Host adaptation and microbial competition drive *Ralstonia solanacearum* phylotype I evolution in South Korea

Maxim Prokchorchik¹, Ankita Pandey¹, Hayoung Moon¹, Wanhui Kim²,³, Hyelim Jeon²,³,
Stephen Poole⁴, Cécile Segonzac²*,⁵, Kee Hoon Sohn¹*,³ and Honour C. McCann⁴,⁶,

¹Department of Life Sciences, Pohang University of Science and Technology, Pohang 37673,
Republic of Korea.
²Department of Plant Science, Plant Genomics and Breeding Institute and Research Institute
of Agriculture and Life Sciences, Seoul National University, Seoul 08826, Republic of
Korea.
³Plant Immunity Research Center, Seoul National University, Seoul 08826, Republic of
Korea.
⁴Department of Molecular, Cellular and Developmental Biology, University of California
Santa Barbara, Santa Barbara, United States of America.
⁵School of Interdisciplinary Bioscience and Bioengineering, Pohang University of Science
and Technology, Pohang 37673, Republic of Korea.
⁶New Zealand Institute for Advanced Study, Massey University Albany, New Zealand.
⁷Max Planck Institute for Developmental Biology, Tübingen, Germany.

*Corresponding authors. Email: csegonzac@snu.ac.kr, khsohn@postech.ac.kr and
honour.mccann@tuebingen.mpg.de.

Keywords: *Ralstonia solanacearum*, population genomics, plant-microbe interactions,
pathogen evolution, microbial competition
Repositories: All genome sequences obtained in this study are deposited to NCBI GeneBank under BioProject number PRJNA593908.

Abstract

Bacterial wilt caused by the *Ralstonia solanacearum* species complex (RSSC) threatens the cultivation of important crops worldwide. The exceptional diversity of type III secreted effector (T3E) families, high rates of recombination and broad host range of the RSSC hinder sustainable disease management strategies. We sequenced 30 phylotype I RSSC strains isolated from pepper (Capsicum annuum) and tomato (Solanum lycopersicum) in South Korea. These isolates span the diversity of phylotype I, have extensive effector repertoires and are subject to frequent recombination. Recombination hotspots among South Korean phylotype I isolates include multiple predicted contact-dependent inhibition loci, suggesting microbial competition plays a significant role in *Ralstonia* evolution. Rapid diversification of secreted effectors present challenges for the development of disease resistant plant varieties. We identified potential targets for disease resistance breeding by testing for allele-specific host recognition of T3Es present among South Korean phylotype I isolates. The integration of pathogen population genomics and molecular plant pathology contributes to the development of location-specific disease control and development of plant cultivars with durable resistance to relevant threats.
Introduction

*Ralstonia solanacearum* is a soil-borne pathogen that causes bacterial wilt disease of over 200 species in 50 different plant families, presenting a major threat to global crop production. This β-proteobacterium colonizes vascular tissues by invading root systems via wounds. Rapid proliferation and biofilm formation block water flow, causing wilt and death [1, 2]. *R. solanacearum* causes severe outbreaks in crops of critical economic and nutritional importance, including banana (*Musa* spp.), potato (*Solanum tuberosum*), tomato (*S. lycopersicum*) and pepper (*Capsicum annuum*) [1]. The broad host range, extensive distribution and genetic heterogeneity of *R. solanacearum* present major obstacles for the design of sustainable disease management strategies.

*R. solanacearum* was initially described as a single species within *Burkholderia* prior to reassignment to *Ralstonia* [3, 4]. Low (< 70%) levels of DNA hybridization between strains indicated *R. solanacearum* is more appropriately described as a species complex (RSSC) [7]. Four RSSC phylotypes were delineated based on 16S–23S rRNA gene intergenic spacer (ITS) regions and the *hrpB* and *egl* genes, coupled with comparative genomic hybridization analysis. Phylotype I is mainly present in Asia, phylotype II mainly in the Americas, phylotype III in Africa and phylotype IV in Japan, Indonesia, Australia and the Philippines [1]. Multi-locus sequence analysis and whole-genome comparisons indicate phylotype I and III are more closely related to each other, while phylotype IV is most divergent. Although we refer to phylotype designations in this paper, taxonomic revisions have been proposed to reclassify the RSSC into three separate species: *Ralstonia pseudosolanacearum* (phylotypes I and III), *R. solanacearum* (phylotype II) and *R. syzygii* (phylotype IV) [5–7].
R. solanacearum phylotypes have ~ 5 Mb bipartite genomes comprised of two replicons 3.5-3.7 Mb and ~ 1.6-2.1 Mb in size. The smaller replicon is conventionally referred to as a megaplasmid, though both replicons encode essential and pathogenesis-related genes. Multipartite genomes are frequently observed among genera harbouring soil and marine bacteria interacting with eukaryotic hosts [8]. The mosaic structure of RSSC genomes attests to the importance of recombination in enhancing the genetic variation evident in this species complex [7, 11]. Natural competence is considered a ubiquitous trait among RSSC phylotypes and has been demonstrated both in vitro and in planta [9, 10]. Coupat et al. (2008) found 80% of 55 RSSC strains tested developed competence in vitro and exhibited single or multiple homologous recombination events up to 90 kb in length. High rates of recombination enhance the potential for rapid adaptation and emergence of novel pathogen variants [11, 12].

Virulence mechanisms in the RSSC include biofilm production, secretion of cell wall degrading enzymes and type III secreted virulence effector (T3E) proteins. Coordinated synthesis of T3Es is achieved by the activity of HrpG/HrpB transcriptional regulators on the hypersensitive response and pathogenicity (hrp) box cis-element in the promoter of T3E genes. RSSC T3Es target various compartments and functions to suppress host resistance and promote bacterial proliferation [13]. The RSSC has a vastly expanded type III effector repertoire compared to other well-studied plant pathogenic bacteria, with over one hundred T3E families identified thus far [14]. In contrast, 70 and 53 orthologous effector families are reported in Pseudomonas syringae and Xanthomonas spp. respectively [15, 16].

Plants have evolved intracellular receptors that monitor for the presence of pathogen effectors and activate a robust defense response often culminating in a form of programmed cell death.
(referred to as a hypersensitive response) around the site of infection and recognition [17].

Effector-dependent host specific interactions have been identified in *R. solanacearum*. A comprehensive survey of *R. solanacearum* phylotype II BS048 effector recognition identified RipAA, RipE1 and RipH2 as cell death inducers in *Nicotiana* spp., tomato and lettuce respectively [18]. The phylotype I GMI1000 RipE1 allele also triggers hypersensitive response in *Nicotiana benthamiana* [19, 20]. RipA family AWR motif-containing effectors are strong inducers of cell death in *Nicotiana* spp. [21]. The loss of RipAA was recently found to be correlated with the emergence of phylotype IIB strains pathogenic on cucurbits; related strains retaining RipAA are pathogenic on banana but not cucurbits [22].

We sequenced 30 RSSC phylotype I isolates from tomato and capsicum to identify the population structure and virulence factor repertoire of pathogens limiting production of these key crops in South Korea. Location and host of isolation do not have a strong impact on the population structure of South Korean phylotype I isolates; diverse phylotype I isolates capable of infecting solanaceous crops are broadly distributed across South Korea. Loci linked with competition are frequent targets of recombination, suggesting microbial interactions play an important role in *Ralstonia* evolution. We examined South Korean phylotype I virulence factor evolution in the context of the global pool of RSSC T3Es and found that most strains carry large effector repertoires with evidence of frequent horizontal transfer. We identified T3Es that elicit stable host recognition and defense responses regardless of which allelic variant is used, indicating they would be suitable targets for disease resistance breeding efforts. We also identified allelic variation in two T3Es resulting in the loss of host recognition. Rapid diversification of secreted effectors present challenges for the development of resistant plant varieties. The integration of pathogen population genomics with molecular plant pathology contributes to location-specific pathogen control strategies, as well as the development of cultivars with durable resistance to relevant threats.
Methods

Genome assembly and annotation

*Ralstonia solanacearum* (*Rso*) were isolated from field-grown pepper and greenhouse tomato plants between 1999 and 2008 (Table 1). Tomato isolates were collected from the north in 2008, while pepper isolates were collected from central and southern regions of the South Korean peninsula. One isolate (*Pe_4*) was collected from Jeju island, southeast of the peninsula. Genomic DNA was extracted using the Promega Wizard Genomic DNA Purification kit and 150 bp paired-end sequencing was run on the Illumina NextSeq platform using Nextera and the library preparation protocol described by Baym et. al (2015) [23].

Reads were filtered to remove adapter sequences (parameters: ktrim=r k=23 mink=11 hdist=1 minlen=50 ftm=5 tpe tbo) and common sequence contaminants according to NCBI Univec database (parameters: k=31 hdist=1), and quality-trimmed using the bbduk program available in the BBmap package (paramters: qtrim=r trimq=10 minlen=50 maq=10 tpe).

Filtered and quality-trimmed reads were used for genome assembly with SPAdes v3.10.0, using --careful assembly with kmer size distribution –k 21, 33, 55 and 77 [24]. The coverage cutoff was set to auto. Coverage was assessed by mapping reads to the assembly using Bowtie2, removing duplicates [25]. Assembly improvement was performed using Pilon v1.14 and assembly statistics obtained using assembly-stats (https://github.com/sanger-pathogens/assembly-stats) [26]. Genomes were annotated with Prokka v.1.12 beta using a custom annotation database created using all 71 *Ralstonia solanacearum* annotated assemblies available on Genbank Nov. 11, 2017 [27]. 68 publicly available *R. solanacearum* genomes were reannotated with Prokka v1.13.3 as well for inclusion in downstream analyses (Table S1).

Nanopore sequencing was performed for five isolates: *Pe_1*, *Pe_3*, *Pe_27*, *Pe_39* and *Pe_57*. DNA extraction was performed using Promega Wizard, and barcoding and library
preparation were performed using Native Barcoding Expansion 1-12 (EXP-NBD104) kit, in conjunction with the Ligation Sequencing Kit (SQK-LSK109) according to manufacturer recommendations. After sequencing run of 48 h, basecalling was performed using Guppy followed by hybrid assembly with Unicycler using Illumina reads for each respective strain [28]. Genomes were annotated with Prokka v1.13.3 using the custom *R. solanacearum* annotation database.

**Core genome phylogeny of Korean phylotype I isolates**

The core genome phylogeny of South Korean phylotype I isolates was obtained by mapping Illumina reads to the completed nanopore assembly of *R. solanacearum* Pe_57. Regions of the reference genome corresponding to mobile genetic elements were removed after identification using the PHASTER phage and ISFinder databases; no integrative and conjugative elements were identified with ICEberg [29–31].

Readmapping of filtered, quality-trimmed reads was performed using BWA mem 0.7.17 -r1188 [32]. Samtools v1.3.1 was used to sort and index reads, remove duplicate reads and generate pileup data [31]. Variants were called with BCFtools v1.3.1 (bcftools mpileup --consensus-caller --ploidy 1 --min-MQ 60) and a set of filtered SNPs obtained using BCFTools v0.1.13 to retain single nucleotide polymorphisms (SNPs) with a minimum read depth of 10 in both directions, non-reference allele frequency of 0.95, minimum quality of 30 and minimum distance from a gap of 10bp [33]. BEDtools v2.26.0 was used to determine positions where the depth of coverage fell below 10 reads [34]. A final consensus relative to the reference was obtained with BCFtools consensus v1.3.1, masking low coverage positions identified by BEDtools, retaining both invariant sites and SNPs passing coverage, non-reference allele frequency and quality thresholds. An alignment was generated using the
consensus sequences of each isolate. Alignment positions with one or more gaps or overlapping mobile genetic elements in the reference genome were removed. Initial tree building was performed with RAxML v7.2.8 using parameters: `-f a -p $RANDOM -x $RANDOM -# 100 -m GTRGAMMA -T 8 [35].

ClonalFrameML v1.178 was employed to identify polymorphisms introduced via homologous recombination using the gap-free core genome alignment and initial maximum likelihood phylogeny as input [36]. Alignment regions identified as likely to have been introduced via recombination were removed from the alignment and tree building was repeated using the non-recombinant alignment in RAxML v7.2.8 (parameters as above).

Trees were visualized using FigTree v1.4.3, displaying only nodes with bootstrap support scores of 70 and above (https://github.com/rambaut/figtree/). This approach was employed to obtain a core genome phylogeny and alignment using the primary and secondary replicons of R. solanacaerum Pe_57.

Recent and ancestral recombination events among South Korean strains phylotype I strains were visualized using maskrc-svg.py with the core genome alignment along with the recombinant regions predicted by ClonalFrameML (https://github.com/kwongj/maskrc-svg). Recombination events are referred to as ancestral when predicted to occur at internal (ancestral) nodes, and extant or recent events when impacting terminal nodes (individual strains). The number of recent recombinant events was calculated with pyGenomeTracks by counting the number of extant (recent) recombination events overlapping with annotated coding sequences in the reference genome GMI1000 chromosome and megaplasmid, excluding recent events in GMI1000 and Pe_57 (https://github.com/deeptools/pyGenomeTracks).
**R. solanacearum species complex phylogeny**

The *R. solanacearum* species complex tree was generated using an alignment of concatenated single-copy core genes. All Korean isolates and 68 publicly available *R. solanacearum* genomes were annotated with Prokka (as above). OrthoFinder was implemented to identify single copy orthologs [37]. Single copy ortholog nucleotide sequences were employed for codon-aware alignment with PRANK [38]. Core gene alignments were concatenated and stripped of gap and invariant positions. Maximum likelihood phylogeny reconstruction was then performed with RAxML v8.2.10 (parameters: -f a -p $RANDOM -x $RANDOM -# 100 -m GTRCATX -T 16) [36]. Trees were visualized using FigTree v1.4.3, displaying only nodes with bootstrap support scores of 70 and above (https://github.com/rambaut/figtree/).

Estimation of the population structure was performed with RhierBAPS using the core SNP alignment with two levels of clustering and setting the number of initial clusters to 20 [39].

**Type III effector identification and comparative analyses**

Effector predictions were performed on the South Korean phylotype I isolates and 68 publicly available genomes. First, sequences from the manually curated *R. solanacearum* effector (RalstoT3E) database were downloaded from (http://iant.toulouse.inra.fr/bacteria/annotation/site/prj/T3Ev3) and used as queries in tBLASTn searches of the 30 Korean *R. solanacearum* genomes, retaining all hits with a minimum coverage of 30% and pairwise identity of 50%. Overlapping hits on the subject sequence were merged to generate the longest hit using BEDtools. The merged hit was then scanned to identify the predicted open reading frame (ORF), searching for start and stop codons and the presence of any premature stop codons. The presence of putative *hrp* boxes up to 500 bp upstream of the effector candidate ORF start codon was retained. Merged hits
containing no start codons and/or premature stop codons were adjusted to reconstitute a functional ORF if possible. Start codon adjustment was performed by searching between the merged hit and putative hrp box (or up to 500 bp upstream of the merged hit for a start codon). Effector candidates were classified as pseudogenes in the absence of a start codon or in the presence of a stop codon resulting in a truncated ORF 30% shorter than the shortest query sequence generating the initial hit.

Individual effector phylogenies were generated using PRANK codon-aware alignments followed by tree building using RAxML 8.2.12. Gene tree topologies were compared with the species complex core gene phylogeny using the normalized Robinson-Foulds metric implemented in ETE3, collapsing identical sequences into the same leaf and disregarding comparison between nodes with bootstrap scores lower than 70 [40].

**Bacterial strains and plant growth conditions**

*Escherichia coli* DH5α and *Agrobacterium tumefaciens* AGL1 were grown in Luria-Bertani (LB) medium supplied with appropriate antibiotics at 37°C and 28°C, respectively. *Nicotiana benthamiana* and *N. tabacum* W38 plants were grown at 25°C in long-day conditions (14 h light/10 h dark).

**Plasmid construction for transient expression of effectors**

*R. solanacearum* genomic DNA was isolated using the Promega Wizard bacterial genomic DNA extraction kit following the manufacturer’s protocol. Allelic variants of RipA1 and RipAA were cloned from South Korean phylotype I isolates as well as GMI1000 and BS048 using Golden Gate [41]. Alleles were PCR amplified using Takara DNA polymerase (Thermo) and cloned into pICH40121 as ~1 kb modules with Smal/T4 ligase (New England Biolabs). The modules in pICH40121 were assembled into binary vector pICH86988 under
control of the constitutive 35S cauliflower mosaic virus promoter and in C-terminal fusion
with 6xHA or 3xFLAG epitope. The binary constructs were mobilized into Agrobacterium
tumefaciens strain AGL1 using electroporation for transient expression experiments in
Nicotiana spp.

**Hypersensitive response and ion leakage assays**

A. tumefaciens AGL1 carrying effector constructs were grown on LB medium supplied with
appropriate antibiotics for 24 h. Cells were harvested and resuspended to final OD_{600} of either
0.4 (all constructs other than RipH2) or 0.8 (RipH2) in infiltration medium (10 mM MgCl₂,
10 mM MES pH 5.6). A blunt syringe was used to infiltrate the abaxial surface of 5-week old
N. benthamiana or N. tabacum leaves. The symptoms of hypersensitive response were
observed and photographed at 4 days post infiltration (dpi) for N. benthamiana or 6-7 dpi for
N. tabacum.

Ion leakage assays were performed after infiltration of the same AGL1 strains into N.
benthamiana or N. tabacum leaves. Plants were infiltrated with bacterial inoculum (OD_{600} 0.4)
and two leaf discs of 9 mm diameter each were harvested daily from 0 to 6 dpi. The leaf discs
were placed in 2 ml of distilled water (Millipore MilliQ system, Bedford MA) with shaking
at 150 rpm for 2 h at room temperature. Ion leakage was measured using a conductivity meter
(LAQUAtwin EC-33, Horiba) by loading 60 µl from each sample well. 3-4 technical
replicates were used for each treatment. Statistical differences between treatments was
assessed using Students two-tailed T-test with a p-value cutoff of 0.05.
Results and Discussion

Diversity and distribution of *R. solanacearum* phylotype I isolates across South Korea

Assemblies of thirty isolates of *R. solanacearum* collected from pepper and tomato in South Korea between 1999 and 2008 were generated from paired-end Illumina sequencing. Hybrid assemblies were generated with additional Nanopore sequence data for five isolates (Table 1). SPAdes Illumina assemblies were on average 5.7 Mb long, with scaffold counts ranging from 200 (Pe_26) to 306 (To_42). Unicycler hybrid assemblies for five isolates (Pe_1, Pe_3, Pe_27, Pe_39 and Pe_57) were on average 5.85 Mb in length across 2-9 scaffolds. Pe_39 and Pe_57 hybrid assemblies are considered complete as they are comprised of 2-4 circularized replicons.

Maximum likelihood trees were built using non-recombinant core genome alignments generated by readmapping to each Pe_57 replicon as reference. Recombination events identified by ClonalFrameML were removed to obtain non-recombinant trees of South Korean phylotype I strains (Fig. 1). Phylogenies produced using either the chromosome or megaplasmid as reference share nearly identical topologies, with four major clades of isolates and a single representative of a divergent lineage sampled in the eastern mainland (Pe_57).

The chromosomal tree is slightly more resolved and has shorter branches leading to clades 1 and 2 than the megaplasmid tree. The two larger clades include strains isolated from both hosts: clade 2 includes isolates from central and northern regions while clade 4 includes spanning the entire country, including Jeju Island. A sister group to clade 2 includes solely pepper isolates (Pe_26, Pe_28, Pe_39 and Pe_1), while a smaller sister group to clade 4 includes a pair of tomato isolates (To_53, To_22). Other than these small clades, there is no strong phylogenetic signal of host of isolation: tomato and pepper isolates are distributed across the phylogeny. Although all but one tomato isolate was sampled from a single northern province and pepper isolates were sampled from seven other provinces, the tomato
isolates show no evidence of phylogeographic clustering. Almost every province sampled has isolates from at least two clades, indicating diverse phylotype I isolates are circulating across many agricultural regions in South Korea.

Recombination of genes linked with virulence and competition

The cumulative impact of recombination on South Korean phylotype I isolates is enormous: 64.2% (2,026,835) of the gap-free non-mobile 3,155,880bp chromosome and 89.0% (1,511,546) of the gap-free non-mobile 1,698,146bp megaplasmid alignments are predicted to be affected by recombination in one or multiple strains at some stage in their evolutionary history (Table 2). Recombination occurs at only a slightly reduced frequency relative to mutation on both replicons but is far more likely to introduce substitutions, with ratios of the effect of recombination/mutation at 4.3 and 5.4 for the chromosome and megaplasmid, respectively. Fewer non-recombinant variant sites were retained for the megaplasmid (614) than the chromosomal (4,260) core alignment.

The location of recombinant events within South Korean phylotype I genomes was assessed to determine whether genes affected by recombination are of relevance to pathogenesis in R. solanacearum. Extensive recombination events are present at ancestral nodes in the tree, while the range of predicted recent strain-specific recombination events varies between single recombination events for Pe_15 and Pe_28 to 27 recombination events across both replicons for To_42 and To_63. Since Pe_57 is a unique representative of a separate lineage in the phylogeny, many apparent strain-specific recombination events (118 and 77 on the chromosome and megaplasmid, respectively) are likely to have occurred earlier in the history of this lineage. There are distinct hotspots of recombination on both replicons (Fig. 2 and...
Table S2). 62 chromosomal and 46 megaplasmid genes are affected by two or more recent recombination events, some of which encompass multiple coding sequences (Table S2).

Many genes subject to recent recombination events are linked with virulence and microbial competition. Strikingly, six of these recombination hotspots overlap with six out of seven clusters of hemagglutinin and hemolysin (shlB) genes in Pe_57 (Fig. 2). Some of these clusters appear to be contact-dependent inhibition (CDI) loci, whose products mediate self/non-self-recognition and direct interference competition [42–44]. Contact-dependent inhibition is the product of two-partner (type V) secretion systems that assemble outer-membrane proteins (CdiB) involved in the export and presentation of large, toxic CdiA effectors on the cell surface [45]. CdiA effector protein binding to outer membrane receptors is followed by the transfer of the C-terminal toxin domain (CdiA-CT) into the susceptible bacterial cell cytoplasm, resulting in cell death. A cognate immunity protein (CdiI) confers resistance to toxin production by neighboring sister cells by binding and neutralizing specific CdiA-CT toxins; CdiI does not confer immunity to heterologous CdiA-CT toxins. CdiA-CT and CdiI sequences are highly variable both within and between species. Two clusters in Pe_57 (shlB_1 and shlB_4) have cdIA-cdiI-cdiO-cdiB(shlB) arrangements typical of CDI loci in Burkholderia, as well as small ORFs and IS elements between cdiI and cdiB. Comparison of the shlB_4 clusters in isolates for which nanopore assemblies are available (Pe_57, Pe_39, Pe_1, Pe_27 and Pe_3) reveals Pe_3 diverges significantly from other strains in the C-terminal domain of CdiA and cognate CdiI protein (Fig. S1). Recombination at CDI loci can be a mechanism promoting the maintenance of immunity between otherwise divergent isolates (e.g. Pe_57 vs Pe_39, Pe_1 and Pe_27) or a means of promoting antagonism via rapid diversification in toxin and cognate immunity proteins (e.g. Pe_27 vs Pe_3). The
presence of multiple CDI loci in a single strain suggest both processes may operate simultaneously.

Recombination targets additional genes involved in microbial interactions and virulence. Up to eleven strain-specific recombination events were identified in genes annotated as D-alanine-polyphoribitol ligase subunit 1 (dltA-4) and tyrocidine synthase tycC-2 genes by Prokka. The closest BLASTn hit for this region in the phylotype I isolate GMI1000 is to a non-ribosomal peptide synthase locus called ralsomycin (rmyA and rmyB) or ralstonin which is linked with *R. solanacearum* invasion of fungal hyphae and the formation of chlamydospores [46, 47]. Five chromosomal (ripA1, ripB, ripG6, ripG7, ripTAL) and three megaplasmid (ripH2, ripD, ripAW) effectors show evidence of two or more strain-specific recombination events among the South Korean phylotype I isolates. Flagellar genes (*fliI* and *fliJ*), a type IV prepilin-like protein (*plpA*, annotated by Prokka as *pilE1*), a type III secretion system translocation protein (*yscU*) and component of the export pathway for enterobactin (*entS*), a high-affinity siderophore are also affected by multiple strain-specific recombination events in South Korean phylotype I isolates. The detection of recombination within gyrase B and endoglucanase is particularly notable given these are frequently used to infer relationships between *Ralstonia* strains.

**Core gene tree of the *R. solanacearum* species complex**

In order to place the new South Korean strains and T3E dynamics in context of the structure of *R. solanacearum* species complex, a maximum-likelihood phylogeny was inferred from a concatenated alignment of core genes identified in an OrthoFinder pangenome analysis (Fig. 3). The core genome of the 30 South Korean phylotype I genomes and 68 NCBI genomes is...
represented by 1,680 single-copy orthogroups; 17.9% of the species complex pangenome. The *R. solanacearum* species complex has a large flexible genome, with 39.7% (3,720/9,366) of orthogroups present in 15 to 97 genomes and 40.1% of orthogroups in the cloud genome (1 to 15 genomes).

The four major phylotypes (and three separate species) are displayed in the core gene tree. The first level of population clustering is consistent with phylotype designations I, IIA, III and IV, while phylotype IIB falls into two separate clusters. Clonal expansion of multiple groups is evident: a phylotype IIB subclade (3.3) typically associated with *S. tuberosum* infections has been sampled from three continents (South America, Africa and Eurasia). The two larger clades of South Korean phylotype I isolates correspond to subclades 6.2 and 6.4, primarily sampled from solanaceous hosts in China, Taiwan and South Korea (Fig. 3). The divergent pepper isolate Pe_57 falls into a cluster of East Asian isolates from different hosts: eggplant (EP1, OE1-1); tomato (PSS1308, PSS216, Rs-To2); tobacco (CQPS-1) and squash (RSCM).

**Distribution of type III secreted effectors across the RSSC**

The distribution of type III secreted effectors across the RSSC was determined to gain an understanding of the extent of variation within Korean isolates and place these in the context of the variation in the RSSC as a whole (Fig. 3B). The average T3E complement varies between phylotypes: phylotype I has on average 71 effector genes per strain, while phylotype II has an average of 61 effector genes per strain. Phylotypes III and IV are each represented by a pair of strains: these have an average of 58 and 64 T3Es per strain respectively. Many T3E families are broadly distributed across the entire species complex: 64 out of 118 (54%) of T3Es are present in at least half of the isolates included here. *RipAB* is present in all strains.
but YC40-M (*Kaempferia* sp., China) and *ripU* is potentially pseudogenised in CFBP3059 (*Solanum melongena*, Burkina Faso) but is otherwise ubiquitous. There are some highly amplified T3E families present across much of the species complex, like *ripA* and *ripB*. Diversification of these families may be linked with host recognition; the broad distribution suggests they encode critical virulence or host subversion functions.

There are both clade- and phylotype-specific patterns of T3E presence and absence. A set of fourteen T3Es (*ripA1, ripAG, ripAK, ripBA, ripBE, ripBJ, ripBK, ripBO, ripBP, ripG1, ripS6, ripS8, ripT and Hyp6*) are present in one or more phylotype I strain, but are absent from phylotype II. This is likewise the case for *ripAH*, though it is present in two phylotype II strains. Three T3Es (*ripJ, ripS5 and ripTAL*) are mostly truncated or pseudogenized in phylotype II, while full-length versions are observed in phylotype I. Conversely, seven T3Es (*ripBC, ripBG, ripBH, ripBI, ripK, ripS7 and ripV2*) and many hypothetical T3Es (*Hyp1, Hyp3, Hyp4, Hyp8, Hyp9, Hyp11 and Hyp12*) are present in phylotype II but absent in phylotype I.

There are some apparent patterns of recurrent T3E loss across phylotypes I and II. In phylotype II these losses are neither clade nor host-specific, affecting IPO1609 (*S. tuberosum*, Netherlands), Po82 (*S. tuberosum*, Mexico), RS2 (*S. tuberosum*, Bolivia), UW551 (*Geranium sp.*, Kenya), UW25 (*S. lycopersicum*, USA), POPS2 (*S. lycopersicum*, China), RS489 (*S. lycopersicum*, Brazil). The apparent losses in phylotype I impact tomato strains PSS190 from Taiwan, UTT-25 from India, KACC10709, To_22 and To_53 from South Korea. In addition, Cars-Mep strain isolated from *Elettaria cardamom* from India shows significant effector losses. These patterns should be interpreted with caution as some apparent losses (or indeed truncations and pseudogenizations) may be due to variable sequencing and assembly quality.
Despite their ubiquity across the species complex, T3Es are nevertheless subject to recombination. Individual T3E gene trees were compared with the core gene tree to determine whether they exhibit evidence of horizontal transfer. The normalized Robinson Fould's (nRF) distance implemented in ETE3 was used as a measure of topological similarity between individual T3E trees and the core gene tree. Some T3Es, including ripAK, ripBD, ripAH, ripP1, ripC2, ripAG, ripG1 and ripT have divergent tree topologies (nRF > 0.8), indicating they have been subject to dynamic exchange throughout their evolutionary histories. T3Es with topologies consistent with the core gene tree are typically limited to specific clades (e.g. ripK, ripBO, ripF2 and ripS7, with nRF < 0.2). There are however three exceptions: ripG5, ripH2 and ripS5 are broadly distributed across the RSSC and are less frequently exchanged, though ripS5 is pseudogenized in many phylotype II strains. Allelic variation is likely to determine fine-scale differences in host range among T3Es with broad distributions across the *Ralstonia* species complex.

**Allelic diversity in T3Es impacts host recognition**

Allelic diversification of T3Es is likely to play an important role in host adaptation. The link between defense induction and allelic diversity among South Korean isolates was examined for six T3Es: ripA1, ripAA, ripAY, ripE1, ripA5 and ripH2. These families were chosen as they were reported to trigger programmed cell death when overexpressed in *Nicotiana* spp. [18–22]. Alleles were classified into groups on the basis of shared amino acid sequence identity; patterns of diversity broadly reflected the phylogenetic groupings in the non-recombinant core genome tree for South Korean strains phylotype I strains (Fig. 4A, B).
Representative alleles from each group (Table S3) were transiently expressed in *N. benthamiana* and *N. tabacum* plants by agroinfiltration in order to score their ability to elicit a programmed cell death and ion leakage, markers of host recognition. Overexpression of RipA1 group 2 and group 4 alleles failed to elicit cell death, while group 1 and group 3 alleles triggered cell death and ion leakage (Fig. 4C, D). One of two substitutions present solely in groups 2 and 4 is likely responsible for the loss of recognition (V1002M); an additional G710D substitution is present in group 4. Host recognition of RipAA alleles varied as well: a single allele (group 2) failed to elicit cell death. A C-terminal truncation preceded by multiple substitutions is likely responsible for the loss of recognition of group 2 RipAA in *N. benthamiana* (Fig. 4C, D). We found RipA1 and RipAA protein variants accumulated to significant levels *in planta*, suggesting the phenotypes observed are not due to variation in protein stability but rather the loss of host recognition (Fig. S2). Both RipAY alleles present among South Korean phylotype I strains trigger weak cell death and intermediate ion leakage, while all RipE1 and RipA5 alleles caused strong cell death and ion leakage (Fig. S3). In contrast, no RipH2 alleles triggered cell death and ion leakage when overexpressed in *N. tabacum* (Fig. S3).
Conclusions

*Ralstonia* infections of staple carbohydrate and vegetable crops threaten regional food security and agricultural productivity. The 30 South Korean isolates shown here encompass much of the diversity of phylotype I. Phylotype I is likely to have been present in East Asia for some time, though recent clonal expansions and dissemination between agricultural regions is apparent among South Korean isolates. High rates of recombination among the *R. solanacearum* species complex enhance the potential for rapid adaptation and the emergence of novel pathogen variants. Recombination affecting multiple virulence-related genes was identified among the South Korean strains. Hotspots of recombination overlap with contact-dependent inhibition loci, indicating microbial competition plays a critical role in *Ralstonia* evolution, perhaps during stages in its lifecycle when it must compete for nutrients in soil and freshwater environments.

The *R. solanacearum* species complex has an expanded T3E repertoire relative to other bacterial plant pathogens. Many T3Es have evolutionary histories inconsistent with the core gene topology, revealing how recombination can reshuffle virulence factor repertoires. Though some effector families may be less frequently horizontally transferred than others, host adaptation and selection imposed by host recognition can result in allelic variation linked with altered host specificity. The integration of pathogen population genomics with molecular plant pathology allows for the identification of conserved pathogen virulence proteins with consistent recognition and elicitation activity, providing suitable targets for breeding crops with durable resistance to *R. solanacearum* phylotype I in South Korea.
Acknowledgements

We gratefully acknowledge support for this work from the following sources: Next-Generation BioGreen 21 Program (Plant Molecular Breeding Center No. PJ01317501), Rural Development Administration (RDA) and the National Research Foundation (NRF) of Korea grants funded by the Korea government (MSIT) (NRF-2019R1A2C2084705 & 2018R1A5A1023599 (SRC)), Republic of Korea (K.H.S.); Creative-Pioneering Researchers Program through Seoul National University (C.S.); Royal Society of New Zealand Marsden Fast-Start (MAU1709) and support from the Max Planck Society (H.C.M.). We thank Sven Kuenzel (MPI for Evolutionary Biology) for Illumina sequencing.

Conflicts of Interest

The authors declare that there are no conflicts of interest.
Figure 1. Diversity and distribution of South Korean phyloptype I isolates. (A) Non-recombinant core chromosomal tree and (B) non-recombinant core megaplasmid tree of *R. solanacearum* phyloptype I strains isolated from tomato (To) and pepper (Pe) hosts across South Korea (C). Core alignments were generated by readmapping to separate Pe_57 reference replicons. Phylogenetic trees were built with RAxML v7.2.8 and visualized using FigTree v1.4.3 displaying only the branches with bootstrap support score above 70. Node labels indicate the bootstrap support.
Figure 2. Recombination in South Korean phylotype I isolates. Recombination events identified by ClonalFrameML in chromosomal (A) and megaplasmid (B) sequences of South Korean R. solanacearum strains. Ancestral recombination events are shown in grey, and recent/strain-specific recombination events shown in red. Sum of strain-specific events shown at the top of each panel. Genes with two or more strain-specific events have gene IDs listed (hypothetical proteins not listed).
Figure 3. Phylogeny and type III secreted effector repertoires of the *R. solanacearum* species complex. (A) Maximum likelihood tree inferred using a non-recombinant core gene alignment of single-copy orthologs. Species and phylotype designation shown above and below branches, respectively. Nodes with bootstrap support values under 70 collapsed. Primary (1°) and secondary (2°) cluster assignment by hierBAPS shown next to strain labels. (B) Type III effector repertoires. Top row displays normalized Robinson Fould’s distance between the effector family gene tree and core gene tree shown at left. Lower values (e.g. blue) indicate greater topological similarity between effector and core gene tree.
Figure 4. Effector alleles of RipA1 and RipAA display variable recognition responses in *N. benthamiana*. (A) Effector alleles were classified into groups (A) based on amino acid sequence alignments (B) and representatives of each group were tested for their ability to trigger programmed cell death (PCD, shown in C) and increased ion leakage (D), markers of host recognition. *Agrobacterium* strains expressing allelic variants were infiltrated into *N. benthamiana* leaves at OD₆₀₀ 0.4 and PCD was photographed 3 days post infection (dpi). Asterisks in (C) indicate absence of PCD for RipA1 allele groups 2 and 4, and RipAA group 2. GFP is a negative control, Pto+AvrPto, GMI1000 RipA1 and BS4048 RipAA refer to positive controls. Ion leakage in (D) was measured at 0, 1, 2 and 3 dpi and plotted as average conductivity of 4 technical replicates, error bars represent standard error of mean (SEM).
Figure S1. Contact-dependent inhibition loci in *Ralstonia* phylotype I isolates. CDI regions homologous to the *shlB_4* region in Pe_57 were identified in nanopore assemblies of South Korean phylotype I isolates as well as the complete genome of GMI1000, *B. thailandensis* E264 and *B. pseudomallei* K96243. Orange CDS refer to (left to right) *cdiA*, *cdiI*, *cdiO* and *cdiB*; grey CDS refer to hypotheticals and purple CDS refer to IS5-family transposases IS1021 (Pe_57) and ISRso18 (Pe_3). The small light purple CDS is a pseudogenized IS (GMI1000). The *cdiA*-CT domain is immediately adjacent to *cdiI*. Vertical blocks indicate regions of shared similarity shaded according to BLASTn similarity.
Figure S2. Loss of host recognition is not a product of differential protein accumulation in planta. RipA1 (A) and RipAA (B) protein accumulation does not explain the loss of programmed cell death phenotype when overexpressed in N. benthamiana leaves. Proteins were overexpressed in N. benthamiana leaves using agroinfiltration (OD₆₀₀ 0.4). Proteins were sampled at 2 day post infiltration and subjected to western blot analysis to visualize their accumulation using anti FLAG and anti HA antibodies. A nonspecific band present in the RipA1 immunoblot is denoted with asterisk. Ponceau S staining provided to confirm equal amounts of protein loaded.
Figure S3. Consistent recognition of T3E alleles in *R. solanacearum* phylotype I isolates. RipAY (A), RipE1 (B), RipA5 (C) and RipH2 (D) were overexpressed in *N. benthamiana* (A,B,C) or *N. tabacum* (D) leaves and PCD symptoms photographed at 4 dpi. Ion leakage was quantified from the infiltrated tissue. Average of 4 technical replicates are plotted on the graph with error bars representing S.E.M.
<table>
<thead>
<tr>
<th>Strain</th>
<th>Year</th>
<th>Location of isolation</th>
<th>Host of isolation</th>
<th>Assembly length (bp)</th>
<th>Scaffolds (^1)</th>
<th>Longest scaffold</th>
<th>NS0</th>
<th>Coverage</th>
<th>Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pe_1(^1)</td>
<td>2000</td>
<td>Chungnam Seosan, South Korea</td>
<td>Capsicum annuum</td>
<td>5,839,983</td>
<td>5</td>
<td>3,681,740</td>
<td>3,681,758</td>
<td>306</td>
<td>WSNR</td>
</tr>
<tr>
<td>Pe_2</td>
<td>2000</td>
<td>Chungbuk Cheongwon, South Korea</td>
<td>Capsicum annuum</td>
<td>5,600,524</td>
<td>213</td>
<td>254,641</td>
<td>85,236</td>
<td>289</td>
<td>WSYW</td>
</tr>
<tr>
<td>Pe_3(^1)</td>
<td>2001</td>
<td>Chungbuk Chungju, South Korea</td>
<td>Capsicum annuum</td>
<td>5,938,020</td>
<td>3</td>
<td>3,851,714</td>
<td>3,851,698</td>
<td>500</td>
<td>WSYS</td>
</tr>
<tr>
<td>Pe_4</td>
<td>2001</td>
<td>Jeju Bukjeju, South Korea</td>
<td>Capsicum annuum</td>
<td>5,593,270</td>
<td>203</td>
<td>317,826</td>
<td>77,130</td>
<td>247</td>
<td>WSYZ</td>
</tr>
<tr>
<td>Pe_9</td>
<td>2002</td>
<td>Chungnam Cheongyang, South Korea</td>
<td>Capsicum annuum</td>
<td>5,723,071</td>
<td>278</td>
<td>233,723</td>
<td>78,017</td>
<td>240</td>
<td>WSYY</td>
</tr>
<tr>
<td>Pe_13</td>
<td>2002</td>
<td>Jeonbuk Imsil, South Korea</td>
<td>Capsicum annuum</td>
<td>5,865,496</td>
<td>248</td>
<td>214,411</td>
<td>77,725</td>
<td>208</td>
<td>WSYX</td>
</tr>
<tr>
<td>Pe_15</td>
<td>2002</td>
<td>Jeonbuk Imsil, South Korea</td>
<td>Capsicum annuum</td>
<td>6,033,637</td>
<td>237</td>
<td>329,719</td>
<td>73,852</td>
<td>291</td>
<td>WSZA</td>
</tr>
<tr>
<td>Pe_18</td>
<td>2002</td>
<td>Jeonnam Haenam, South Korea</td>
<td>Capsicum annuum</td>
<td>6,029,178</td>
<td>254</td>
<td>241,872</td>
<td>83,596</td>
<td>261</td>
<td>JAABKA</td>
</tr>
<tr>
<td>Pe_24</td>
<td>2002</td>
<td>Chungnam Gungju, South Korea</td>
<td>Capsicum annuum</td>
<td>5,917,705</td>
<td>253</td>
<td>421,926</td>
<td>85,211</td>
<td>319</td>
<td>JAABIZ</td>
</tr>
<tr>
<td>Pe_26</td>
<td>2003</td>
<td>Chungnam Seosan, South Korea</td>
<td>Capsicum annuum</td>
<td>5,747,672</td>
<td>200</td>
<td>235,891</td>
<td>77,520</td>
<td>286</td>
<td>JAABJY</td>
</tr>
<tr>
<td>Pe_27(^1)</td>
<td>2003</td>
<td>Chungbuk Cheongwon, South Korea</td>
<td>Capsicum annuum</td>
<td>5,853,871</td>
<td>9</td>
<td>1,965,711</td>
<td>1,655,033</td>
<td>450</td>
<td>WSYT</td>
</tr>
<tr>
<td>Pe_28</td>
<td>1999</td>
<td>Gyeonggi Hwasoeong, South Korea</td>
<td>Capsicum annuum</td>
<td>5,830,258</td>
<td>247</td>
<td>236,369</td>
<td>74,138</td>
<td>243</td>
<td>JAABIX</td>
</tr>
<tr>
<td>Pe_30</td>
<td>2005</td>
<td>Jeonbuk Imsil, South Korea</td>
<td>Capsicum annuum</td>
<td>5,894,388</td>
<td>227</td>
<td>246,051</td>
<td>83,264</td>
<td>271</td>
<td>JAABIW</td>
</tr>
<tr>
<td>Pe_39(^1)</td>
<td>2005</td>
<td>Jeonbuk Jeongeup, South Korea</td>
<td>Capsicum annuum</td>
<td>5,919,856</td>
<td>4</td>
<td>3,808,376</td>
<td>3,808,376</td>
<td>490</td>
<td>WSYU</td>
</tr>
<tr>
<td>Pe_42</td>
<td>2005</td>
<td>Jeonnam Naju, South Korea</td>
<td>Capsicum annuum</td>
<td>5,788,311</td>
<td>272</td>
<td>170,277</td>
<td>74,548</td>
<td>281</td>
<td>JAABJV</td>
</tr>
<tr>
<td>Pe_45</td>
<td>2005</td>
<td>Chungnam Gungju, South Korea</td>
<td>Capsicum annuum</td>
<td>5,702,018</td>
<td>211</td>
<td>317,902</td>
<td>72,734</td>
<td>246</td>
<td>JAABJU</td>
</tr>
<tr>
<td>Pe_49</td>
<td>2005</td>
<td>Chungnam Cheongyang, South Korea</td>
<td>Capsicum annuum</td>
<td>5,645,699</td>
<td>223</td>
<td>241,872</td>
<td>73,220</td>
<td>300</td>
<td>JAABIT</td>
</tr>
<tr>
<td>Pe_51</td>
<td>2005</td>
<td>Chungnam Taean, South Korea</td>
<td>Capsicum annuum</td>
<td>5,722,228</td>
<td>229</td>
<td>406,488</td>
<td>86,970</td>
<td>285</td>
<td>JAABJ</td>
</tr>
<tr>
<td>Pe_52</td>
<td>2005</td>
<td>Chungbuk Goesan, South Korea</td>
<td>Capsicum annuum</td>
<td>5,688,143</td>
<td>240</td>
<td>205,520</td>
<td>78,443</td>
<td>250</td>
<td>JAABJR</td>
</tr>
<tr>
<td>Pe_56</td>
<td>2005</td>
<td>Chungbuk Eumseong, South Korea</td>
<td>Capsicum annuum</td>
<td>5,642,649</td>
<td>225</td>
<td>292,052</td>
<td>86,052</td>
<td>258</td>
<td>JAABQ</td>
</tr>
<tr>
<td>Pe_57(^1)</td>
<td>2005</td>
<td>Gyeongbuk Cheongsong, South Korea</td>
<td>Capsicum annuum</td>
<td>5,720,795</td>
<td>2</td>
<td>3,754,354</td>
<td>3,754,354</td>
<td>490</td>
<td>WSYV</td>
</tr>
<tr>
<td>Pe_61</td>
<td>2005</td>
<td>Gyeongbuk Goesan, South Korea</td>
<td>Capsicum annuum</td>
<td>5,719,158</td>
<td>246</td>
<td>269,792</td>
<td>82,771</td>
<td>247</td>
<td>JAABJP</td>
</tr>
<tr>
<td>To_1</td>
<td>2008</td>
<td>Gangwon Hoenge-seong, South Korea</td>
<td>Solanum lycopersicum</td>
<td>5,878,275</td>
<td>235</td>
<td>216,471</td>
<td>84,441</td>
<td>245</td>
<td>JAABJO</td>
</tr>
<tr>
<td>To_7</td>
<td>2008</td>
<td>Gangwon Hwacheon, South Korea</td>
<td>Solanum lycopersicum</td>
<td>5,666,531</td>
<td>228</td>
<td>269,904</td>
<td>78,595</td>
<td>215</td>
<td>JAABN</td>
</tr>
<tr>
<td>To_22</td>
<td>2008</td>
<td>Gangwon Cheorwon, South Korea</td>
<td>Solanum lycopersicum</td>
<td>5,357,968</td>
<td>223</td>
<td>323,913</td>
<td>62,467</td>
<td>232</td>
<td>JAABIM</td>
</tr>
<tr>
<td>To_28</td>
<td>2008</td>
<td>Gangwon Yang-gu, South Korea</td>
<td>Solanum lycopersicum</td>
<td>5,656,151</td>
<td>235</td>
<td>275,581</td>
<td>69,771</td>
<td>245</td>
<td>JAABIL</td>
</tr>
<tr>
<td>To_36</td>
<td>2008</td>
<td>Gangwon Pyeongchang, South Korea</td>
<td>Solanum lycopersicum</td>
<td>5,785,858</td>
<td>252</td>
<td>213,926</td>
<td>79,781</td>
<td>238</td>
<td>JAABK</td>
</tr>
<tr>
<td>To_42</td>
<td>2008</td>
<td>Gangwon Chuncheon, South Korea</td>
<td>Solanum lycopersicum</td>
<td>5,811,979</td>
<td>306</td>
<td>214,411</td>
<td>61,263</td>
<td>112</td>
<td>JAABJI</td>
</tr>
<tr>
<td>To_53</td>
<td>2008</td>
<td>Gangwon Hongcheon, South Korea</td>
<td>Solanum lycopersicum</td>
<td>5,357,324</td>
<td>231</td>
<td>297,172</td>
<td>54,396</td>
<td>226</td>
<td>JAABJ</td>
</tr>
<tr>
<td>To_63</td>
<td>2008</td>
<td>Gyeongbuk Bonghwa, South Korea</td>
<td>Solanum lycopersicum</td>
<td>5,612,635</td>
<td>208</td>
<td>210,182</td>
<td>82,829</td>
<td>203</td>
<td>JAABJH</td>
</tr>
</tbody>
</table>

\(^1\) Hybrid assemblies

\(^2\) Contigs shorter than 500bp not included
Table 2. Impact of recombination on South Korean phylotype I isolates

<table>
<thead>
<tr>
<th></th>
<th>Alignment length (bp)</th>
<th>Variable sites</th>
<th>Recombinant sites</th>
<th>R/theta②</th>
<th>1/delta</th>
<th>delta③ (bp)</th>
<th>nu④</th>
<th>r/m⑤</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosome</td>
<td>1,129,045</td>
<td>4,260</td>
<td>2,026,835 (64.2%)</td>
<td>0.805</td>
<td>0.0013</td>
<td>758</td>
<td>0.007</td>
<td>4.30</td>
</tr>
<tr>
<td>Megaplasmid</td>
<td>186,600</td>
<td>614</td>
<td>1,511,546 (89.0%)</td>
<td>0.863</td>
<td>0.0010</td>
<td>1025</td>
<td>0.006</td>
<td>5.40</td>
</tr>
</tbody>
</table>

①Gap-free nonrecombinant alignment length
②Rate of recombination/mutation
③Average length of recombinant regions
④Average relative divergence of recombinant sequence
⑤Relative effect of recombination to mutation
<table>
<thead>
<tr>
<th>Strain</th>
<th>Phylotype</th>
<th>Host of isolation</th>
<th>Location of isolation</th>
<th>WGS Project</th>
<th>Genbank Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM1000</td>
<td>I</td>
<td>Solanum lycopersicum</td>
<td>Kourou, French Guyana</td>
<td></td>
<td>GCA_00009125_1</td>
</tr>
<tr>
<td>Rs-09-161</td>
<td>I</td>
<td>Solanum melongena</td>
<td>Goa, India</td>
<td>JHBO01</td>
<td>GCA_00067133_1</td>
</tr>
<tr>
<td>CaRs-Mep</td>
<td>I</td>
<td>Elettaria cardamom</td>
<td>Kerala, India</td>
<td>MCBM01</td>
<td>GCA_00185549_1</td>
</tr>
<tr>
<td>UW757</td>
<td>I</td>
<td>Osteosperum</td>
<td>Escuintla, Guatemala</td>
<td>LFJP01</td>
<td>GCA_00164572_1</td>
</tr>
<tr>
<td>UTT-25</td>
<td>I</td>
<td>Solanum lycopersicum</td>
<td>India</td>
<td>PPFC02</td>
<td>GCA_00293008_2</td>
</tr>
<tr>
<td>Rs-10-244</td>
<td>I</td>
<td>Capsicum annuum</td>
<td>Andaman Islands, India</td>
<td>JHAM01</td>
<td>GCA_00067131_1</td>
</tr>
<tr>
<td>SEPPX05</td>
<td>I</td>
<td>Sesasum indicum</td>
<td>China</td>
<td></td>
<td>GCA_00216201_1</td>
</tr>
<tr>
<td>EP1</td>
<td>I</td>
<td>Solanum melongena</td>
<td>Guangdong, China</td>
<td></td>
<td>GCA_00189110_1</td>
</tr>
<tr>
<td>PSS1308</td>
<td>I</td>
<td>Solanum lycopersicum</td>
<td>Taiwan</td>
<td>MOLO01</td>
<td>GCA_00187080_1</td>
</tr>
<tr>
<td>OE1-1</td>
<td>I</td>
<td>Solanum melongena</td>
<td>Japan</td>
<td></td>
<td>GCA_00187956_1</td>
</tr>
<tr>
<td>RSCM</td>
<td>I</td>
<td>Cucurbita maxima</td>
<td>Guangdong, China</td>
<td></td>
<td>GCA_00289428_1</td>
</tr>
<tr>
<td>PSS216</td>
<td>I</td>
<td>Solanum lycopersicum</td>
<td>Hsinchu, Taiwan</td>
<td>MOL01</td>
<td>GCA_00187697_1</td>
</tr>
<tr>
<td>Rs-T02</td>
<td>I</td>
<td>Solanum lycopersicum</td>
<td>Guangxi, China</td>
<td>LKVH01</td>
<td>GCA_00148409_1</td>
</tr>
<tr>
<td>CQPS-1</td>
<td>I</td>
<td>Nicotiana tabacum</td>
<td>Chongqing, China</td>
<td></td>
<td>GCA_00222046_1</td>
</tr>
<tr>
<td>PSS4</td>
<td>I</td>
<td>Solanum lycopersicum</td>
<td>Taiwan</td>
<td>MOLK01</td>
<td>GCA_00187698_1</td>
</tr>
<tr>
<td>RD15</td>
<td>I</td>
<td>Raphanus sativus</td>
<td>Taiwan</td>
<td>MNCM01</td>
<td>GCA_00185426_1</td>
</tr>
<tr>
<td>Y45</td>
<td>I</td>
<td>Nicotiana tabacum</td>
<td>Yunnan, China</td>
<td>AFWL01</td>
<td>GCA_00223115_1</td>
</tr>
<tr>
<td>P781</td>
<td>I</td>
<td>Mandevilla sp.</td>
<td>Florida, USA</td>
<td>JXLK01</td>
<td>GCA_00164486_1</td>
</tr>
<tr>
<td>FQY_4</td>
<td>I</td>
<td>Solanum sp.</td>
<td>Guizhou, China</td>
<td></td>
<td>GCA_00034854_1</td>
</tr>
<tr>
<td>YC40-M</td>
<td>I</td>
<td>Kaempferia sp.</td>
<td>Shajiang, China</td>
<td></td>
<td>GCA_00166341_1</td>
</tr>
<tr>
<td>KACC10709</td>
<td>I</td>
<td>Solanum lycopersicum</td>
<td>South Korea</td>
<td></td>
<td>GCA_00170852_1</td>
</tr>
<tr>
<td>SD54</td>
<td>I</td>
<td>Zingiber officinale</td>
<td>China</td>
<td>ASQR02</td>
<td>GCA_00043092_1</td>
</tr>
<tr>
<td>FJAT-91</td>
<td>I</td>
<td>Solanum lycopersicum</td>
<td>Fujian, China</td>
<td></td>
<td>GCA_00215524_1</td>
</tr>
<tr>
<td>FJAT-1458</td>
<td>I</td>
<td>Solanum lycopersicum</td>
<td>China</td>
<td></td>
<td>GCA_00188753_1</td>
</tr>
<tr>
<td>PSS190</td>
<td>I</td>
<td>Solanum lycopersicum</td>
<td>Taiwan</td>
<td>MOLP01</td>
<td>GCA_00187082_1</td>
</tr>
<tr>
<td>UW25</td>
<td>IIA</td>
<td>Solanum lycopersicum</td>
<td>USA</td>
<td>NCTK01</td>
<td>GCA_00225169_1</td>
</tr>
<tr>
<td>CIP120</td>
<td>IIA</td>
<td>Solanum tuberosum</td>
<td>Peru</td>
<td>JXAY01</td>
<td>GCA_00164479_1</td>
</tr>
<tr>
<td>RS_489</td>
<td>IIA</td>
<td>Solanum lycopersicum</td>
<td>Brazil</td>
<td></td>
<td>GCA_00254981_1</td>
</tr>
<tr>
<td>PS97</td>
<td>IIA</td>
<td>Solanum lycopersicum</td>
<td>USA</td>
<td>JIBY01</td>
<td>GCA_00164480_1</td>
</tr>
<tr>
<td>IBSBF1900</td>
<td>IIA</td>
<td>Musa sp.</td>
<td>Brazil</td>
<td>CDRW01</td>
<td>GCA_00137327_1</td>
</tr>
<tr>
<td>BS0</td>
<td>IIA</td>
<td>Musa sp.</td>
<td>Brazil</td>
<td>CDMA01</td>
<td>GCA_00082578_1</td>
</tr>
<tr>
<td>Grenada_9-1</td>
<td>IIA</td>
<td>Musa sp.</td>
<td>Grenada</td>
<td>CDLW01</td>
<td>GCA_00082584_1</td>
</tr>
<tr>
<td>UW181</td>
<td>IIA</td>
<td>Musa sp.</td>
<td>Venezuela</td>
<td>CDSB01</td>
<td>GCA_00137331_1</td>
</tr>
<tr>
<td>CFBP7014</td>
<td>IIB</td>
<td>Anthurium andreanum</td>
<td>Trinidad</td>
<td>CDRJ01</td>
<td>GCA_00137325_1</td>
</tr>
<tr>
<td>P673</td>
<td>IIB</td>
<td>Epipremnum aureum</td>
<td>USA</td>
<td>JAL01</td>
<td>GCA_00052561_1</td>
</tr>
<tr>
<td>CFBP6783</td>
<td>IIB</td>
<td>Heliconia sp.</td>
<td>Martinique</td>
<td>JXAZ01</td>
<td>GCA_00164481_1</td>
</tr>
<tr>
<td>IBSBF1503</td>
<td>IIB</td>
<td>Cucumis sativus</td>
<td>Brazil</td>
<td></td>
<td>GCA_00158715_1</td>
</tr>
<tr>
<td>Po82</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>Mexico</td>
<td></td>
<td>GCA_00021646_1</td>
</tr>
</tbody>
</table>
Table S1. Publicly available RSSC assemblies included in this work - continued

<table>
<thead>
<tr>
<th>Strain</th>
<th>Phylotype</th>
<th>Host of isolation</th>
<th>Location of isolation</th>
<th>WGS Project</th>
<th>Genbank Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>UW179</td>
<td>IIB</td>
<td>Musa sp.</td>
<td>Colombia</td>
<td>CDLZ01</td>
<td>GCA_000825805.2</td>
</tr>
<tr>
<td>UW163_2</td>
<td>IIB</td>
<td>Musa sp.</td>
<td>Peru</td>
<td>GCA_001587135.1</td>
<td></td>
</tr>
<tr>
<td>CFBP1416</td>
<td>IIB</td>
<td>Musa sp.</td>
<td>Costa Rica</td>
<td>CDLX01</td>
<td>GCA_000825925.2</td>
</tr>
<tr>
<td>MolK2</td>
<td>IIB</td>
<td>Musa sp.</td>
<td>Philippines</td>
<td>CAHW02</td>
<td>GCA_000212635.2</td>
</tr>
<tr>
<td>CIPA17</td>
<td>IIB</td>
<td>Musa sp.</td>
<td>Philippines</td>
<td>CDM01</td>
<td>GCA_000825825.2</td>
</tr>
<tr>
<td>23-10BR</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>Brazil</td>
<td>JQOI01</td>
<td>GCA_000749995.1</td>
</tr>
<tr>
<td>CFBP3858</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>Netherlands</td>
<td>CDLS01</td>
<td>GCA_001373315.1</td>
</tr>
<tr>
<td>UW491</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>Colombia</td>
<td>LVKU01</td>
<td>GCA_001669645.1</td>
</tr>
<tr>
<td>NCPPB_282</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>Colombia</td>
<td>JQSH01</td>
<td>GCA_000750575.1</td>
</tr>
<tr>
<td>RS2</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>Bolivia</td>
<td>CDRX01</td>
<td>GCA_001373295.1</td>
</tr>
<tr>
<td>IPO1609</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>Netherlands</td>
<td>CDGL01</td>
<td>GCA_001050995.1</td>
</tr>
<tr>
<td>UW551</td>
<td>IIB</td>
<td>Geranium sp.</td>
<td>Kenya</td>
<td>NCTI01</td>
<td>GCA_0002251655.1</td>
</tr>
<tr>
<td>GEO_99</td>
<td>IIB</td>
<td>Capsicum annuum</td>
<td>Georgia</td>
<td>MZNBO1</td>
<td>GCA_002029865.1</td>
</tr>
<tr>
<td>UW24</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>Israel</td>
<td>LVKTO1</td>
<td>GCA_001669635.1</td>
</tr>
<tr>
<td>GEO_230</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>Georgia</td>
<td>PHHL01</td>
<td>GCA_002894795.1</td>
</tr>
<tr>
<td>POPS2</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>China</td>
<td>JQSO01</td>
<td>GCA_000750585.1</td>
</tr>
<tr>
<td>GEO_304</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>Georgia</td>
<td>PHHO01</td>
<td>GCA_002894775.1</td>
</tr>
<tr>
<td>CFIA906</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>India</td>
<td>MZNA01</td>
<td>GCA_002029895.1</td>
</tr>
<tr>
<td>GEO_96</td>
<td>IIB</td>
<td>Solanum lycopersicum</td>
<td>Georgia</td>
<td>MZNA01</td>
<td>GCA_002029895.1</td>
</tr>
<tr>
<td>UW365</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>China</td>
<td>LVKSO1</td>
<td>GCA_001669865.1</td>
</tr>
<tr>
<td>GEO_6</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>Georgia</td>
<td>PHHM01</td>
<td>GCA_002894765.1</td>
</tr>
<tr>
<td>NCPPB_909</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>Egypt</td>
<td>JNGDO1</td>
<td>GCA_000710695.1</td>
</tr>
<tr>
<td>GEO_57</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>Georgia</td>
<td>MXAM01</td>
<td>GCA_002029885.1</td>
</tr>
<tr>
<td>GEO_55</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>Georgia</td>
<td>PHHO01</td>
<td>GCA_002894845.1</td>
</tr>
<tr>
<td>UY031</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>Uruguay</td>
<td>PHHP01</td>
<td>GCA_001299555.1</td>
</tr>
<tr>
<td>GEO_81</td>
<td>IIB</td>
<td>Solanum tuberosum</td>
<td>Georgia</td>
<td>PHHN01</td>
<td>GCA_002894785.1</td>
</tr>
<tr>
<td>KACC_10722</td>
<td>IV</td>
<td>Solanum tuberosum</td>
<td>South Korea</td>
<td>GCA_001586135.1</td>
<td></td>
</tr>
<tr>
<td>PSIO7</td>
<td>IV</td>
<td>Solanum lycopersicum</td>
<td>Indonesia</td>
<td>GCA_000283475.1</td>
<td></td>
</tr>
<tr>
<td>CMR15</td>
<td>III</td>
<td>Solanum lycopersicum</td>
<td>Cameroon</td>
<td>GCA_000427195.1</td>
<td></td>
</tr>
<tr>
<td>CFBP3059</td>
<td>III</td>
<td>Solanum melongena</td>
<td>Burkina Faso</td>
<td>JXBA01</td>
<td>GCA_001644855.1</td>
</tr>
</tbody>
</table>
### Table S2. Genes with 2+ strain-specific recombination events

<table>
<thead>
<tr>
<th>Reference</th>
<th>Gene ID</th>
<th>Highest event #</th>
<th>Locus tag</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>2</td>
<td>GO298_00008</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome pilE1</td>
<td>4</td>
<td>GO298_00322</td>
<td>Fimbrial protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>3</td>
<td>GO298_00323</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>7</td>
<td>GO298_00324</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome pqiA</td>
<td>3</td>
<td>GO298_00365</td>
<td>Paraquat-inducible protein A</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>15</td>
<td>GO298_00626</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>4</td>
<td>GO298_00632</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>3</td>
<td>GO298_00633</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>4</td>
<td>GO298_00634</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome RipG6</td>
<td>5</td>
<td>GO298_01109</td>
<td>RipG6</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome RipS4</td>
<td>3</td>
<td>GO298_01266</td>
<td>RipS4</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome dltA_1</td>
<td>2</td>
<td>GO298_01291</td>
<td>D-alanine--polyphosphoribitol ligase subunit 1</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome fyuA</td>
<td>2</td>
<td>GO298_01295</td>
<td>Pesticin receptor</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome dltA_2</td>
<td>2</td>
<td>GO298_01296</td>
<td>D-alanine--polyphosphoribitol ligase subunit 1</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome shlB_2</td>
<td>2</td>
<td>GO298_01323</td>
<td>Hemolysin transporter protein ShlB</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>4</td>
<td>GO298_01377</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome maeB_1</td>
<td>2</td>
<td>GO298_01811</td>
<td>NADP-dependent malic enzyme</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome RipA1</td>
<td>2</td>
<td>GO298_01826</td>
<td>RipA1</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>3</td>
<td>GO298_02385</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome sasA_5</td>
<td>3</td>
<td>GO298_02386</td>
<td>Adaptive-response sensory-kinase SasA</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>8</td>
<td>GO298_02866</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>5</td>
<td>GO298_02867</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome acor</td>
<td>2</td>
<td>GO298_02881</td>
<td>Acetoin catabolism regulatory protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>2</td>
<td>GO298_02882</td>
<td>Alcohol dehydrogenase</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome ldha</td>
<td>2</td>
<td>GO298_02883</td>
<td>D-lactate dehydrogenase</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome sutR_1</td>
<td>2</td>
<td>GO298_02884</td>
<td>HTH-type transcriptional regulator SutR</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>2</td>
<td>GO298_02885</td>
<td>L-methionine sulfoximine/L-methionine sulfone acetyltransferase</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>2</td>
<td>GO298_02886</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>6</td>
<td>GO298_02888</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>3</td>
<td>GO298_02944</td>
<td>putative parvulin-type peptidyl-prolyl cis-trans isomerase</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome shlB_3</td>
<td>3</td>
<td>GO298_02945</td>
<td>Hemolysin transporter protein ShlB</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>2</td>
<td>GO298_02946</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>3</td>
<td>GO298_02952</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>3</td>
<td>GO298_02953</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>5</td>
<td>GO298_02954</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>6</td>
<td>GO298_02955</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>6</td>
<td>GO298_02991</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>2</td>
<td>GO298_02992</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Gene ID</td>
<td>Highest event #</td>
<td>Locus tag</td>
<td>Product</td>
</tr>
<tr>
<td>-----------</td>
<td>---------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>2</td>
<td>GO298_02993</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>2</td>
<td>GO298_02994</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>4</td>
<td>GO298_02995</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>4</td>
<td>GO298_02996</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>8</td>
<td>GO298_02997</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>8</td>
<td>GO298_02999</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>8</td>
<td>GO298_03000</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>8</td>
<td>GO298_03001</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>8</td>
<td>GO298_03003</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>6</td>
<td>GO298_03004</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>6</td>
<td>GO298_03005</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>6</td>
<td>GO298_03006</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>6</td>
<td>GO298_03007</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>6</td>
<td>GO298_03008</td>
<td>hypothetical protein, CDS overlap</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>8</td>
<td>GO298_03009</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>13</td>
<td>GO298_03010</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>13</td>
<td>GO298_03011</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>6</td>
<td>GO298_03012</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome RipAX1</td>
<td>2</td>
<td>GO298_03055</td>
<td>RipAX1</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>8</td>
<td>GO298_03087</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>5</td>
<td>GO298_03186</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>3</td>
<td>GO298_03353</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome gyrB</td>
<td>2</td>
<td>GO298_03363</td>
<td>DNA gyrase subunit B</td>
<td></td>
</tr>
<tr>
<td>Pe_57 chromosome CDS</td>
<td>3</td>
<td>GO298_03380</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 megaplasmid CDS</td>
<td>5</td>
<td>GO298_03596</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 megaplasmid egl</td>
<td>2</td>
<td>GO298_03600</td>
<td>Endoglucanase</td>
<td></td>
</tr>
<tr>
<td>Pe_57 megaplasmid CDS</td>
<td>2</td>
<td>GO298_03602</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 megaplasmid CDS</td>
<td>3</td>
<td>GO298_03615</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 megaplasmid CDS</td>
<td>2</td>
<td>GO298_03635</td>
<td>Secreted chorismate mutase</td>
<td></td>
</tr>
<tr>
<td>Pe_57 megaplasmid CDS</td>
<td>2</td>
<td>GO298_03649</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 megaplasmid CDS</td>
<td>2</td>
<td>GO298_03830</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 megaplasmid fiI</td>
<td>2</td>
<td>GO298_03831</td>
<td>Flagellum-specific ATP synthase</td>
<td></td>
</tr>
<tr>
<td>Pe_57 megaplasmid fiI</td>
<td>2</td>
<td>GO298_03832</td>
<td>Flagellar FilI protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 megaplasmid CDS</td>
<td>2</td>
<td>GO298_03833</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 megaplasmid CDS</td>
<td>3</td>
<td>GO298_03927</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 megaplasmid CDS</td>
<td>3</td>
<td>GO298_03928</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 megaplasmid CDS</td>
<td>2</td>
<td>GO298_03929</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 megaplasmid CDS</td>
<td>6</td>
<td>GO298_03951</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>Pe_57 megaplasmid shIB_4</td>
<td>3</td>
<td>GO298_03957</td>
<td>Hemolysin transporter protein ShIB</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Gene ID</td>
<td>Highest event #</td>
<td>Locus tag</td>
<td>Product</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------</td>
<td>-----------------</td>
<td>-------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>CDS</td>
<td>2</td>
<td>GO298_03996</td>
<td>hypothetical protein, CDS overlap</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>CDS</td>
<td>2</td>
<td>GO298_03997</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>dltA_4</td>
<td>5</td>
<td>GO298_04044</td>
<td>D-alanine–poly(phosphoribitol) ligase subunit 1</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>tycC_2</td>
<td>8</td>
<td>GO298_04045</td>
<td>Tyrocidine synthase 1</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>CDS</td>
<td>2</td>
<td>GO298_04046</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>CDS</td>
<td>2</td>
<td>GO298_04047</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>CDS</td>
<td>2</td>
<td>GO298_04048</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>acdS</td>
<td>2</td>
<td>GO298_04049</td>
<td>1-aminocyclopropane-1-carboxylate deaminase</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>btuD_6</td>
<td>2</td>
<td>GO298_04128</td>
<td>Vitamin B12 import ATP-binding protein</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>livF_7</td>
<td>2</td>
<td>GO298_04129</td>
<td>High-affinity branched-chain amino acid transport ATP-binding protein</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>CDS</td>
<td>2</td>
<td>GO298_04151</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>CDS</td>
<td>4</td>
<td>GO298_04219</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>RipAN</td>
<td>4</td>
<td>GO298_04262</td>
<td>RipAN</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>RipA3</td>
<td>2</td>
<td>GO298_04263</td>
<td>RipA3</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>RipA4</td>
<td>2</td>
<td>GO298_04264</td>
<td>RipA4</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>yscU</td>
<td>2</td>
<td>GO298_04281</td>
<td>Yop proteins translocation protein U</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>RipAB</td>
<td>2</td>
<td>GO298_04293</td>
<td>RipAB</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>RipX</td>
<td>2</td>
<td>GO298_04294</td>
<td>PopA1/RipX</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>CDS</td>
<td>2</td>
<td>GO298_04295</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>adc</td>
<td>2</td>
<td>GO298_04301</td>
<td>putative acetoacetate decarboxylase</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>bdhA_2</td>
<td>2</td>
<td>GO298_04462</td>
<td>D-beta-hydroxybutyrate dehydrogenase</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>CDS</td>
<td>4</td>
<td>GO298_04470</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>CDS</td>
<td>8</td>
<td>GO298_04472</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>entS</td>
<td>2</td>
<td>GO298_04581</td>
<td>Enterobactin exporter EntS</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>CDS</td>
<td>3</td>
<td>GO298_04583</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>cdIA2_4</td>
<td>3</td>
<td>GO298_04735</td>
<td>tRNA nuclease CdIA-2</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>CDS</td>
<td>2</td>
<td>GO298_04736</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>shlB_6</td>
<td>2</td>
<td>GO298_04737</td>
<td>Hemolysin transporter protein ShlB</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>CDS</td>
<td>2</td>
<td>GO298_04747</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>Pe_57 megaplasmid</td>
<td>CDS</td>
<td>2</td>
<td>GO298_04898</td>
<td>hypothetical protein</td>
</tr>
</tbody>
</table>
Table S3. Strains used as a source of representative effector alleles for testing host recognition

<table>
<thead>
<tr>
<th>RipA1 Group</th>
<th>Pe_57</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>Pe_13</td>
</tr>
<tr>
<td>Group 3</td>
<td>Pe_49</td>
</tr>
<tr>
<td>Group 4</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RipAA Group</th>
<th>Pe_1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>Pe_56</td>
</tr>
<tr>
<td>Group 3</td>
<td>Pe_4</td>
</tr>
<tr>
<td>Group 4</td>
<td>Pe_13</td>
</tr>
<tr>
<td>Group 5</td>
<td>To_22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RipAY Group</th>
<th>Pe_57</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>Pe_13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RipE1 Group</th>
<th>Pe_57</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>Pe_13</td>
</tr>
<tr>
<td>Group 3</td>
<td>Pe_1</td>
</tr>
<tr>
<td>Group 4</td>
<td>To_22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RipA5 Group</th>
<th>Pe_57</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>Pe_27</td>
</tr>
<tr>
<td>Group 3</td>
<td>To_22</td>
</tr>
<tr>
<td>Group 4</td>
<td>Pe_13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RipH2 Group</th>
<th>Pe_57</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>Pe_1</td>
</tr>
<tr>
<td>Group 3</td>
<td>Pe_4</td>
</tr>
<tr>
<td>Group 4</td>
<td>To_22</td>
</tr>
</tbody>
</table>


42. Anderson MS, Garcia EC, Cotter PA. The Burkholderia bcpAIOB Genes Define Unique Classes of Two-Partner Secretion and Contact Dependent Growth Inhibition Systems. Plos Genet 2012;8:e1002877.


