SolI regulated expression of *aidA* in R3B2 strains (Meng et al., 2015). *AidA* is not present in the phylotype I and III strains but it is present in all phylotype II strains compared in this work including the NPLT strains K60, P597, P795 (CIP20), and CFBP2957. Interestingly a deletion of *aidA* in UW551 (CV strain) reduced the virulence of the strain in tomato significantly only at low temperature, while a deletion in *solI* did not (Meng et al., 2015). The presence of *aidA* in NPLT strains suggests that *aidA* might be involved in other functions besides cool virulence and/or that there could be more unknown regulatory pathways that might affect other genes related to cool virulence.

Another two component regulation pair of interest is *phcS/phcR*. This pair of genes with *phcB* conform an operon responsive to the quorum sensing signaling molecule 3-OH_PAME. *PhcB* is required for production of 3-OH_PAME and is autoregulated by it. 3-OH-PAME also regulates *PhcS/R*. *PhcS/PhcR* function together to reduce production of *PhcA*-regulated factors. The operon in absence of 3-OH-PAME (low density cells) regulates negatively production or activity of the global regulator *PhcA* at the top a complex hierarchical regulation chain of virulence and survival factors (Clough et al., 1997; Poussier et al., 2003). Additionally, two undescribed transcriptional regulators (*Rsp0645* and *Rsc1610*) in this group were differentially regulated in GMI1000 and UW551 at low temperature (Meng et al., 2015). This fact makes them interesting candidates to be investigated.

Approximately 20–30% of genes with conserved SNPs in CV strains were related to responses to SA presence in the cellular processes, metabolic processes, and stress response categories. In motility and pathogenicity categories, the percentage was 100 and 60%, respectively. Additionally, several key transcriptional regulators such as SolR and *PhcS/PhcS* responsive to different

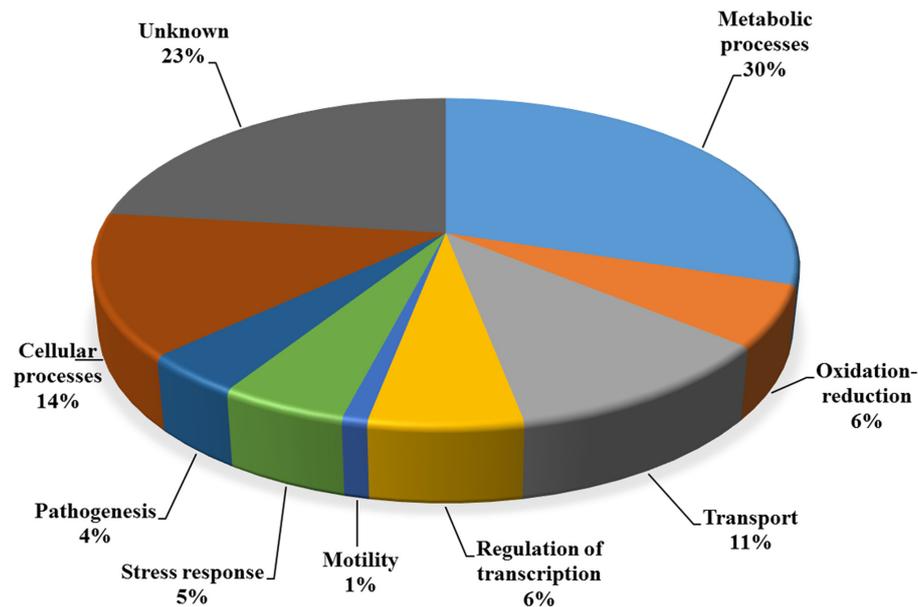


FIGURE 6 | Functional distribution of non-synonymous SNPs conserved in CV strains. GO functional annotations were assigned to genes by biological annotation categories.

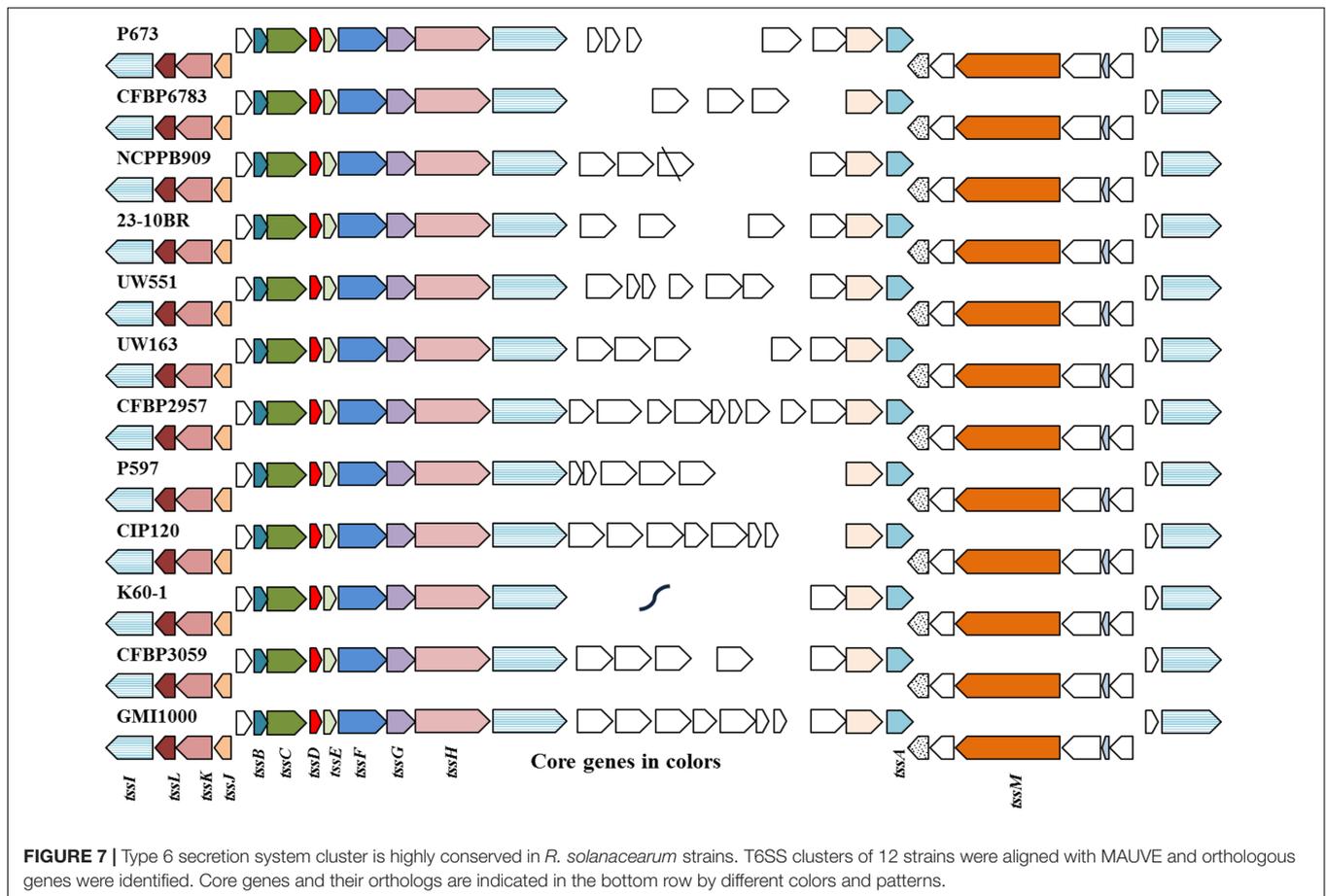
secreted external molecules involved in quorum sensing have conserved SNPs in CV strains. This indicates that those genes directly or indirectly respond to environmental stimuli. It would be logical to think that these genes may respond also to temperature changes and it is possible they will respond differently in CV strains than in NPLT strains.

Type 6 Secretion System

The T6SS is a secretion system recently identified in gram-negative bacteria (Mougous et al., 2006; Schwarz et al., 2010a). It is highly conserved and ubiquitous in many gram-negative bacterial species (Das and Chaudhuri, 2003; Pukatzki et al., 2006). It has been associated with varied functions (Russell et al., 2014), including antimicrobial (Hood et al., 2010; MacIntyre et al., 2010), pathogenesis (Schell et al., 2007; Schwarz et al., 2010b), nodulation inhibitor (Bladergroen et al., 2003), and intra and inter-species competition (Schwarz et al., 2010b; Ma et al., 2014; Shyntum et al., 2015). T6SS is of interest in this comparison because T6SS its expression is different at low temperature between CV strains and NPLT ones (Bocsanczy et al., 2014b). Thirteen core genes have been identified in different bacterial species (Zheng and Leung, 2007), and have been classified in five different clades based on their sequence and structure in the cluster (Barret et al., 2013; Li et al., 2015). In our work, all *R. solanacearum* compared strains encode the 13 core genes and the cluster architecture is a type I subclade 4b. Type I are found in proteobacteria and are the most widespread type of T6SS clusters. As expected, the core genes are highly conserved among all the strains compared (Figure 7). For a complete list of T6SS core genes see Supplementary Table S6A. P799 T6SS cluster was not included in the comparison because the sequence of this region is incomplete in the draft genome. We added a closely related strain

(CMR15) to show that the T6SS cluster conservation includes *R. solanacearum* Phylotype III.

Type 6 secretion system effectors are diverse, relatively unknown secreted proteins, thought to be translocated directly into the host cell cytosol. They can target eukaryotic cells (Pukatzki et al., 2007) or bacterial cells (Hood et al., 2010) depending of their function. We used a combination of Secret6 database searches (Li et al., 2015) and T6SS associated domains defined in the literature to scan the genomes for putative T6SS effectors. We identified 33 putative T6SS effector families (Supplementary Table S6B). VGR were the first effector domains to be identified in T6SS systems (Pukatzki et al., 2007). VGR is a protein that forms a complex with hemolysin coregulated protein (Hcp) to produce the cap structure thought to pierce the hosts cell membrane in an analogous process to virus injection (Pukatzki et al., 2007; Leiman et al., 2009; Silverman et al., 2013). VGR proteins can be found alone or fused with functional domains at their C-terminus. The modular structure of these effectors suggests that VGR in complex with Hcp may facilitate the secretion and translocation of the functional domains (Ho et al., 2014). We identified 21 VGR containing genes in the genomes of the compared strains (Supplementary Table S6B). Three families contained only the VGR domain and the rest were fusions of VGR, Rhs elements, and DUF2345 domain of unknown function. Rhs domains are also common in T6SS effectors, and they are composed of repeats associated with mobile elements and horizontal transfer (Jackson et al., 2009). It has been shown that Rhs proteins are T6SS secreted and mediate intercellular competition in *Dickeya dadantii* (Koskiniemi et al., 2013). We also identified four gene families containing a PAAR motif. Recently PAAR domains have been associated with T6SS effectors (Shneider et al., 2013). In our analysis, we found two



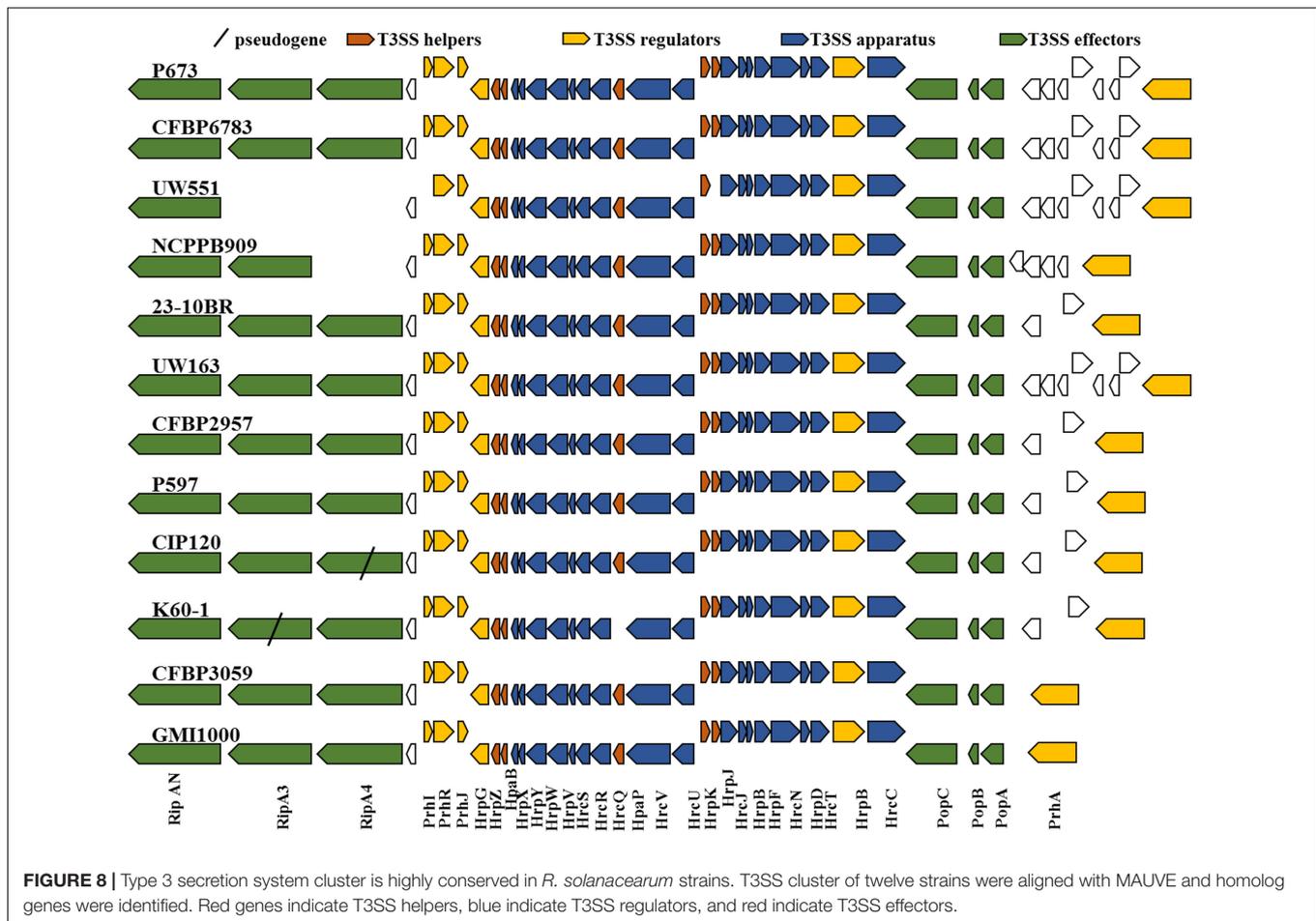
families with only the PAAR motif domain and no conserved domain following it, however, there are VGR-Rhs proteins in close proximity. It is likely that the sequences were not properly annotated and there might be assembly errors due to the presence of the repetitive elements. One family contains the PAAR domain fused with Rhs repeats but no toxin module. Another family has a toxin module domain at the C-terminus. Other domains identified by similarity with known effectors in other bacterial species, included two peptidases, one with fused Rhs and discoidin domains, hemolysin translocator, a DUF1795 unknown conserved domain, two putative alpha/beta hydrolase and a lectin containing a DUF3274 of unknown function. The T6SS cluster contains two VGR proteins and four conserved proteins in all strains compared. None of the VGR proteins seem to be functional in all the strains (Supplementary Table S6A), suggesting that the VGR proteins may be interchangeable during secretion. Conserved proteins in the cluster included a putative lipoprotein, a glycosyl transferase, an OMPA-like cell envelope protein, and a hypothetical protein of unknown function. The fact that those proteins are conserved and complete in all the compared *R. solanacearum* strains, even in *Burkholderia* sp. strains, suggests that they might have an important function in the T6SS secretion. We could not identify any Tae4 superfamily in the genomes analyzed. The Tae4 superfamily is a newly identified form of toxin-antitoxin system associated with T6SS system,

and has anti-microbial functions against other bacterial species (Russell et al., 2012). The lack of these pair of genes suggests that T6SS in *R. solanacearum* may not have antimicrobial function, but it may have a role in inter or intra-species competition. Although there is a high diversity in T6SS effectors, especially in the VGR domain containing families, we could identify only one putative operon that seems to be present exclusively in NPLT strains (Supplementary Table S6B). This operon (Rsp1137, Rsp1138, and Rsp1139 in GMI1000) is composed of a PAAR-Rhs toxin, and VGR domains, a hallmark of an “evolved” VGR T6 effector (Pukatzki et al., 2009), suggesting a role in translocation of toxins by T6SS. The results of our comparison have uncovered several potential functional T6 effectors that we are investigating in more detail.

Type 3 Secretion System

The T3SS cluster is highly conserved in all the different strains. All the core and T3SS accessory genes were present in all the strains compared (Figure 8). This comparison confirms previous observations of high conservation in the T3 cluster genes (Van Gijsegem et al., 1993).

Type 3 secretion system effectors were predicted using the “ScanYourGenome” software and RalstoT3E database (Peeters et al., 2013). “ScanYourGenome” software predicts ORFs based on similarity to known T3E in other bacterial species, presence



of HrpII boxes (to predict regulation by HrpB), and presence of T3SS-dependent export pattern (Petnicki-Ocwieja et al., 2002). The software classifies the predicted T3E by similarity with orthologous families defined in the RalstoT3E database and assigns a unified nomenclature: *Ralstonia* Injected Proteins (Rip). Predicted T3E were further inspected for homology in the group of compared strains, and pseudogenes status verified manually. Frameshifts were assumed as functional in WGS genomes, since a frameshift may be due to errors in sequencing and assembly. The repertoires vary from one strain to the other. We identified 17 core effectors (complete in all the compared strains) of which 14 coincide with the core effectors identified in Peeter's T3E repertoires (Peeters et al., 2013) (Supplementary Table S7). Analogous to Peeter's analysis looking for a correlation between T3E repertoires and host range, we did not find a correlation of T3E repertoire and cool virulence as there was not one putative effector that would be present in all the CV strains compared and absent in the NPLT ones or vice-versa. These results suggest that the T3 effectors are not involved in the CV phenotype.

This genomic comparative analysis of *R. solanacearum* strains complements previous works on transcriptomics and proteomics in the search for temperature related virulence factors. Currently identified CV strains are restricted to phylotype IIB clade sequevar 1 and 4, supporting the idea that the phenotype arose

before the split of sequevars. A high percentage of transcriptional regulators and a practical absence of known virulence related genes that are present only in CV strains supports the hypothesis that regulation of virulence genes could explain the difference in virulence at low temperatures between the groups compared. This study as previous ones on cool virulence factors supports the hypothesis that cool virulence depends on multiple genes and multiple regulation pathways. This comparison contributes to find new possible connections of temperature dependent virulence to the described complex regulatory system involving quorum-sensing, phenotype conversion (phcA), acyl-HSL production and responses to SA. It also identified seven transcriptional regulators with high likelihood to be involved in regulation of cool virulence. This comparative work has also identified for the first time, putative T6 effectors that could be involved in differential fitness or aggressiveness of the strains compared.

AUTHOR CONTRIBUTIONS

AB designed the experiments, performed the assemblies, comparative analysis and experiments, and wrote the manuscript. JH-T supported the bioinformatics work for the comparative

analysis and reviewed the manuscript. DN conceived the work, supervised, coordinated all activities and provided critical review to the manuscript.

ACKNOWLEDGMENTS

We acknowledge the USDA Floral Industry Task Force specific cooperative agreement and the University of Florida Institute of Food and Agricultural Sciences for providing the funds for this

REFERENCES

- Ailloud, F., Lowe, T., Cellier, G., Roche, D., Allen, C., and Prior, P. (2015). Comparative genomic analysis of *Ralstonia solanacearum* reveals candidate genes for host specificity. *BMC Genomics* 16:270. doi: 10.1186/s12864-015-1474-8
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Arnold, D. L., Gibbon, M. J., Jackson, R. W., Wood, J. R., Brown, J., Mansfield, J. W., et al. (2001). Molecular characterization of avrPphD, a widely-distributed gene from *Pseudomonas syringae* pv. *phaseolicola* involved in non-host recognition by pea (*Pisum sativum*). *Physiol. Mol. Plant Pathol.* 58, 55–62. doi: 10.1006/pmpp.2000.0315
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556
- Barret, M., Egan, F., and O’Gara, F. (2013). Distribution and diversity of bacterial secretion systems across metagenomic datasets. *Environ. Microbiol. Rep.* 5, 117–126. doi: 10.1111/j.1758-2229.2012.00394.x
- Bladergroen, M. R., Badelt, K., and Spaink, H. P. (2003). Infection-blocking genes of a symbiotic *Rhizobium leguminosarum* strain that are involved in temperature-dependent protein secretion. *Mol. Plant Microbe Interact.* 16, 53–64. doi: 10.1094/MPMI.2003.16.1.53
- Bocsanczy, A. M., Achenbach, U. C., Mangravita-Novo, A., Yuen, J. M., and Norman, D. J. (2012). Comparative effect of low temperature on virulence and twitching motility of *Ralstonia solanacearum* strains present in Florida. *Phytopathology* 102, 185–194. doi: 10.1094/PHYTO-05-11-0145
- Bocsanczy, A. M., Achenbach, U. C. M., Mangravita-Novo, A., Chow, M., and Norman, D. J. (2014a). Proteomic comparison of *Ralstonia solanacearum* strains reveals temperature dependent virulence factors. *BMC Genomics* 15:280. doi: 10.1186/1471-2164-15-280
- Bocsanczy, A. M., Hugué-Tapia, J. C., and Norman, D. J. (2013). Genome sequencing and analysis of P673, a cool virulent strain of *Ralstonia solanacearum* can reveal virulence factors at low temperatures. *Phytopathology* 103, 18.
- Bocsanczy, A. M., and Norman, D. J. (2013). Twitching motility and a type VI secretion system contribute to virulence in P673, a cool virulent strain of *Ralstonia solanacearum*. *Phytopathology* 103, 17–18.
- Bocsanczy, A. M., Yuen, J. M. F., Palmateer, A. J., and Norman, D. J. (2014b). Comparative genomics of *Ralstonia solanacearum* strain P781 that infects Mandevilla and Dipladenia plants. *Phytopathology* 104, 16.
- Buddenhagen, I., and Kelman, A. (1964). Biological and physiological aspects of bacterial wilt caused by *Pseudomonas solanacearum*. *Annu. Rev. Phytopathol.* 2, 203–230. doi: 10.1146/annurev.py.02.090164.001223
- Burdman, S., Bahar, O., Parker, J. K., and De La Fuente, L. (2011). Involvement of type IV pili in pathogenicity of plant pathogenic bacteria. *Genes* 2, 706–735. doi: 10.3390/genes2040706
- Carver, T., Harris, S. R., Berriman, M., Parkhill, J., and McQuillan, J. A. (2012). Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 28, 464–469. doi: 10.1093/bioinformatics/btr703
- Ciampi, L., and Sequeira, L. (1980). Influence of temperature on virulence of race 3 strains of *Pseudomonas solanacearum*. *Am. Potato J.* 57, 307–317. doi: 10.1007/BF02854025

research. Thanks to Caitilyn Allen at the University of Wisconsin for providing *R. solanacearum* strains for this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2017.01565/full#supplementary-material>

- Clarke, C. R., Studholme, D. J., Hayes, B., Runde, B., Weisberg, A., Cai, R., et al. (2015). Genome-enabled phylogeographic investigation of the quarantine pathogen *Ralstonia solanacearum* race 3 biovar 2 and screening for sources of resistance against its core effectors. *Phytopathology* 105, 597–607. doi: 10.1094/PHYTO-12-14-0373-R
- Clough, S. J., Lee, K. E., Schell, M. A., and Denny, T. P. (1997). A two-component system in *Ralstonia (Pseudomonas) solanacearum* modulates production of PhcA-regulated virulence factors in response to 3-hydroxypalmitic acid methyl ester. *J. Bacteriol.* 179, 3639–3648. doi: 10.1128/jb.179.11.3639-3648.1997
- Conesa, A., Gotz, S., Garcia-Gomez, J. M., Terol, J., Talon, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610
- Darling, A. C., Mau, B., Blattner, F. R., and Perna, N. T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 14, 1394–1403. doi: 10.1101/gr.2289704
- Das, S., and Chaudhuri, K. (2003). Identification of a unique IAHP (IcmF associated homologous proteins) cluster in *Vibrio cholerae* and other proteobacteria through in silico analysis. *In Silico Biol.* 3, 287–300.
- Deberdt, P., Guyot, J., Coranson-Beaudu, R., Launay, J., Noreksal, M., Riviere, P., et al. (2014). Diversity of *Ralstonia solanacearum* in French Guiana expands knowledge of the “emerging ecotype”. *Phytopathology* 104, 586–596. doi: 10.1094/PHYTO-09-13-0264-R
- Delcher, A. L., Phillippy, A., Carlton, J., and Salzberg, S. L. (2002). Fast algorithms for large-scale genome alignment and comparison. *Nucleic Acids Res.* 30, 2478–2483. doi: 10.1093/nar/30.11.2478
- Denny, T. P., Ganova-Raeva, L. M., Huang, J., and Schell, M. A. (1996). Cloning and characterization of tek, the gene encoding the major extracellular protein of *Pseudomonas solanacearum*. *Mol. Plant Microbe Interact.* 9, 272–281. doi: 10.1094/MPMI-9-0272
- Desai, A., Marwah, V. S., Yadav, A., Jha, V., Dhaygude, K., Bangar, U., et al. (2013). Identification of optimum sequencing depth especially for *de novo* genome assembly of small genomes using next generation sequencing data. *PLOS ONE* 8:e60204. doi: 10.1371/journal.pone.0060204
- Doran, A. G., and Creevey, C. J. (2013). Snpdat: easy and rapid annotation of results from *de novo* snp discovery projects for model and non-model organisms. *BMC Bioinformatics* 14:45. doi: 10.1186/1471-2105-14-45
- Elphinstone, J. G. (2005). “The current bacterial wilt situation: a global overview,” in *Bacterial Wilt Disease and the Ralstonia solanacearum Species Complex*, eds C. Allen, P. Prior, and A. C. Hayward (St. Paul, MN: APS Press), 9–28.
- Ewing, B., and Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* 8, 186–194. doi: 10.1101/gr.8.3.186
- Ewing, B., Hillier, L., Wendl, M. C., and Green, P. (1998). Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* 8, 175–185. doi: 10.1101/gr.8.3.175
- Fegan, M., and Prior, P. (2005). “How complex is the “*Ralstonia solanacearum* species complex,” in *Bacterial Wilt Disease and the Ralstonia solanacearum Species Complex*, eds P. P. C. Allen and A. C. Hayward (Madison, WI: APS Press), 449–462.
- Feldbauer, R., Schulz, F., Horn, M., and Rattei, T. (2015). Prediction of microbial phenotypes based on comparative genomics. *BMC Bioinformatics* 16(Suppl. 14):S1. doi: 10.1186/1471-2105-16-S14-S1
- Flavier, A. B., Ganova-Raeva, L. M., Schell, M. A., and Denny, T. P. (1997). Hierarchical autoinduction in *Ralstonia solanacearum*: control of acyl-homoserine lactone production by a novel autoregulatory system responsive to

- 3-hydroxypalmitic acid methyl ester. *J. Bacteriol.* 179, 7089–7097. doi: 10.1128/jb.179.22.7089-7097.1997
- Gabriel, D. W., Allen, C., Schell, M., Denny, T. P., Greenberg, J. T., Duan, Y. P., et al. (2006). Identification of open reading frames unique to a select agent: *Ralstonia solanacearum* race 3 biovar 2. *Mol. Plant Microbe Interact.* 19, 69–79. doi: 10.1094/MPMI-19-0069
- Gardner, S. N., Slezak, T., and Hall, B. G. (2015). kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics* 31, 2877–2878. doi: 10.1093/bioinformatics/btv271
- Goh, C. S., Gianoulis, T. A., Liu, Y., Li, J., Paccanaro, A., Lussier, Y. A., et al. (2006). Integration of curated databases to identify genotype-phenotype associations. *BMC Genomics* 7:257. doi: 10.1186/1471-2164-7-257
- Hayward, A. C. (1964). Characteristics of *Pseudomonas solanacearum*. *J. Appl. Microbiol.* 27, 265–277. doi: 10.1111/j.1365-2672.1964.tb04912.x
- Ho, B. T., Dong, T. G., and Mekalanos, J. J. (2014). A view to a kill: the bacterial type VI secretion system. *Cell Host Microbe* 15, 9–21. doi: 10.1016/j.chom.2013.11.008
- Hood, R. D., Singh, P., Hsu, F., Guvener, T., Carl, M. A., Trinidad, R. R., et al. (2010). A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. *Cell Host Microbe* 7, 25–37. doi: 10.1016/j.chom.2009.12.007
- Huang, Q., and Allen, C. (2000). Polygalacturonases are required for rapid colonization and full virulence of *Ralstonia solanacearum* on tomato plants. *Physiol. Mol. Plant Pathol.* 57, 77–83. doi: 10.1006/pmpp.2000.0283
- Illumina (2011). *Quality Scores for Next-Generation Sequencing. Assessing Sequencing Accuracy Using Phred Quality Scoring. Technical Note: Sequencing.* Available at: http://www.illumina.com/documents/products/technotes/technote_Q-Scores.pdf
- Jackson, A. P., Thomas, G. H., Parkhill, J., and Thomson, N. R. (2009). Evolutionary diversification of an ancient gene family (rhs) through C-terminal displacement. *BMC Genomics* 10:584. doi: 10.1186/1471-2164-10-584
- Jim, K., Parmar, K., Singh, M., and Tavazoie, S. (2004). A cross-genomic approach for systematic mapping of phenotypic traits to genes. *Genome Res.* 14, 109–115. doi: 10.1101/gr.1586704
- Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2014). Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.* 42, D199–D205. doi: 10.1093/nar/gkt1076
- Kang, Y., Liu, H., Genin, S., Schell, M. A., and Denny, T. P. (2002). *Ralstonia solanacearum* requires type 4 pili to adhere to multiple surfaces and for natural transformation and virulence. *Mol. Microbiol.* 46, 427–437. doi: 10.1046/j.1365-2958.2002.03187.x
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199
- Koskiniemi, S., Lamoureux, J. G., Nikolakakis, K. C., t'Kint de Roodenbeke, C., Kaplan, M. D., Low, D. A., et al. (2013). Rhs proteins from diverse bacteria mediate intercellular competition. *Proc. Natl. Acad. Sci. U.S.A.* 110, 7032–7037. doi: 10.1073/pnas.1300627110
- Lambert, C. D. (2002). Agricultural bioterrorism protection act of 2002: possession, use, and transfer of biological; agents and toxins; interim and final rule. (7 CFR Part 331). *Fed. Regist.* 67, 76908–76938.
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Leiman, P. G., Basler, M., Ramagopal, U. A., Bonanno, J. B., Sauder, J. M., Pukatzki, S., et al. (2009). Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4154–4159. doi: 10.1073/pnas.0813360106
- Li, H., Handsaker, B., Wysocker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Li, J., Yao, Y., Xu, H. H., Hao, L., Deng, Z., Rajakumar, K., et al. (2015). SecReT6: a web-based resource for type VI secretion systems found in bacteria. *Environ. Microbiol.* 17, 2196–2202. doi: 10.1111/1462-2920.12794
- Lowe-Power, T. M., Jacobs, J. M., Ailloud, F., Fochs, B., Prior, P., and Allen, C. (2016). Degradation of the plant defense signal salicylic acid protects *Ralstonia solanacearum* from toxicity and enhances virulence on tobacco. *mBio* 7:e00656-16. doi: 10.1128/mBio.00656-16
- Ma, L. S., Hachani, A., Lin, J. S., Filloux, A., and Lai, E. M. (2014). *Agrobacterium tumefaciens* deploys a superfamily of type VI secretion DNase effectors as weapons for interbacterial competition in planta. *Cell Host Microbe* 16, 94–104. doi: 10.1016/j.chom.2014.06.002
- MacIntyre, D. L., Miyata, S. T., Kitaoka, M., and Pukatzki, S. (2010). The *Vibrio cholerae* type VI secretion system displays antimicrobial properties. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19520–19524. doi: 10.1073/pnas.1012931107
- Maddocks, S. E., and Oyston, P. C. (2008). Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* 154(Pt 12), 3609–3623. doi: 10.1099/mic.0.2008/022772-0
- Madeira, H. M. F., and Gabriel, J. E. (2007). Regulation of gene expression in Mycoplasmas: contribution from *Mycoplasma hyopneumoniae* and *Mycoplasma synoviae* genome sequences. *Genet. Mol. Biol.* 30, 277–282. doi: 10.1590/S1415-47572007000200016
- Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., et al. (2012). Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol. Plant Pathol.* 13, 614–629. doi: 10.1111/j.1364-3703.2012.00804.x
- Mattick, J. S. (2002). Type IV pili and twitching motility. *Annu. Rev. Microbiol.* 56, 289–314. doi: 10.1146/annurev.micro.56.012302.160938
- Meng, F., Babujee, L., Jacobs, J. M., and Allen, C. (2015). Comparative transcriptome analysis reveals cool virulence factors of *Ralstonia solanacearum* race 3 biovar 2. *PLOS ONE* 10:e0139090. doi: 10.1371/journal.pone.0139090
- Miller, J. R., Koren, S., and Sutton, G. (2010). Assembly algorithms for next-generation sequencing data. *Genomics* 95, 315–327. doi: 10.1016/j.ygeno.2010.03.001
- Milling, A., Meng, F. H., Denny, T. P., and Allen, C. (2009). Interactions with hosts at cool temperatures, not cold tolerance, explain the unique epidemiology of *Ralstonia solanacearum* race 3 biovar 2. *Phytopathology* 99, 1127–1134. doi: 10.1094/Phyto-99-10-1127
- Mougous, J. D., Cuff, M. E., Raunser, S., Shen, A., Zhou, M., Gifford, C. A., et al. (2006). A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science* 312, 1526–1530. doi: 10.1126/science.1128393
- Norman, D. J., and Yuen, J. M. F. (1997). Pothos (*Epipremnum aureum*), a new host of *Ralstonia (Pseudomonas) solanacearum*. *Phytopathology* 87(Suppl. 6), S70.
- Norman, D. J., and Yuen, J. M. F. (1999). First report of *Ralstonia (Pseudomonas) solanacearum* infecting pot *Anthurium* production in Florida. *Plant Dis.* 83, 300. doi: 10.1094/PDIS.1999.83.3.300A
- Peeters, N., Carrere, S., Anisimova, M., Plener, L., Cazale, A. C., and Genin, S. (2013). Repertoire, unified nomenclature and evolution of the Type III effector gene set in the *Ralstonia solanacearum* species complex. *BMC Genomics* 14:859. doi: 10.1186/1471-2164-14-859
- Pellegrini, M., Marcotte, E. M., Thompson, M. J., Eisenberg, D., and Yeates, T. O. (1999). Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. *Proc. Natl. Acad. Sci. U.S.A.* 96, 4285–4288. doi: 10.1073/pnas.96.8.4285
- Petnicki-Ocwieja, T., Schneider, D. J., Tam, V. C., Chancey, S. T., Shan, L., Jamir, Y., et al. (2002). Genomewide identification of proteins secreted by the Hrp type III protein secretion system of *Pseudomonas syringae* pv. *tomato* DC3000. *Proc. Natl. Acad. Sci. U.S.A.* 99, 7652–7657. doi: 10.1073/pnas.112183899
- Poussier, S., Thoquet, P., Trigalet-Demery, D., Barthet, S., Meyer, D., Arlat, M., et al. (2003). Host plant-dependent phenotypic reversion of *Ralstonia solanacearum* from non-pathogenic to pathogenic forms via alterations in the phcA gene. *Mol. Microbiol.* 49, 991–1003. doi: 10.1046/j.1365-2958.2003.03605.x
- Pukatzki, S., Ma, A. T., Revel, A. T., Sturtevant, D., and Mekalanos, J. J. (2007). Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15508–15513. doi: 10.1073/pnas.0706532104
- Pukatzki, S., Ma, A. T., Sturtevant, D., Krastins, B., Sarracino, D., Nelson, W. C., et al. (2006). Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. *Proc. Natl. Acad. Sci. U.S.A.* 103, 1528–1533. doi: 10.1073/pnas.0510322103
- Pukatzki, S., McAuley, S. B., and Miyata, S. T. (2009). The type VI secretion system: translocation of effectors and effector-domains. *Curr. Opin. Microbiol.* 12, 11–17. doi: 10.1016/j.mib.2008.11.010

- Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Bournsnel, C., et al. (2012). The Pfam protein families database. *Nucleic Acids Res.* 40, D290–D301. doi: 10.1093/nar/gkr1065
- Read, T. D., and Massey, R. C. (2014). Characterizing the genetic basis of bacterial phenotypes using genome-wide association studies: a new direction for bacteriology. *Genome Med.* 6, 109. doi: 10.1186/s13073-014-0109-z
- Remenant, B., Babujee, L., Lajus, A., Medigue, C., Prior, P., and Allen, C. (2012). Sequencing of K60, type strain of the major plant pathogen *Ralstonia solanacearum*. *J. Bacteriol.* 194, 2742–2743. doi: 10.1128/JB.00249-12
- Remenant, B., Coupat-Goutaland, B., Guidot, A., Cellier, G., Wicker, E., Allen, C., et al. (2010). Genomes of three tomato pathogens within the *Ralstonia solanacearum* species complex reveal significant evolutionary divergence. *BMC Genomics* 11:379. doi: 10.1186/1471-2164-11-379
- Russell, A. B., Peterson, S. B., and Mougous, J. D. (2014). Type VI secretion system effectors: poisons with a purpose. *Nat. Rev. Microbiol.* 12, 137–148. doi: 10.1038/nrmicro3185
- Russell, A. B., Singh, P., Brittnacher, M., Bui, N. K., Hood, R. D., Carl, M. A., et al. (2012). A widespread bacterial type VI secretion effector superfamily identified using a heuristic approach. *Cell Host Microbe* 11, 538–549. doi: 10.1016/j.chom.2012.04.007
- Saile, E., McGarvey, J. A., Schell, M. A., and Denny, T. P. (1997). Role of extracellular polysaccharide and endoglucanase in root invasion and colonization of tomato plants by *Ralstonia solanacearum*. *Phytopathology* 87, 1264–1271. doi: 10.1094/PHYTO.1997.87.12.1264
- Salanoubat, M., Genin, S., Artiguenave, F., Gouzy, J., Mangenot, S., Arlat, M., et al. (2002). Genome sequence of the plant pathogen *Ralstonia solanacearum*. *Nature* 415, 497–502. doi: 10.1038/415497a
- Schell, M. A., Ulrich, R. L., Ribot, W. J., Brueggemann, E. E., Hines, H. B., Chen, D., et al. (2007). Type VI secretion is a major virulence determinant in *Burkholderia mallei*. *Mol. Microbiol.* 64, 1466–1485. doi: 10.1111/j.1365-2958.2007.05734.x
- Schwarz, S., Hood, R. D., and Mougous, J. D. (2010a). What is type VI secretion doing in all those bugs? *Trends Microbiol.* 18, 531–537. doi: 10.1016/j.tim.2010.09.001
- Schwarz, S., West, T. E., Boyer, F., Chiang, W. C., Carl, M. A., Hood, R. D., et al. (2010b). Burkholderia type VI secretion systems have distinct roles in eukaryotic and bacterial cell interactions. *PLOS Pathog.* 6:e1001068. doi: 10.1371/journal.ppat.1001068
- Shneider, M. M., Buth, S. A., Ho, B. T., Basler, M., Mekalanos, J. J., and Leiman, P. G. (2013). PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature* 500, 350–353. doi: 10.1038/nature12453
- Shyntum, D. Y., Theron, J., Venter, S. N., Moleleki, L. N., Toth, I. K., and Coutinho, T. A. (2015). *Pantoea ananatis* utilizes a type vi secretion system for pathogenesis and bacterial competition. *Mol. Plant Microbe Interact.* 28, 420–431. doi: 10.1094/MPMI-07-14-0219-R
- Silverman, J. M., Agnello, D. M., Zheng, H., Andrews, B. T., Li, M., Catalano, C. E., et al. (2013). Haemolysin coregulated protein is an exported receptor and chaperone of type VI secretion substrates. *Mol. Cell* 51, 584–593. doi: 10.1016/j.molcel.2013.07.025
- Simao, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351
- Sonnhammer, E. L., Eddy, S. R., Birney, E., Bateman, A., and Durbin, R. (1998). Pfam: multiple sequence alignments and HMM-profiles of protein domains. *Nucleic Acids Res.* 26, 320–322. doi: 10.1093/nar/26.1.320
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447–D452. doi: 10.1093/nar/gku1003
- Tamura, M., and D'Haeseleer, P. (2008). Microbial genotype-phenotype mapping by class association rule mining. *Bioinformatics* 24, 1523–1529. doi: 10.1093/bioinformatics/btn210
- Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28, 33–36. doi: 10.1093/nar/28.1.33
- Tatusova, T., Ciufu, S., Federhen, S., Fedorov, B., McVeigh, R., O'Neill, K., et al. (2015). Update on RefSeq microbial genomes resources. *Nucleic Acids Res.* 43, D599–D605. doi: 10.1093/nar/gku1062
- Vallenet, D., Belda, E., Calteau, A., Cruveiller, S., Engelen, S., Lajus, A., et al. (2013). MicroScope—an integrated microbial resource for the curation and comparative analysis of genomic and metabolic data. *Nucleic Acids Res.* 41, D636–D647. doi: 10.1093/nar/gks1194
- Valls, M., Genin, S., and Boucher, C. (2006). Integrated regulation of the type III secretion system and other virulence determinants in *Ralstonia solanacearum*. *PLOS Pathog.* 2:e82. doi: 10.1371/journal.ppat.0020082
- Van Gijsegem, F., Genin, S., and Boucher, C. (1993). Conservation of secretion pathways for pathogenicity determinants of plant and animal bacteria. *Trends Microbiol.* 1, 175–180. doi: 10.1016/0966-842X(93)90087-8
- Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLOS ONE* 9:e112963. doi: 10.1371/journal.pone.0112963
- Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191. doi: 10.1093/bioinformatics/btp033
- Yuan, K. X., Cullis, J., Levesque, C. A., Tambong, J., Chen, W., Lewis, C. T., et al. (2015). Draft genome sequences of *Ralstonia solanacearum* race 3 biovar 2 strains with different temperature adaptations. *Genome Announc.* 3:e00815-15. doi: 10.1128/genomeA.00815-15
- Zheng, J., and Leung, K. Y. (2007). Dissection of a type VI secretion system in *Edwardsiella tarda*. *Mol. Microbiol.* 66, 1192–1206. doi: 10.1111/j.1365-2958.2007.05993.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Bocsanczy, Huguet-Tapia and Norman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.