

# Genomic Analysis of the Kiwifruit Pathogen *Pseudomonas syringae* pv. *actinidiae* Provides Insight into the Origins of an Emergent Plant Disease

Honour C. McCann<sup>1,2</sup>, Erik H. A. Rikkerink<sup>3</sup>, Frederic Bertels<sup>1,4</sup>, Mark Fiers<sup>5</sup>, Ashley Lu<sup>5</sup>, Jonathan Rees-George<sup>3</sup>, Mark T. Andersen<sup>3</sup>, Andrew P. Gleave<sup>3</sup>, Bernhard Haubold<sup>6</sup>, Mark W. Wohlers<sup>3</sup>, David S. Guttman<sup>2</sup>, Pauline W. Wang<sup>2</sup>, Christina Straub<sup>1</sup>, Joel Vanneste<sup>7</sup>, Paul B. Rainey<sup>1,6†\*</sup>, Matthew D. Templeton<sup>3,8†\*</sup>

**1** New Zealand Institute for Advanced Study and Allan Wilson Centre, Massey University, Auckland, New Zealand, **2** Centre for the Analysis of Genome Evolution and Function, University of Toronto, Toronto, Canada, **3** The New Zealand Institute for Plant and Food Research Limited, Auckland, New Zealand, **4** Biozentrum, University of Basel and Swiss Institute of Bioinformatics, Basel, Switzerland, **5** The New Zealand Institute for Plant and Food Research Limited, Lincoln, New Zealand, **6** Max Planck Institute for Evolutionary Biology, Plön, Germany, **7** The New Zealand Institute for Plant and Food Research Limited, Ruakura, Hamilton, New Zealand, **8** School of Biological Sciences, University of Auckland, Auckland, New Zealand

## Abstract

The origins of crop diseases are linked to domestication of plants. Most crops were domesticated centuries – even millennia – ago, thus limiting opportunity to understand the concomitant emergence of disease. Kiwifruit (*Actinidia* spp.) is an exception: domestication began in the 1930s with outbreaks of canker disease caused by *P. syringae* pv. *actinidiae* (*Psa*) first recorded in the 1980s. Based on SNP analyses of two circularized and 34 draft genomes, we show that *Psa* is comprised of distinct clades exhibiting negligible within-clade diversity, consistent with disease arising by independent samplings from a source population. Three clades correspond to their geographical source of isolation; a fourth, encompassing the *Psa*-V lineage responsible for the 2008 outbreak, is now globally distributed. *Psa* has an overall clonal population structure, however, genomes carry a marked signature of within-pathovar recombination. SNP analysis of *Psa*-V reveals hundreds of polymorphisms; however, most reside within PPHGI-1-like conjugative elements whose evolution is unlinked to the core genome. Removal of SNPs due to recombination yields an uninformative (star-like) phylogeny consistent with diversification of *Psa*-V from a single clone within the last ten years. Growth assays provide evidence of cultivar specificity, with rapid systemic movement of *Psa*-V in *Actinidia chinensis*. Genomic comparisons show a dynamic genome with evidence of positive selection on type III effectors and other candidate virulence genes. Each clade has highly varied complements of accessory genes encoding effectors and toxins with evidence of gain and loss via multiple genetic routes. Genes with orthologs in vascular pathogens were found exclusively within *Psa*-V. Our analyses capture a pathogen in the early stages of emergence from a predicted source population associated with wild *Actinidia* species. In addition to candidate genes as targets for resistance breeding programs, our findings highlight the importance of the source population as a reservoir of new disease.

**Citation:** McCann HC, Rikkerink EHA, Bertels F, Fiers M, Lu A, et al. (2013) Genomic Analysis of the Kiwifruit Pathogen *Pseudomonas syringae* pv. *actinidiae* Provides Insight into the Origins of an Emergent Plant Disease. PLoS Pathog 9(7): e1003503. doi:10.1371/journal.ppat.1003503

**Editor:** Jeffery L. Dangl, The University of North Carolina at Chapel Hill, United States of America

**Received:** January 20, 2013; **Accepted:** May 28, 2013; **Published:** July 25, 2013

**Copyright:** © 2013 McCann et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was funded in part by the New Zealand Ministry for Business, Innovation and Employment ([www.mbie.govt.nz](http://www.mbie.govt.nz)) contract C06X0812 (to MDT and EHAR), and Allan Wilson Centre for Molecular Ecology and Evolution (<http://www.allanwilsoncentre.ac.nz/>) (to PBR). PBR is grateful for support from Zespri International, Mt Maunganui, New Zealand (<http://www.Zespri.com>). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: rainey@evolbio.mpg.de (PBR); matt.templeton@plantandfood.co.nz (MDT)

† PBR and MDT are joint senior authors on this work.

## Introduction

Despite considerable improvements in the management of plant diseases, modern agriculture remains vulnerable to losses caused by microbial pathogens. Plant diseases conservatively account for the loss of at least 10% of annual global food production [1]. The intensive cultivation of clonally propagated plants with low genetic diversity heightens opportunities for the emergence and rapid spread of infectious disease [1–3].

The origins of agricultural plant diseases are unclear [3,4]. The earliest pathogens are likely to have evolved from commensals or pathogens colonizing wild relatives of plants selected for human domestication [5–10]. Given that domestication of staple crop plants took place centuries (and often millennia) ago, the signal of this evolutionary past – the nature of the initial pathogen population, its relationship with commensal types, its diversity and genetic structure, plus factors and processes that might have led to the first outbreaks of disease – is obscured by the passage of

## Author Summary

Despite considerable scientific advances in plant protection during the last century, agricultural crops remain vulnerable to infection by pathogens. The intensive cultivation particularly of clonally propagated crop plants increases the potential for the emergence and rapid spread of new diseases. *Pseudomonas syringae* pv. *actinidiae* was first reported as a canker-causing pathogen of kiwifruit in the mid-1980s. However, a new outbreak of the disease occurred in 2008 and this strain has spread rapidly throughout growing regions of the world. In order to determine the origin, population structure and defining features of this pathogen, a large-scale sequencing project was established. This clarified the phylogenetic relationships between the different *Psa* isolates and identified the outbreak-specific gene sets associated with the aggressive systemic infection strategy exhibited by the virulent strain. This information is invaluable in developing robust long-term solutions for this serious disease. Given that kiwifruit production on a commercial scale is a relatively recent event, this analysis provides a unique insight into the evolution of this pathogen with its host, from its first emergence to the latest global outbreak. This understanding should aid in the mitigation of devastating outbreaks in the future.

time. Kiwifruit (*Actinidia* spp.) is an exception. Domestication of kiwifruit is recent and clearly documented; outbreaks of disease are recorded and the pathogens responsible have been preserved.

The genus *Actinidia* comprises 55 species and about 76 taxa native to eastern Asia, with the greatest abundance and diversity in the southwestern provinces of China [11]. Fruit have long been collected from the wild, yet commercial cultivation only began to gain momentum in the 1980s, based on the success of *A. deliciosa* ‘Hayward’ developed in New Zealand during the 1930s. Subsequently cultivars from the species *A. chinensis*, such as ‘Hort16A’, ‘Jin Tao’ and ‘Hongyang’ were commercialized in several kiwifruit growing regions, including New Zealand, China (1998), Italy (2001) and Chile (2003) [12].

A small group of fungal and bacterial diseases on vines, roots and fruit of ‘Hayward’ were recognized once commercial plantings became substantial in the 1980s [13]. Kiwifruit canker disease caused by *Pseudomonas syringae* pv. *actinidiae* (*Psa*) was first reported and characterized on *A. deliciosa* in Shizuoka, Japan in 1984 [14]; in that same year bacterial canker disease was also reported in an orchard in Hunan, China [15]. Canker disease was subsequently observed in Korea (1988) and Italy (1992) [16,17]. Symptoms of infection include late winter die-back of young canes, frequently accompanied by rust red exudates from canes and trunks, and the presence of necrotic lesions with chlorotic halos on leaves during the spring [18,19].

In 2008 an aggressive form of *Psa* was reported in Italy on *A. chinensis*. Multi-locus sequencing showed the pathogenic strains to be divergent from earlier Italian and Japanese isolates [20–23]). The virulent form of *Psa* was subsequently detected in neighboring European countries [24], China, Chile and New Zealand [23,25]. Additional genomic analyses confirmed the clonal nature of the disease outbreak and its distinctive genetic composition [19,22,26,27]. Throughout the paper we refer to this recent epidemic as the “2008 outbreak” with the strains responsible being referred to collectively as “*Psa-V*”.

Rapid transmission and increased severity of infection arising from the 2008 outbreak has had devastating effects leading to complete destruction of orchards. In 2010, *Psa-V* was detected in

New Zealand, where kiwifruit is the most valuable horticultural export [12,25]. *Psa-V* spread rapidly from its initial incursion site in the Bay of Plenty. Within two years the number of infected orchards rose from three to 1232 (37% of New Zealand orchards) and continues to increase [28].

Draft genome sequencing of one Italian isolate from the 2008 outbreak, plus another strain from an epidemic in 1992, showed the earlier Italian strain to be identical to a 1984 Japanese isolate, but different to the 2008 (*Psa-V*) outbreak strains, particularly with respect to effector inventories [26]. Draft genomes of four additional isolates from the 2008 Italian outbreak plus three isolates from China and two from Chile also belong to the same *Psa-V* lineage although one Chinese strain (M228) appears divergent [22,27]. Strains of the *Psa-V* lineage were shown to differ by as few as six SNPs. A genomic island with similarity to PPHGI from *Pseudomonas syringae* pv. *phaseolicola* (*Pph*) was characterized and shown to differ between the European, Chilean and Chinese/New Zealand isolates [22,27,29]. The full extent of the genetic distinctiveness of *Psa-V* relative to strains from early outbreaks remains uncertain.

While work to date shows that the 2008 epidemic is caused by a strain distinct from previous outbreaks of canker disease, it is not clear whether *Psa-V* evolved from earlier outbreaks or whether it has independent origins. Clarification can come from studies that determine patterns of nucleotide diversity from strains sampled from different disease outbreaks, at different time points, from different geographical locations [30–33]. From the analysis of polymorphism data it becomes possible to infer underlying population processes and, provided recombination is not extensive, phylogeny [34,35]. Leaving aside opportunities for insight into the evolution of a newly emergent pathogen, knowledge of population structure also has important implications for the development of strategies for disease control, including protocols to prevent future outbreaks.

Here we report an in depth analysis of *Psa* evolution based upon complete genome sequences of a *Psa-V* strain from the New Zealand outbreak and the type strain (J-35, ICMP9617) originally isolated in Japan, plus 34 additional draft genomes that encompass strains from previously known outbreaks. Our analyses provide evidence of a single source population from which outbreaks of disease have arisen *via* independent transmission events to different kiwifruit growing regions. Overall the population is clonal, however, approximately 10% of the genome shows evidence of homologous recombination marked by gene conversion. Comparative analyses reveal dynamic genomes with positive selection affecting type III effectors and candidate virulence genes with *Psa-V* containing numerous genes found in pathovars of vascular plants. *In planta* growth data provide a link between genomic inferences and ecological performance on host cultivars. Together these data provide understanding of the processes and factors affecting the emergence of a new plant pathogen.

## Materials and Methods

### Isolates

The 25 strains of *P. syringae* pv. *actinidiae* were isolated from *Actinidia* spp. in Japan, Korea, Italy and New Zealand. Details of dates and locations of isolations are in Table 1. Bacteria were maintained on King’s B (KB) agar plates [36] and stored at  $-80^{\circ}\text{C}$  in glycerol. Where possible, cultures have been deposited in the International Collection of Micro-organisms from Plants (ICMP; www.landcareresearch.co.nz). Several isolates were sourced from overseas collections, including the Korean Agricultural Culture Collection (KACC), Suwon, Republic of Korea, the National

**Table 1.** Strain table and assembly statistics.

Isolate ID	WGS origin	Host plant	Country	Year	ICMP number	Other collection/alias	Genbank accession	Contigs <sup>2</sup>	N50	Longest scaffold
<i>Psa</i> C-1	Mazzaglia <i>et al.</i> (2012)	<i>A. chinensis</i> 'Hongyang'	China, Shaanxi	2010		CH2010-6	AGUH	342	51,971	198,200
<i>Psa</i> C-9	Butler <i>et al.</i> 2013		China, Shaanxi	2010		M228	ANJ1	3259	4561	57,016
<i>Psa</i> Cl-4	Butler <i>et al.</i> 2013	<i>A. deliciosa</i>	Chile, Maule	2010	19439	Psa1B	ANJM	477	36,822	142,020
<i>Psa</i> Cl-5	Butler <i>et al.</i> 2013	<i>A. deliciosa</i>	Chile, Maule	2010	19455	286532	ANJK	415	39,400	127,854
<i>Psa</i> I-1	Marceletti <i>et al.</i> (2011)	<i>A. deliciosa</i> 'Hayward'	Italy	1992		NCPFB 3871/I-Psa	AFTF	466	27,730	122,209
<i>Psa</i> I-10	Butler <i>et al.</i> 2013	<i>A. deliciosa</i>	Italy, Rome	2010	18744	CRA-FRU 11.41	ANGD	422	35,308	143,214
<i>Psa</i> I-12	This study	<i>A. chinensis</i> 'Hort16A'	Italy, Latina	2010	19079	I.27.4.10	AOKL	3,045	4,064	57,811
<i>Psa</i> I-2	Marceletti <i>et al.</i> (2011)	<i>A. chinensis</i> 'Hort16A'	Italy	2008		CRAFRU8.43/I2-Psa	AFTG	590	22,372	85,982
<i>Psa</i> I-3	Mazzaglia <i>et al.</i> (2012)	<i>A. chinensis</i> 'Hort16A'	Italy, Lazio	2008		CFBP 7286	AGNO	352	43,501	139,438
<i>Psa</i> J-1	Baltrus <i>et al.</i> (2011)	<i>A. deliciosa</i>	Japan	1984		MAFF 302091	AEAL	138	69,188	N/A
<i>Psa</i> J-2	Mazzaglia <i>et al.</i> (2012)	<i>A. chinensis</i>	Japan	1988		PA459	AGNQ	393	49,861	303,211
<i>Psa</i> J-25	Mazzaglia <i>et al.</i> (2012)	<i>A. deliciosa</i> 'Hayward'	Japan, Shizuoka	1984	9855	KW41	AGNP	429	38,609	136,008
<i>Psa</i> J-29	This study	<i>A. arguta</i>	Japan, Kanagawa	1987	19102	MAFF 302133/JpSar1	AOKA	548	27,853	113,869
<i>Psa</i> J-30	This study	<i>A. arguta</i>	Japan, Kanagawa	1987	19103	MAFF 302134/JpSar2	AOJQ	527	30,580	145,811
<i>Psa</i> J-31	This study	<i>A. deliciosa</i> 'Hayward'	Japan, Kanagawa	1987	19070	MAFF 302143/JpKiwa4	AOJY	2,844	3,960	27,127
<i>Psa</i> J-32	This study	<i>A. deliciosa</i> 'Hayward'	Japan, Wakayama	1988	19068	MAFF 302145/JpWa1	AOJX	790	35,303	133,336
<i>Psa</i> J-33	This study	<i>A. deliciosa</i> 'Hayward'	Japan, Wakayama	1988	19104	MAFF 302146/JpWa2	AOJZ	496	34,688	133,960
<i>Psa</i> J-35	This study	<i>A. deliciosa</i> 'Hayward'	Japan, Shizuoka	1984	9617	NCPFB 3739/Kw11	AOKP <sup>3</sup>	2	158,837	6,466,847
<i>Psa</i> K-26	This study	<i>A. chinensis</i>	Korea, Jeonnam	1997	19072	KACC 10584/Kr26	AOJW	350	43,039	169,668
<i>Psa</i> K-27	This study	<i>A. chinensis</i>	Korea, Jeonnam	1998	19073	KACC 10594/Kr27	AOJR	926	17,194	105,639
<i>Psa</i> K-28	This study	<i>A. chinensis</i>	Korea, Jeonnam	1997	19071	KACC 10754/Kr28	AOJS	459	29,527	107,707
<i>Psa</i> NZ LV-11	This study	<i>A. chinensis</i> 'Hort16A'	New Zealand, Bay of Plenty	2010	18804		AOJU	4,811	1,704	34,764
<i>Psa</i> NZ LV-14	This study	<i>A. deliciosa</i> 'Hayward'	New Zealand, Tauranga	2010	18807		AOKG	3,227	3,876	38,937
<i>Psa</i> NZ LV-17	This study	<i>A. chinensis</i> 'Hort16A'	New Zealand, Katikati	2010	19096		AOKF	4,832	2,333	19,396
<i>Psa</i> NZ LV-18	This study	<i>A. chinensis</i> 'Hort16A'	New Zealand, Motueka	2010	19098		AOKE	391	49,576	136,769
<i>Psa</i> NZ LV-19	This study	<i>A. chinensis</i> 'Hort16A'	New Zealand, Te Puke	2010	19099		AOKD	417	43,791	164,509
<i>Psa</i> NZ LV-20	This study	<i>A. chinensis</i> 'Hort16A'	New Zealand, Te Puke	2010	19100		AOKC	631	26,586	113,072
<i>Psa</i> NZ LV-5	This study	<i>A. chinensis</i> 'Hort16A'	New Zealand, Hawke's Bay	2010	18803		AOKK	351	55,333	144,243
<i>Psa</i> NZ LV-6	This study	<i>A. chinensis</i> 'Hort16A'	New Zealand, Te Puke	2010	19094		AOJU	586	26,883	111,577
<i>Psa</i> NZ LV-8	This study	<i>A. chinensis</i> 'Hort16A'	New Zealand, Te Puke	2010	19095		AOKI	849	15,642	70,310
<i>Psa</i> NZ LV-9	This study	<i>A. deliciosa</i> 'Hayward'	New Zealand, Golden Bay	2010	18883		AOKH	502	34,879	147,278
<i>Psa</i> NZ V-1	This study	<i>A. deliciosa</i> 'Hayward'	New Zealand, Te Puke	2010	18886		AOJT	2,016	6,501	53,135
<i>Psa</i> NZ V-13	This study	<i>A. deliciosa</i> 'Hayward'	New Zealand, Te Puke	2010	18884		AOKO	2	163,512	6,504,601
<i>Psa</i> NZ V-15	This study	<i>A. deliciosa</i> 'Hayward'	New Zealand, Te Puke	2010	19101		AOKM	2,171	6,108	73,690
<i>Psa</i> NZ V-16	This study	<i>A. chinensis</i> 'Hort16A'	New Zealand, Te Puke	2010	18801		AOKQ	800	23,499	100,404

Table 1. Cont.

Isolate ID	WGS origin	Host plant	Country	Year	ICMP number	Other collection/alias	Genbank accession	Contigs <sup>2</sup>	N50	Longest scaffold
<i>Psa</i> NZ V-23	This study	<i>A. chinensis</i> 'Hort16A'	New Zealand, Te Puke	2010	19097	10787	AOKN	843	20,872	80,234
<i>Pmp</i> M302280	Baltrus et al. (2011)	<i>Prunus</i> spp.	England	1931		MAFF 302280	AEAE	969	15161	N/A
<i>Pmp</i>	N/A <sup>1</sup>	<i>Prunus avium</i>	New Zealand, Alexandra	1972	3676			N/A	N/A	N/A
<i>Pth</i> 2598	Mazzaglia et al. (2012)	<i>Camellia sinensis</i>	Japan	1970	3923	NCPBP 2598	AGNN	218	7937	242,733
<i>Pro</i> DC3000	Buell et al. (2003)	<i>Solanum lycopersicon</i>	England	1960	2844	NCPBP 1106	AE016853	2	N/A	6,397,126

<sup>1</sup>A New Zealand *Pmp* isolate was used solely for the pathogenicity assay.

<sup>2</sup>Italicized contig entries refer to the number of scaffolds in the assembly. Various assembly programs were used by different authors so these data are not directly comparable.

<sup>3</sup>A prior version of NCPBP3739 was deposited in NCBI with the Genbank accession AFTH.

doi:10.1371/journal.ppat.1003503.t001

Institute of Agrobiological Sciences (NIAS), 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan, the National Collection of Plant Pathogenic Bacteria (NCPBP), Food and Environment Research Agency, Sand Hutton, York, UK.

### DNA extraction

Cultures were grown overnight in KB or nutrient broth. DNA for whole genome sequencing was purified using the PureGene DNA Isolation kit (Qiagen, Hilden, Germany) with some modifications. For each 750  $\mu$ L of culture extracted, reagent volumes were doubled and the protein precipitation step was carried out twice. Purified DNA was quantified and its purity assessed using a nanodrop spectrophotometer (NanoDrop Technologies, Rockland, DE).

### Genome sequencing and assembly

Paired-end libraries were generated from genomic DNA and sequenced using the Illumina GAI instrument at the Center for Genome Analysis and Function (CAGEF, Toronto). The raw data were filtered using fastq-mcf (<https://code.google.com/p/ea-utils/wiki/FastqMcf>) and quality checked using fastqc (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Contigs were built using the SOAP *de novo* assembler (<http://soap.genomics.org.cn/soapdenovo.html>) with a k-mer value of 37. Sequencing and assembly statistics for each strain are shown in Table 1. Additional sequencing was performed on two *Psa* isolates, NZ V-13 (a virulent (*Psa*-V) strain isolated during the 2010 outbreak in New Zealand), and the *Psa* type strain (J-35, ICMP9617 originally isolated in 1984 from *Actinidia* in Japan) using a 10 kb mate-end library on the Roche 454 platform by Macrogen, Korea ([www.macrogen.com](http://www.macrogen.com)). Scaffolds were generated using the Roche Newbler assembler, and resulted in 14 and 12 scaffolds greater than 2 kb, for NZ V-13 and J-35, respectively. The quality of all assemblies was improved by iteratively filling in short tracts of Ns using GapCloser (version 1.12, <http://soap.genomics.org.cn/soapdenovo.html>) and scaffolding with the mate-pair data using SSPACE [37]. The two Newbler assemblies were used as a basis for generating high quality reference sequences by manually integrating the Illumina-based contigs using Geneious (version 5.6.3, Biomatters, <http://www.geneious.com/>). The remaining large scaffolds were linked by designing primers 1–2 kb from the 5' and 3' ends and using long-template PCR (Takara Bio inc., Shiga, Japan) with amplification conditions recommended by the manufacturer. The quality of the genome assemblies was visually inspected using Hagfish (<https://github.com/mfiers/hagfish/>) to align paired-end reads back to the draft genome. Misaligned and dubious contigs were manually filtered. Additional gap-filling was performed on *Psa* NZ V-13 with primers flanking each gap. Products were purified and Sanger sequenced (Macrogen, Korea, [www.macrogen.com](http://www.macrogen.com)). Assemblies for all strains were submitted to the Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP, <https://www.ncbi.nlm.nih.gov/genomes/static/Pipeline.html>) for gene prediction and annotation: they are available from GenBank (Table 1).

### Plasmid preparation and Pulsed-Field Gel Electrophoresis analysis

Plasmid DNA was extracted from 250 mL cultures using an alkaline lysis with SDS extraction method (Protocol 3: Preparation of Plasmid DNA by Alkaline Lysis with SDS: Maxipreparation) [38]. Nutrient Broth (250 mL) was inoculated with a 2 mL overnight culture and left shaking for 48 h at 28°C. Cells were harvested by centrifugation at 12,000 *g* and either used fresh or

stored at  $-20^{\circ}\text{C}$  until required. Cells were resuspended in 18 mL Alkaline Lysis Solution I (50 mM Glucose, 25 mM Tris (pH 8.0), 10 mM EDTA (pH 8.0)), and then 2 mL freshly prepared Lysozyme solution (10 mg/mL) was added. Two volumes (40 mL) of freshly prepared Alkaline Lysis Solution II (0.2 M NaOH, 1% SDS) were added, the cells mixed by gentle inversion several times, and then incubated at room temperature (RT). After 5–10 min of incubation, 20 mL ice-cold Alkaline Lysis Solution III (5 M KAcetate 60 mL, glacial acetic acid, 11.5 mL  $\text{H}_2\text{O}$ , 28.5 mL) was added, and contents mixed by gentle but effective swirling. Samples were incubated on ice for 10 min and then centrifuged at 20,000  $g$  for 30 min at  $5^{\circ}\text{C}$ . The supernatant was separated from the pellet, the volume measured, 0.6 vol of isopropanol was added and samples were incubated for 10 min at RT. After centrifugation at 12,000  $g$ , the supernatant was discarded, and the pellet washed in 70% ethanol, drained and left to dry (without vacuum) for 5–10 min at RT. The pellet was resuspended in 500  $\mu\text{L}$  Tris/EDTA (pH 8.0). Concentration of samples was achieved by further ethanol precipitation.

Plasmid digests were run using a Chef-DR III system (Bio-Rad, Hercules CA). Volumes of up to 20  $\mu\text{L}$  were loaded onto 1.0% 1 $\times$ TAE agarose gels and run in 1 $\times$ TAE buffer for 18 h at  $15^{\circ}\text{C}$  at 6 V/cm. The switch time was 5–15 s with an angle of  $120^{\circ}$ . Gels were stained with ethidium bromide according to the protocol of Sambrook and Russell (2001).

### SNP identification and phylogenetic analysis

Single nucleotide polymorphisms (SNPs) were identified using REALPHY (F Bertels, P. B. Rainey, O. K. Silander and E. van Nimwegen, unpublished) with reference to both *Pto* DC3000 and *Psa* V-13 for phylogenetic analyses between all *Psa* strains and within the *Psa* V clade, respectively. A minimum read coverage of 10, a minimum PHRED score of 20 and a proportion of unequal nucleotide sites of less than 5% was used. Alignments of SNPs and all conserved invariant sites were built using bowtie2 (<http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>) allowing one nucleotide mismatch in each seed alignment (bowtie2 parameter -N). Previously assembled contigs were employed for genomes deposited in Genbank by other groups. Phylogenetic relationships were inferred using RAxML with the GTRGAMMA model (general time reversible substitution model, gamma distributed rate variation) and SplitsTree with 100 bootstrap replicates [39]. Evidence of recombination was obtained initially from SNP alignments against the NZ V-13 genome, followed by Split Decomposition analysis and the Phi test as implemented in SplitsTree. Subsequent analysis used statistical approaches described by Sawyer (1989) [40], and implemented in GENECONV (<http://www.math.wustl.edu/~sawyer/geneconv/>).

### Calculation of the core and flexible genomes

The core genome of the virulent *Psa* genomes was identified using OrthoMCL, which uses a likelihood algorithm to place genes in ortholog groups [41]. The core genome was defined as those genes present amongst all isolates within the set: species-specific gene family expansion was excluded so that downstream analysis would not be affected by the inclusion of in-paralogs. The size of the clade specific genes refers to any ortholog cluster exclusively present in that clade and no other, however not all members of the clade need have a representative ortholog. Genes private to particular strains or sets of strains were identified using either ALFY [42], or a BLAST approach in which, for a given query and a set of subject genomes, we looked for regions unique to the query by extracting the regions that had no ortholog in the subject genomes. The homology search was carried out using translated

BLAST with a maximum E-value of  $10^{-10}$ . Only regions of 1 kb or more were considered. This procedure was implemented in the AWK-script “findUniqueBlast.awk”, which is available in Protocol S1.

### Alignment of core orthologs

Alignments were created for each core ortholog group using PRANK (<http://code.google.com/p/prank-msa/>) alignment software [43,44]. PRANK improves on classical global alignment methods by applying a phylogeny-aware approach that distinguishes between gaps created by insertions or deletions. Independent insertions are prevented from being matched during the progressive alignment process even when they occur at the same position, and the resulting gaps created by these insertions are not penalized during subsequent alignments [45]. Codon-based alignment was employed, with the guide tree inferred by PRANK from the ortholog sequences and default anchoring of pairwise alignments using Exonerate to speed up the alignment process [46].

### Selection analysis of core genomes

PRANK alignments and trees of core orthologous groups were used as input for selection analyses using Codeml in PAML 4.0. Codeml runs a likelihood calculation for multiple sequences on a phylogeny to estimate the ratio of non-synonymous to synonymous mutations ( $\omega$ ) at every codon site. The  $\omega$  ratio is a measure of the direction and magnitude of selection on amino acid changes:  $\omega < 1$  or  $\omega = 1$  indicate purifying or negative and neutral selection, respectively, and  $\omega > 1$  indicates positive selection [43]. The variation in  $\omega$  among sites is modeled by allowing codons to fall into site classes with  $\omega$  values ranging from 0 to 1 (null model 7) and comparing the results with a more general model allowing codons to have  $\omega > 1$ . The likelihood ratio test (LRT) statistic was used to identify the most likely model of sequence evolution at a 1% significance level. Bayes Empirical Bayes estimates of model 8 were then used to identify positive selection acting at individual sites [47,48]. Subcellular localization was predicted using PSORTb 3.0 for core proteins exhibiting statistically significant signatures of positive selection [49].

### Effector distribution and evolution in *Psa* isolates

The set of type 3 secreted effectors (T3SEs) present in each genome was identified by tBLASTx (E value  $< 1e^{-5}$ , 20% minimum identity) using T3SE sequence queries obtained from the T3SE public database (<http://pseudomonas-syringae.org/>). The subject database included all virulent and low virulent *Psa* genomes, as well as *P. syringae* pv. *theae* (*Pth*). Partial hits or T3SEs disrupted by contig breaks are recorded as truncated/disrupted effectors. Name assignment of T3SE orthologs was performed using phylogenetic analyses [50]. tBLASTx was subsequently also used to record significant sequence variation between strains including insertion, deletion, frame shifts and/or translocation events.

### Identification of genomic islands in *Psa* NZ V-13

IslandPath and IslandPick, implemented by IslandViewer [51], were used as guides for the identification of genomic islands (GI) in NZ V-13. Six genomes were selected on the basis of parameter cutoffs intended to produce a robust dataset of highly probable GIs: *P. syringae* pv. *syringae* (*P<sub>sy</sub>*) B728A (NC\_007005.1) [52]; *P. fluorescens* F113 (NC\_016830.1) [53]; *P. fluorescens* SBW25 (NC\_012660.1) [54]; *P. fluorescens* Pf0-1 (NC\_007492.2) [54]; *P. protegens* Pf-5 (NC\_004129.6) [55]; *P. entomophila* L48

(NC\_008027.1) [56]; *P. putida* BIRD-1 (NC\_017530.1) [57]. Predicted GIs were subject to manual curation and delineation of the probable GI size using the following criteria: presence of mobile elements such as transposases and integrases; over-representation of virulence-related genes, genes annotated as hypothetical proteins, and/or outbreak clade-specific genes; and the presence of adjacent tRNA genes, which are common phage integration sites [51].

### Growth assay of *Psa* isolates

*A. deliciosa* ‘Hayward’ and *A. chinensis* ‘Hort16A’ plantlets were inoculated with a representative strain from each clade of *Psa* or a stonefruit canker causing isolate of *P. syringae* pv. *morsprunorum* (*Pmp*) to compare bacterial growth and spread in infected tissues. Each strain was cultured overnight on solid KB medium. The next morning, an inoculation solution was prepared by diluting cells in 10 mM MgSO<sub>4</sub> to a concentration of approximately 10<sup>5</sup> colony forming units (cfu)/mL. Three-month-old clonally propagated plantlets were inoculated by dipping a needle into the inoculation solution and stabbing the stem, followed by the placement of a 2 µL drop of inoculum on top of the wound site (Figure S1A). Four separate plantlets were stab inoculated for each treatment. Plants were maintained at 17–20°C with a light/dark photoperiod of 14/10 hours. Humidity was constant at 70%. Bacterial growth was monitored subsequent to inoculation on day 0 and on days 4, 8 and 14 in both the site of inoculation and the first leaf above the inoculation site. Stem segments and 1 cm<sup>2</sup> leaf discs from the base, middle, tip and margin of the leaf (Figure S1B) were surface sterilized in 70% ethanol and individually ground in 200 µL 10 mM MgSO<sub>4</sub>. Serial dilutions of the homogenate were plated on solid KB medium and incubated for two days at 28°C. The cfu calculated was used to determine the cfu/cm<sup>2</sup> in the stem and leaf tissue.

Statistical analysis involved fitting Analysis of Variance (ANOVA) models using Genstat 14.1 on log<sub>10</sub> transformed colony count values [58]. This allowed for the model assumption of normally distributed residuals with stable variance to be met. As there were a number of zero observations present in the data, a value of 1 was added to the density data before the log<sub>10</sub> transformation was applied. For data collected from stems, F-tests were used to assess if the isolate, time, and their interaction effects were statistically significant. Post hoc tests to compare fitted means were also calculated, with Fisher’s protected Least Significant Difference (LSD) method used to account for multiple testing. The analyses were carried out separately for each of the two cultivars. The leaf data were analyzed similarly, but also included an area factor that indicated the section of the leaf from which the sample came. The ANOVA tested the area, treatment, and time effects along with all two-way and three-way interactions. In addition, plant ID was treated as a blocking factor because there were multiple measurements on the same plant. All tests were carried out at the 5% level of significance. The back-transformed means are presented in the text with the respective Standard Error Ratio (S.E.R), which is also the back-transformed standard error.

## Results and Discussion

### Complete and draft genome sequencing

High-quality draft sequences were assembled for 23 strains of *Psa* from Illumina data (Table 1). This included five strains from Japan isolated between 1984 and 1988, three from Korea (1997–98), one from Italy (2010) and 14 from NZ (all from 2010). The NZ strains were isolated from plants exhibiting either canker disease or foliar symptoms alone and verified as *Psa* based on

diagnostic primers [59]. Strains causing canker disease were consistent with *Psa*-V, whereas the low virulent (LV) form produces mild leaf spotting, but no cankers [28]. In subsequent analyses these 23 draft genomes were supplemented by eleven previously sequenced isolates from China, Italy, Chile and Japan (Table 1).

In addition, the genomes of NZ V-13 and the type strain (J-35) were sequenced with Roche 454 resulting in circularization of each chromosome into a single megascaffold. At approximately 6.5 Mb, both chromosomes are larger than other *P. syringae* pathovars that have been fully sequenced (Table S1). Each strain possesses an autonomously replicating plasmid: NZ V-13 has a 70.1 kb plasmid and the Japanese pathotype strain has one of 32 kb. Restriction digests of plasmid preparations resolved using pulse field gel-electrophoresis (PFGE) matched the predicted pattern and confirm the size of the plasmids (Figure S2). These plasmids are half the size of those reported previously [26].

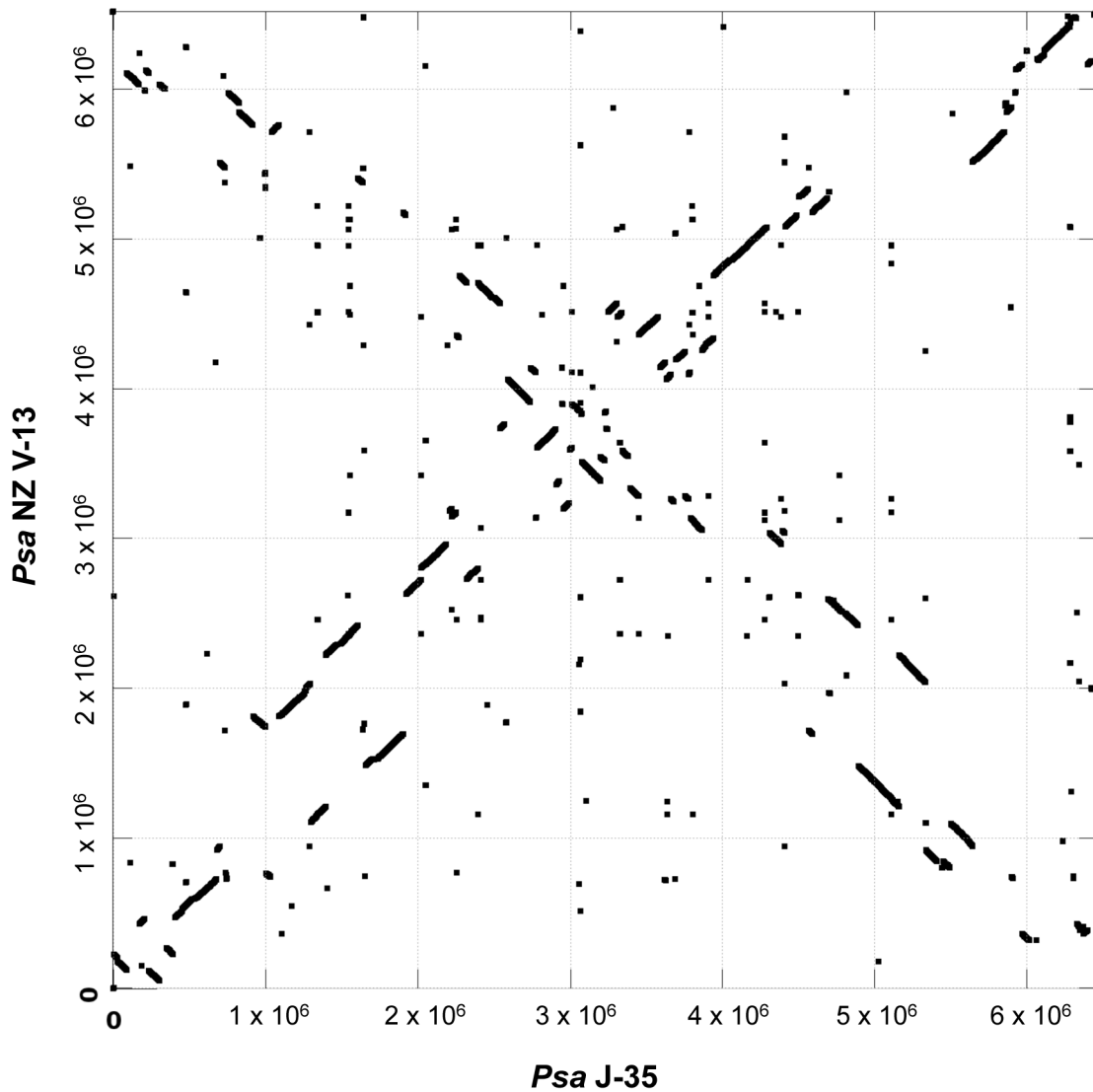
A synteny plot of the two closed genomes of NZ V-13 and J-35 shows evidence of substantial genomic rearrangement (Figure 1). This is unexpected for two strains from a single pathovar. Of particular note is the “X” like structure which indicates multiple translocations and inversions centered about the origin.

### Phylogeny and biogeography

Initial population analyses were based on single nucleotide polymorphisms (SNPs) obtained by aligning DNA sequence reads to the *Pseudomonas syringae* pv. *tomato* (*Pto*) DC3000 genome (NC\_004632.1) [60]. This resulted in 15,329 SNPs and 463,396 invariant sites. Overall diversity as determined by Watterson’s theta was 0.008 (per site) [61]. Phylogenetic analyses identified four distinct clades, of which three exhibit strict phylogeographic structure (Figure 2A). The low-virulent (LV) strains responsible for mild symptoms of foliar infection form a single cluster. Leaving aside the divergent NZ LV-14 isolate, the remaining LV strains differ by just 20 SNPs. The three Korean strains, which were isolated in 1997 and 1998, form a clade defined by 147 SNPs. The Japanese strains isolated between 1984 and 1988 also form a monophyletic group (52 SNPs distinguish these strains) and include the *Psa* I-1 strain, which was isolated in 1992 during an early incursion of *Psa* into Italy [17]. The phylogeographic signal displayed by earlier canker-causing and foliar *Psa* is absent from the group of strains isolated from the 2008 outbreak of *Psa* (*Psa*-V): these form a single clade comprising isolates from New Zealand, China and Italy. These strains show little polymorphism, with just five segregating SNPs detected and none separating the *Psa*-V strains from NZ (but see below for results of additional analysis of polymorphism within *Psa*-V).

While each of the four major clades shows little within-clade diversity, the clades are distinguished by substantial SNP diversity. The LV clades differ from the other three clades by ~4,000 SNPs: the Japanese, Korean and outbreak clades each differ by ~1,000 SNPs. This level of between-lineage diversity is surprising. While reminiscent of differences among three known *Pto* lineages, contemporary *Pto* is comprised of just a single homogeneous (T1) lineage [62]. Consistent with low diversity, *Pto* T1 shows evidence of repeated selective sweeps with signatures of recurrent arms races with its tomato host [62]. A similar picture exists for *Pph* and *P. syringae* pv. *glycinea* (*Pgy*) [63]. In contrast, different outbreaks of *Psa* have been caused by three related, but nonetheless genetically distinct lineages sampled from a diverse source population. Whether these distinct lineages persist into the future is a matter of interest.

That there are such distinct lineages of *Psa* has significant implications, particularly in light of geographic origin. The first



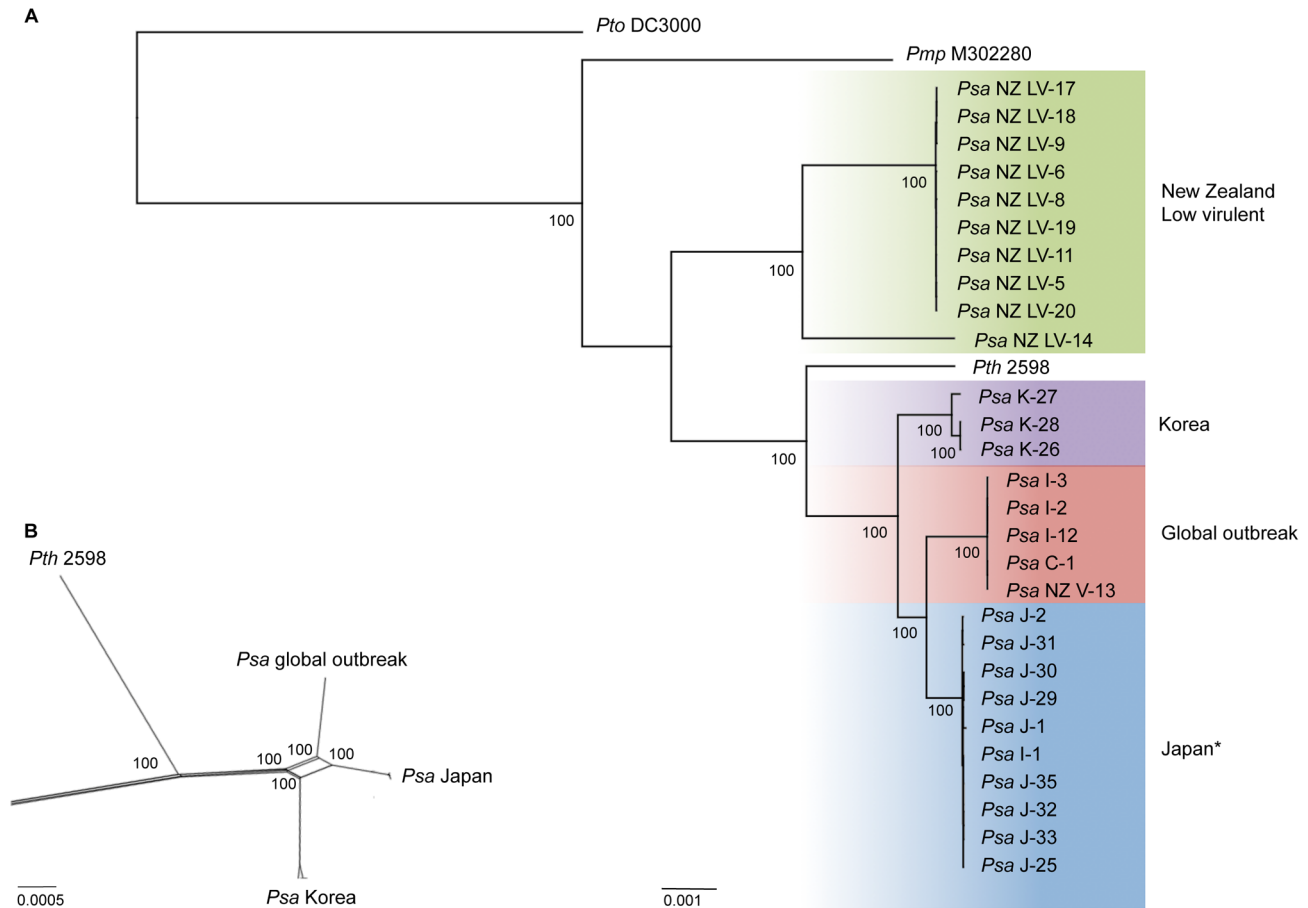
**Figure 1. Synteny plot of *Psa* NZ V-13 and *Psa* J-35.** MUMmer dotplot displaying stretches of conserved sequence between the genomes of *Psa* NZ V-13 and J-35 as lines with slope=1. Inverted and translocated stretches of conserved sequence are displayed as lines with slope = -1.  
doi:10.1371/journal.ppat.1003503.g001

outbreak of canker disease occurred in Japan in 1984 [14], but with the exception of a single transmission event to Italy (recorded in 1992 [17]), the Japanese lineage remained restricted to Japan. Reports of outbreaks in China were also evident in 1984, however no strains are available from these incidents [15]. The second recorded outbreak was from Korea, but our phylogenetic data clearly show that this was not the consequence of transmission from the population established in Japan, but was caused by a separate lineage distinguished by numerous SNPs and hundreds of unique genes (described below). The recent (2008) outbreak of *Psa*-V represents a third independent lineage, again, distinguished from the earlier lineages by large numbers of SNPs and numerous unique genes (described below). Similarly, the LV lineage detected in NZ in 2010 during surveillance for *Psa*-V represents a fourth clade, with the likelihood that *Psa*-LV14 is a fifth (*Psa*-LV14 differs from the other LV strains by  $\sim 1,400$  SNPs) [27]. The simplest explanation for this striking genetic structuring is the existence of a discrete but diverse source population from which independent transmission events have led to separate outbreaks of *Psa* disease.

#### Within pathovar recombination

While the phylogenetic reconstruction of *Psa* received strong bootstrap support (Figure 2A), we nonetheless explored the possibility of recombination. The Phi test [64], based on 6,346 informative sites, revealed statistically significant evidence of recombination ( $p < 0.0001$ ). Split Decomposition analysis [65] suggests this is a result of recombination among the Japanese, Korean and *Psa*-V strains (Figure 2B) [39]. An analysis based on 5,506 informative SNPs (from these three clades) showed 3,633 are congruent with the phylogeny depicted in Figure 2A, with the remainder being homoplasies due to recombination.

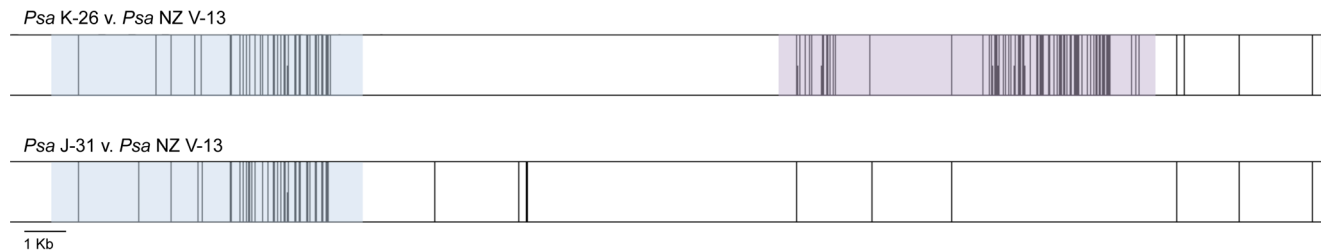
Homologous recombination among the Japanese, Korean and 2008 (*Psa*-V) outbreak lineages would add further weight to the notion of a source population comprised of a set of co-existing and co-evolving strains. If correct, and if recombination has taken place over sufficiently large genomic tracts, and in recent - as well as past - evolutionary time, then a signature of recombination should be evident by regions of clustered SNPs. To this end we mapped SNPs from a representative Korean (K-26) and Japanese



**Figure 2. Phylogeny of *Psa* and recombination between canker-causing *Psa* clades.** RAxML Maximum likelihood phylogenetic analysis of 32 draft and complete genome sequences based on 15,329 SNPs and 463,396 invariant sites (A). Each phylogenetic group is assigned its own color. With the exception of a single Italian strain (\*) isolated in 1992 grouping with the Japanese clade, the canker-causing Japanese, Korean and low-virulent NZ isolates form monophyletic clades reflecting their geographic origin, while global isolates from the 2008–2010 outbreak form a single clade. Bootstrap scores shown are based on 100 replicates. A SplitsTree analysis of recombination predicts recombination between canker-causing clades of *Psa* (B). All bootstrap scores are 100 (shown and otherwise). doi:10.1371/journal.ppat.1003503.g002

(J-31) strain onto the fully sequenced NZ V-13 genome to clarify the positional distribution of polymorphisms. Visual observation showed highly distinctive patterns of clustered SNPs shared between pairs of strains (Figure 3).

Such striking patterns, reminiscent of admixture in sexual populations, led to a statistical analysis of gene conversion (homologous recombination) events using GENECONV [40], which identifies conversion events between pairs of strains in an



**Figure 3. Shared and unique SNPs in *Psa* K-26 and J-31 compared to *Psa* NZ V-13.** Sample view from Artemis showing SNPs from *Psa* J-31 and K-26 aligned against a ~30 kb region of the NZ V-13 genome. Each line represents a SNP that distinguishes J-31 and/or K-26 from NZ V-13. Over this region, K-26 differs from NZ V-13 by 155 SNPs; J-31 differs by 65 SNPs. In the blue region there are 50 SNPs that distinguish J-31 and K-26 from NZ V-13. Of these, J-31 and K-26 are identical at 42 positions. GENECONV predicts that this region represents a gene conversion event between J-31 and K-26. The purple region denotes a gene conversion event into K-26 from a strain outside the set analyzed here. The set of Artemis input files allowing representation of the full set of SNPs, coverage of SNPs and regions identified by GENECONV as statistically supported regions involved in gene conversion events are available as Supplementary Dataset 1. doi:10.1371/journal.ppat.1003503.g003



alignment (and gene conversion events from strains outside the alignment). The results are shown in Table 2 and confirm the visual observations made above (data files for input into Artemis depicting positions of SNPs arising from comparison between K-26, J-31 and NZ V-13 are available as supplementary data (Dataset S1)). Mapping of the GENECONV-predicted gene conversion events onto the NZV-13 genome showed a close correspondence with regions of highly clustered SNPs (Dataset S1).

GENECONV predicts that approximately 10% of each genome is the result of homologous recombination with gene conversion often being between strains included in the data set. For example, Korean and Japanese lineages have exchanged ~300 kb (158 events of ~1.9 kb in length) with each other (Table 2). NZ V-13 shares most with J-31 (~427 kb scattered across 173 regions and comprising ~2.5 kb per gene conversion event).

While the overall patterns of SNPs, combined with analysis based on Split Decomposition, analysis of homoplasies, the Phi test and GENECONV strongly support the hypothesis that clustered SNPs are a consequence of gene conversion, it is conceivable that these regions might arise through positive or diversifying selection on specific genes, such as those involved in interactions between *Psa* and its plant host [66]. If so, then genes containing highly clustered SNPs are likely to be those with known or suspected roles in arms races with plant resistance genes. A scan of the genome suggests this not to be the case: most clustered regions span housekeeping genes. Nonetheless, using SNAP [67] to estimate numbers of synonymous and non synonymous nucleotide substitutions [68], we tested to see whether regions of the genome identified by GENECONV as resulting from gene conversion had more non-synonymous substitutions compared to regions of the genome identified by GENECONV as being involved in gene conversion. An analysis of the normalized ratio of the proportion of non-synonymous to synonymous substitutions (pN/pS) based on pairwise comparisons of ~4000 SNPs from K-26, J-31 and NZ V-13 showed pN/pS values less than 1 (the neutral condition) indicative of purifying – and not positive – selection: mean pN/pS ( $\pm$  standard deviation) for putative non-recombinant regions is  $0.131 \pm 0.004$  and for recombinant regions is  $0.088 \pm 0.002$ . We

therefore conclude that regions of clustered SNPs identified between pairs of strains by GENECONV are a consequence of gene conversion.

Notable from observations of patterns of SNPs mapped onto the NZ V-13 genome (Dataset S1) is the fact that some of the shared regions identified by highly clustered SNPs are identical, or almost identical, between pairs of strains (Figure 3). This indicates that some of the gene conversion events have taken place in recent evolutionary time. This, combined with the fact that each genome carries a signature of gene conversion events scattered around the chromosome, points to the fact that admixture among strains has been a persistent feature of the evolutionary history of these lineages. These patterns of gene conversion events substantiate the notion of a source population. It is difficult to imagine how such patterns could emerge without the sampled lineages having co-existed in close proximity for the bulk of their evolutionary history.

GENECONV provides a measure of the overall contribution of recombination to SNPs, relative to mutation. This ranges from ~8 to 21% for recombination events due to either gene conversion between known or unknown pairs of strains. Overall, ~44% of all polymorphic sites are likely to be the result of a recombination event. This means that SNPs are approximately twice as likely to be generated by mutation as by recombination. This level of recombination is twice as high as previously reported for the species [69], but still toward the clonal end of the spectrum of population structures [70].

#### SNP analysis of *Psa*-V from the 2008 outbreak lineage

The origin of the current global epidemic is the subject of interest and speculation, with reports that China is the likely source [22,27,29]. Previous studies identified six SNPs that distinguished European from Chinese isolates [22]; upon inclusion of two Chilean isolates, nine diagnostic SNPs were identified [27]. Our sequencing of multiple additional isolates of *Psa*-V confirms that the 2008 outbreak is from a single clone that has rapidly spread around the globe [22]. Whereas Mazzaglia et al [22] relied on read-mapping *Psa*-V to the partial genome sequence of *Pth* and Butler et al [27] adopted a similar approach using DC3000 as the

**Table 2.** Recombination events between *Psa* strains.

Strain pair <sup>1</sup>	Total length <sup>2</sup> (kb)	Regions <sup>3</sup>	Average length <sup>4</sup> (kb)	Proportion of genome <sup>5</sup>	Recombinant SNPs <sup>6</sup>	Proportion of recombinant SNPs
K-26 <sup>7</sup>	427.8	172	2.49	0.083	6,085	0.214
J-31 <sup>7</sup>	184.4	134	1.38	0.036	2,413	0.085
NZ V-13 <sup>7</sup>	302.4	159	1.90	0.059	4,218	0.149
NZ V-13 v. J-31	427.5	173	2.47	0.083	6,073	0.214
K-26 v. J-31	299.8	158	1.90	0.058	4,179	0.147
NZ V-13 v. K-26	183.9	132	1.39	0.036	2,371	0.083

The first three rows depict gene conversion events likely to have arisen from recombination events outside of the three compared strains. Rows 4–6 depict gene conversion events likely to have arisen from recombination events between pairs of strains (as listed). 12,716 unique recombinant SNPs were identified among all strains, out of a total of 28,403 SNPs.

<sup>1</sup>Strains used for pairwise comparison of recombination using GENECONV.

<sup>2</sup>Total length of genome affected by recombination events.

<sup>3</sup>Individual (discrete) regions involved in a recombination event ( $P < 0.05$ , GENECONV simulation with 10,000 permutations).

<sup>4</sup>Average length of recombination event.

<sup>5</sup>Proportion of genome affected by recombination.

<sup>6</sup>Total identified SNPs.

<sup>7</sup>Recombination events predicted by GENECONV to have arisen from outside the analyzed set of three genomes are largely the reciprocal of recombination events identified between pairs of strains.

doi:10.1371/journal.ppat.1003503.t002

**Table 3.** SNPs distinguishing *Psa-V* isolates.

	<i>Psa I-2</i>	<i>Psa I-3</i>	<i>Psa I-10</i>	<i>Psa I-12</i>	<i>Psa CI-4</i>	<i>Psa CI-5</i>	<i>Psa V-13</i>	<i>Psa C-1</i>	<i>Psa C-9</i> <sup>1</sup>
<i>Psa I-2</i>	0	52	28	36	37	34	38	38	365
<i>Psa I-3</i>	71 (19)	0	32	40	41	38	42	42	369
<i>Psa I-10</i>	45 (17)	34 (2)	0	16	17	14	18	18	345
<i>Psa I-12</i>	60 (24)	49 (9)	23 (7)	0	25	22	26	26	353
<i>Psa CI-4</i>	115 (78)	104 (63)	78 (61)	93 (68)	0	5	23	23	350
<i>Psa CI-5</i>	112 (78)	101 (63)	75 (61)	90 (68)	5 (0)	0	20	20	347
<i>Psa V-13</i>	346 (308)	335 (293)	309 (291)	324 (298)	337 (314)	334 (314)	0	24	351
<i>Psa C-1</i>	345 (307)	334 (296)	308 (290)	323 (297)	336 (313)	333 (313)	25 (1)	0	351
<i>Psa C-9</i>	558 (193)	547 (178)	521 (176)	536 (183)	587 (237)	584 (237)	818 (467)	817 (466)	0

SNPs due to mutation appear above the diagonal; total number of SNPs appear below the diagonal with the number due to recombination included in brackets.

<sup>1</sup>*Psa C-9* is included here for comparative purposes but is divergent from the global outbreak strains, as shown in Figure 4.

doi:10.1371/journal.ppat.1003503.t003

reference genome, we were able to map reads to the complete genome of NZ V-13. We also included the draft genome of the divergent Chinese strain C-9 (M228) in our analysis [27].

Overall we identified 979 SNPs – many of which are unique to C-9: 515 SNPs distinguish Italian, Chinese (excluding C-9), Chilean and New Zealand *Psa-V* strains (Table 3). However, most of these SNPs are located within integrative conjugative elements (see below) and are therefore the result of lateral gene transfer (the number of SNPs across these elements is in the tens of thousands, but are too polymorphic to be detected via read-mapping). Because SNPs due to recombination stand to mislead phylogenetic analyses recombinant SNPs were removed (Table 3) for subsequent analyses.

As evident in Figure 4, the Chinese strain C-9 is highly divergent [27]. It shares no SNPs with any of the other *Psa-V* isolates. On the basis of the available data – and excluding SNPs due to recombination – C-9 differs from C-1 by 351 SNPs (and other *Psa-V* strains by similar numbers).

Leaving aside the Chinese strain C-9, the Italian strains encompass the greatest level of polymorphism – a finding confirmed through SNP analysis of the genomes of six newly acquired (unpublished) Italian strains. This is consistent with the fact that *Psa-V* was first reported in Italy [17]. The Chilean strains are distinguished by five SNPs. No SNPs were identified among the five NZ *Psa-V* strains isolated immediately after the outbreak

in November 2010. Of the SNPs that distinguish the four Italian strains, two are shared. No SNPs are shared between NZ V-13 and C-1. The lack of informative sites means the phylogeny of *Psa-V* is star-like (Figure 4). The contemporary isolation of the divergent C-9 isolate is consistent with the suggestion that *Psa* is endemic on wild *Actinidia* relatives [27]. However, while C-9 clearly shares a common ancestor with *Psa-V*, C-9 is not the source of the global outbreak.

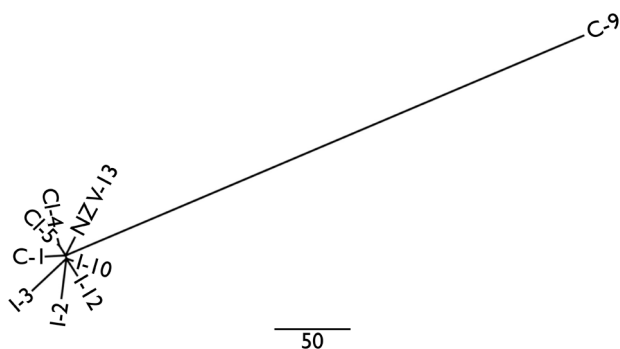
Assuming a mutation rate in the order of  $\sim 10^{-7}$  SNPs per site per year [32,71] and the fact that the most divergent clones differ by 52 SNPs, establishment of the *Psa-V* clone is likely to have occurred less than 10 years ago. C-9 is likely to have diverged from the outbreak clade  $\sim 100$  years ago.

### Identification of the core and accessory genomes

Identification of genes in the accessory pool often provides clues as to traits of ecological significance. Although there is considerable gene conservation within each *Psa* clade, there is also evidence of extensive gain and loss. The core genome of the *Psa-LV*, *Pth* 2598 and canker-causing Korean, Japanese and outbreak *Psa-V* strains includes 4425 orthologs, and the flexible genome includes an additional 4710 orthologs (Figure 5). The *P. syringae* core genome encodes many proteins contributing to its success in the phyllosphere and endophytic environments, however it is likely that some of these core proteins may present microbe-associated molecular patterns (MAMPs) recognized by plant host receptors [72]. MAMP-triggered immunity limits bacterial growth *in planta*, and can result in selection for amino acid diversification in elicitor-active regions of conserved pathogen proteins [73].

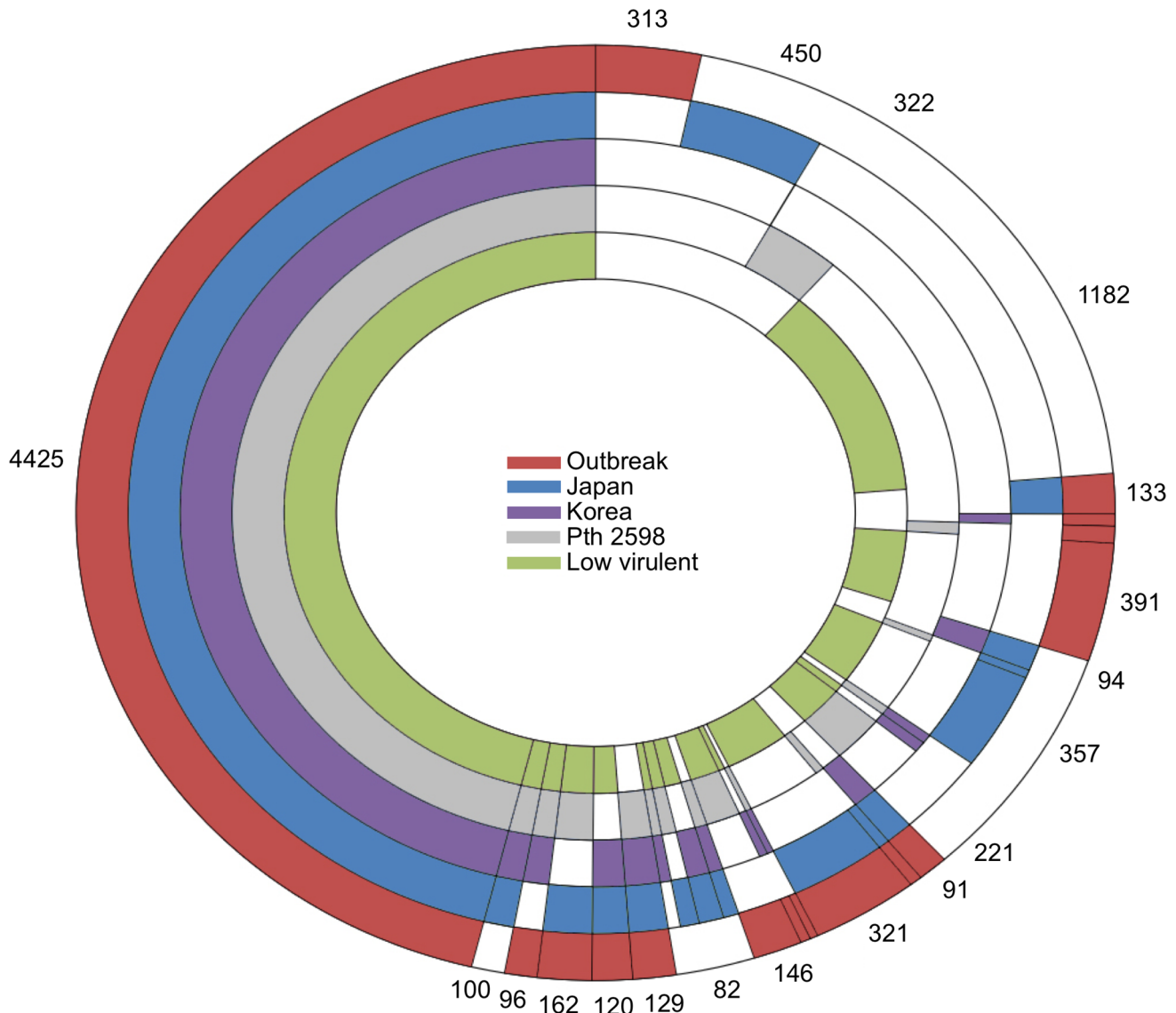
Of the 313 orthologs represented once or more among *Psa-V* strains, a set of 137 orthologs are found in all outbreak strains, and are absent from the Korean and Japanese *Psa* clades (Figure 5). These genes do not share a common history. The most similar orthologs of 39% of these genes are found in different pathogens across 4 different phylogroups of *P. syringae* and the remainder have close orthologs in plant-growth promoting soil bacteria as well as aquatic and insect associated bacteria, reflecting *P. syringae*'s cosmopolitan lifestyle (Table S3) [72]. Almost 10% of the outbreak-specific genes are annotated as phage integrases or transposases.

There are 18 genes present solely within outbreak strains which have top BLAST matches to known vascular and woody pathogens: these represent candidates most likely to confer a virulence advantage to *Psa-V*. These include genes from *P. syringae*



**Figure 4. Phylogeny of *Psa-V* isolates and the divergent Chinese isolate C-9.** Neighbor joining tree of *Psa-V* and C-9 isolates built in PHYLIP using SNPs due to mutation alone. These distances are also displayed in the upper right section of Table 3.

doi:10.1371/journal.ppat.1003503.g004



**Figure 5. Unique and shared ortholog groups between clades.** Numbers outside the rainbow plot show the number of ortholog groups with at least one representative ORFs per strain per clade.  
doi:10.1371/journal.ppat.1003503.g005

*pv. lachrymans* (*Pla*) and *P. syringae* *pv. pisi* (*Ppi*), which invade vascular tissues, and the strains responsible for bacterial canker disease in horse chestnut, plum, and citrus trees (*P. syringae* *pv. aesculi* (*Pae*); *Pmp* and *Xanthomonas axonopodis* *pv. citri* str. 306, respectively) [74–77].

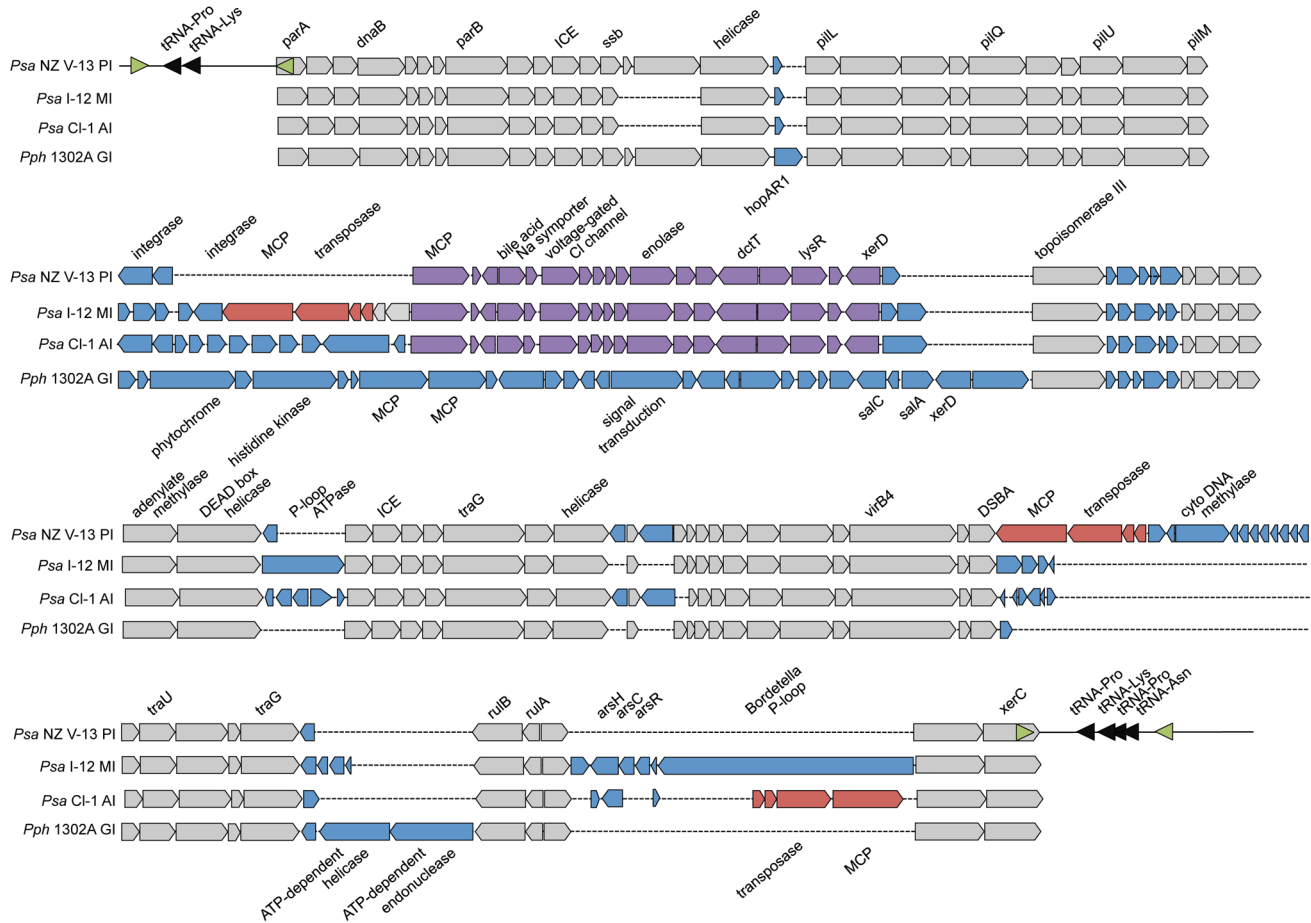
#### Diversifying selection in the core genome

An analysis of selection based on the core genome of all three canker-causing clades of *Psa* revealed 17 genes that exhibit significant signatures of positive selection (Table S2). Five of these candidates have non-synonymous substitutions exclusive to the outbreak clade, representing a shortlist of genes whose signatures of diversifying selection could be due to interactions with pattern-recognition receptors of *Actinidia* species. Curiously, this shortlist includes the flagellar P-ring protein FlgI, which forms a ring structure of the basal body in the peptidoglycan layer. The flagellar filament, comprised of multiple subunits of FlgC, is known to harbor an epitope that induces a strong innate immune

response in plants [62], but there are no reports suggesting this of FlgI [78]. Seven of the 17 candidates with high-confidence predictions of subcellular localization are targeted to different layers of the bacterial cell wall, consistent with their potential function as MAMPs [49].

#### The plasmid from *Psa* NZ V-13 possesses novel effector fusions and genes found in pathogens of woody plants

The NZ V-13 70.1 kb plasmid is a low copy number plasmid based on the small amount of DNA extracted from large-scale plasmid preparations. Consistent with this is the presence of the *stbD* and *stbE* genes which encode a putative segregational stability mechanism probably encoding a toxin-antitoxin system [79]. Analysis of the genes in the plasmid revealed a number of novel features. It possesses three genes encoding predicted T3SEs, HopAU1, HopAV1 and HopAA1-2. Two of these effectors possess novel fusions or insertions. HopAV1 has a C-terminal in-frame fusion to a TrbC conjugal transfer protein; and HopAA1-2



**Figure 6. Structure of the Pacific, Mediterranean, Andean and PPHGI-1 islands.** Grey genes are orthologs with ~75% nucleotide identity. Blue genes are variable accessory genes. Purple genes are accessory genes with complete conservation across the Pacific, Andean and Mediterranean islands. Red genes have translocated via transposon-mediated insertion events. Each island is bounded by 52 bp *att* sequences overlapping tRNA-Lys. Primer sites for the confirmation of excision and chromosomal integration are shown in green (Figure S4). doi:10.1371/journal.ppat.1003503.g006

has a 106 bp insertion causing a frame-shift. The plasmid also possesses two adjacent gene clusters involved in aromatic carbon metabolism. Both pathways are found in the woody pathogen *Pae* [72], and some of the genes are present in the xylem-limited pathogen *Xylella fastidiosa* [80]. The first cluster consists of two genes involved in the biosynthesis of anthranilate. The second cluster has the hallmarks of a secondary metabolic pathway for a secreted compound including putative regulatory and efflux proteins plus three genes with predicted roles in secondary metabolism. Intriguingly, the first of these is a putative phenylacetate CoA ligase for which anthranilate may be a substrate. It suggests this cluster is making an active compound with a role in the infection of vascular tissue.

### Genomic and pathogenicity islands in *Psa-V*

Many outbreak-specific genes are found clustered with mobile genetic elements in genomic and pathogenicity islands (GIs, PAIs), which are frequently linked to virulence and other environmental adaptations [51,81]. Sixteen genomic and pathogenicity islands (PAI) ranging in length from 10 to 100 kb were identified in NZ V-13 with IslandViewer. The PAIs are enriched with T3SEs as well as outbreak-specific genes. In addition, there are two prophages with homology to PSPPH06 and a Mu-like prophage. A total of 15 transposons harboring T3SEs or outbreak-specific genes were also identified.

### Integrative and conjugative elements

Mazzaglia et al [22] identified a polymorphic region with similarity to PPHGI-1, an integrative and conjugative element (ICE) identified in *Pph* 1302A [82]. This element was further characterized and shown to be capable of excision from the genome and circularization [27]. A hagfish plot mapping the Illumina reads from this region in *Psa* NZ V-13 showed no significant additional coverage compared to the rest of the genome suggesting that it is present as a single copy per genome (Figure S3).

Detailed analyses of the island present in *Psa* NZ V-13, referred to as the Pacific Island, revealed a 100 kb region comprised of ~100 predicted open reading frames bounded by *para* (left) and *xerC* (right) and integrated at a 52 base pair *att* locus (*att-1*) within a tRNA<sup>(lys)</sup> (Figure 6). Another copy of this *att* site (*att-2*) was found in the *Psa* NZ V-13 genome 2.7Mb from *att-1*. This site appears to be unoccupied in *Psa* NZ V-13, however the ICE from *Psa* I-12 is located at *att-2* and the corresponding *att-1* site is empty (Figure S4).

Comparison with the PPHGI-1 ICE shows striking similarities and marked differences. The “core” genes from the *Psa* NZ V-13 ICE are syntenous with PPHGI-1, although they share only ~75% nucleotide identity. Saturation at the third codon position indicates these elements have an ancient evolutionary past.

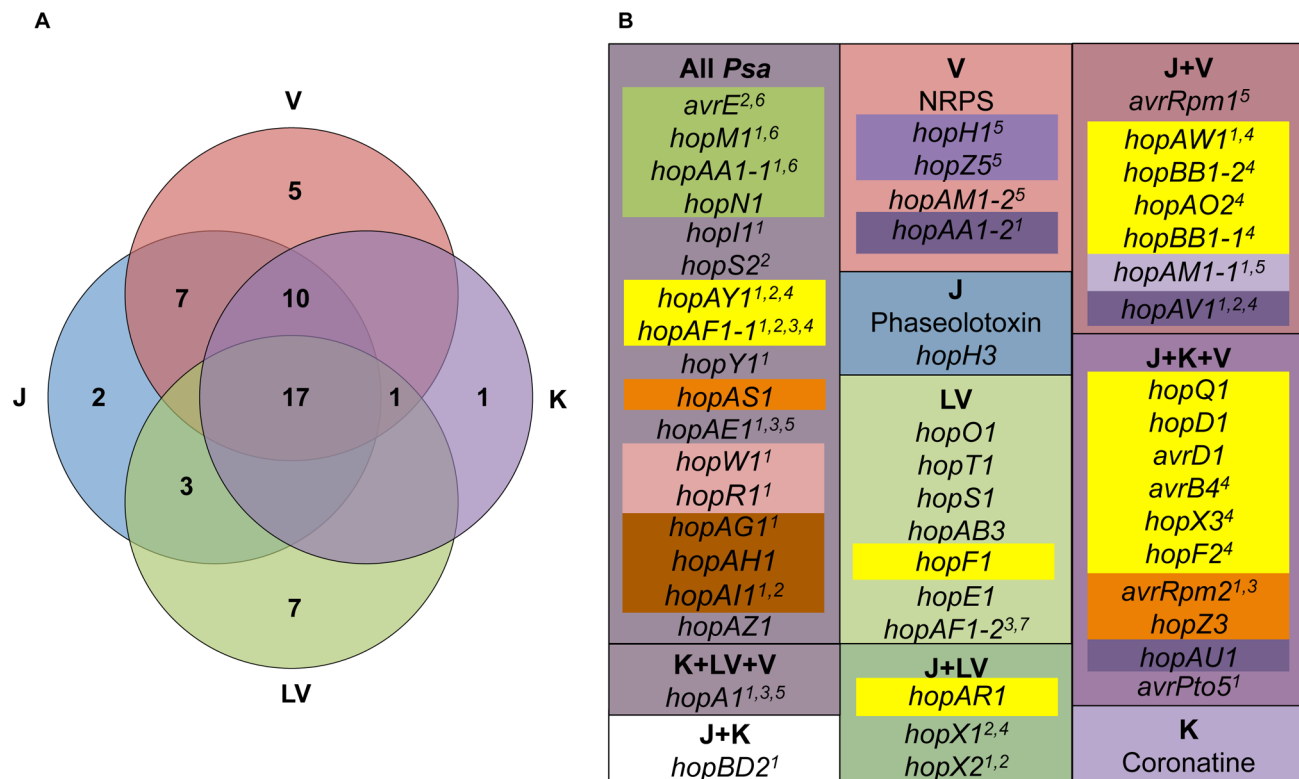
PPHGI-1 carries a number of accessory genes that are not present in the Pacific Island, which for example lacks the *hopAR1* effector. The Pacific Island also carries a novel set of genes located between the Pil locus and the conserved topoisomerase III (purple genes in Figure 6). This set of genes includes a predicted enolase and various transporters, including an ortholog of *dctT*, a putative dicarboxylic acid transporter and a methyl-accepting chemotaxis protein (MCP) predicted to be involved in taxis toward malate. Interestingly, the DctT transporter carries, at its N-terminus, a sequence predicted by SecretomeP version 2.0 to be used to target proteins for secretion (SecP score = 0.64 [83]). Analysis using EffectiveT3 to predict proteins targeted to the type 3 secretion system (T3SS) (trained on effector sequences from DC3000) returned a highly significant probability score (0.97), strongly suggestive of type 3-targeting [84]. In accordance with this prediction we note that 14% of the first 50 amino acids of the protein are either Ser or Pro [85]. No other transporter in this region of accessory genes carries a similar signal.

While this set of accessory genes awaits experimental analysis, putative functions of a number of the predicted proteins together suggest a role in manipulation of host cell metabolism. *Pseudomonas* has a preference for growth on TCA cycle intermediates and *dctA1* a dicarboxylate transporter gene is important for the virulence of *Pto* DC3000 [86]. It is possible that the putative T3SS-targeted di-carboxylic acid transporter, enters the plant cell and incorporates into a host cell membranes to facilitate export of this sugar acid in a manner analogous to that recently suggested for a

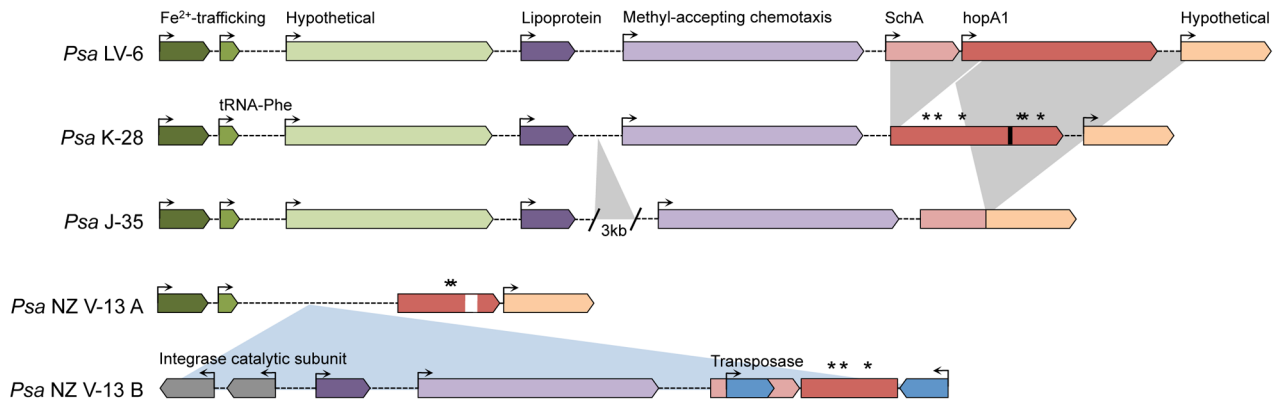
keto glutarate in *Xanthomonas oryzae* pv. *oryzae* [87]. Enolase is the penultimate step in glycolysis and the enzyme is exported as part of the RNA degradosome [88]. The role of the enolase in the ICE is not clear. One possibility is that it plays a role in the conversion of dicarboxylic acids to glucose, alternatively it might enhance activity of glycolysis in plant cells. In this regard it is of interest to note that a study of proteins differentially expressed in olive, with and without the pathogen *P. savastanoi* included *dctT* and plant enolase [89].

The NZ V-13 ICE is also present in the genome of the Chinese outbreak strain C-1, however, as previously indicated [22,29], this region is highly divergent in the four Italian and two Chilean outbreak strains [27]. A comparison of the Italian (Mediterranean Island), Chilean (Andean Island) and the Pacific Island is depicted in Figure 6.

Several notable findings emerge from the comparison of these Islands. Firstly, the core genes of the *Psa* islands are syntenous with those of PPHGI-1, but divergent at the nucleotide level. In fact the core genes of the Mediterranean Island are no more similar to the Pacific Island core genes as to PPHGI-1. Secondly, the enolase/*dctT* region is present all three *Psa* Islands and differs by just four SNPs. This suggests that this region has a recent evolutionary history and has an important function in *Psa* virulence. Thirdly, while there is near identity at the enolase/*dctT* locus there are a number of significant differences in the accessory genes. Most notable is the presence of a putative P-loop protein encoded by a gene of ~6 kb with highest similarity to a gene found in *Bordetella*



**Figure 7. Type 3 secreted effector and toxin distribution in *Psa* clades.** The numbers inside each region of the Venn diagram represent T3Es with orthologs present in the low-virulent (LV), Korean (K), Japanese (J) and recent global outbreak (V) clades (A). The outer boxes in B reflect the clade-specific distribution of T3Es (italicized) and toxins, while the color of the internal text boxes refers to their occurrence on genomic islands. <sup>1</sup>premature terminations, partial translocations, frame shifts, and out of frame indels in some strains. <sup>2</sup>variation between alleles (>2% variation, in frame fusions/indels). <sup>3</sup>indicates the T3SE may not be in all *Psa* K strains. <sup>4</sup>indicates the T3SE may not be in all *Psa* J strains. <sup>5</sup>T3SE that occur within a transposon region. <sup>6</sup>T3SE is present in the conserved effector locus. <sup>7</sup>LV has two effectors in the HopAF1 group, HopAF1-2 is most closely related to HopAF1-1 in the Korean clade while HopAF1-1 is most closely related to HopAF1-1 in the outbreak and Japanese clades. doi:10.1371/journal.ppat.1003503.g007



**Figure 8. Rearrangements, insertions and deletions in the *hopA1* locus.** The deletion of the *schA* chaperone and N-terminus of *hopA1* in K-28 and deletion of *hopA1* in J-35 are displayed with grey triangles. The transposon-mediated excision of a region in the *hopA1* locus and its reintegration 500 kb away in the *Psa* NZ V-13 genome is shown with a blue triangle. Arrows at the 5' end of the coding sequence indicate which alleles are functional. Stars above *hopA1* indicate the presence of non-synonymous (NS) mutations. The white bar in NZ V-13 refers to a deletion and the black bar in *Psa* K-28 indicates the position of a nonsense mutation in *hopA1*. The deletion in NZ V-13 includes one of the non-synonymous mutations in K-28.

doi:10.1371/journal.ppat.1003503.g008

(Figure 6). Finally, all *Psa* Islands carry an identical MCP/transposase region (red genes in Figure 6), which has recently been translocated – presumably via the associated transposon – in one or other of the ICEs. The picture of evolutionary change encompassed by these three laterally transferred ICEs is remarkable. Each displays a distinctly different evolutionary history that is independent of the host genome. Nonetheless, the presence of the conserved *enolase/dctT* region suggests that accessory gene content maybe determined by host-driven selection. Use of such dynamic elements as the basis upon which to infer *Psa*-V phylogeny is likely to generate erroneous conclusions.

### Extensive variation of *Psa* effector repertoires

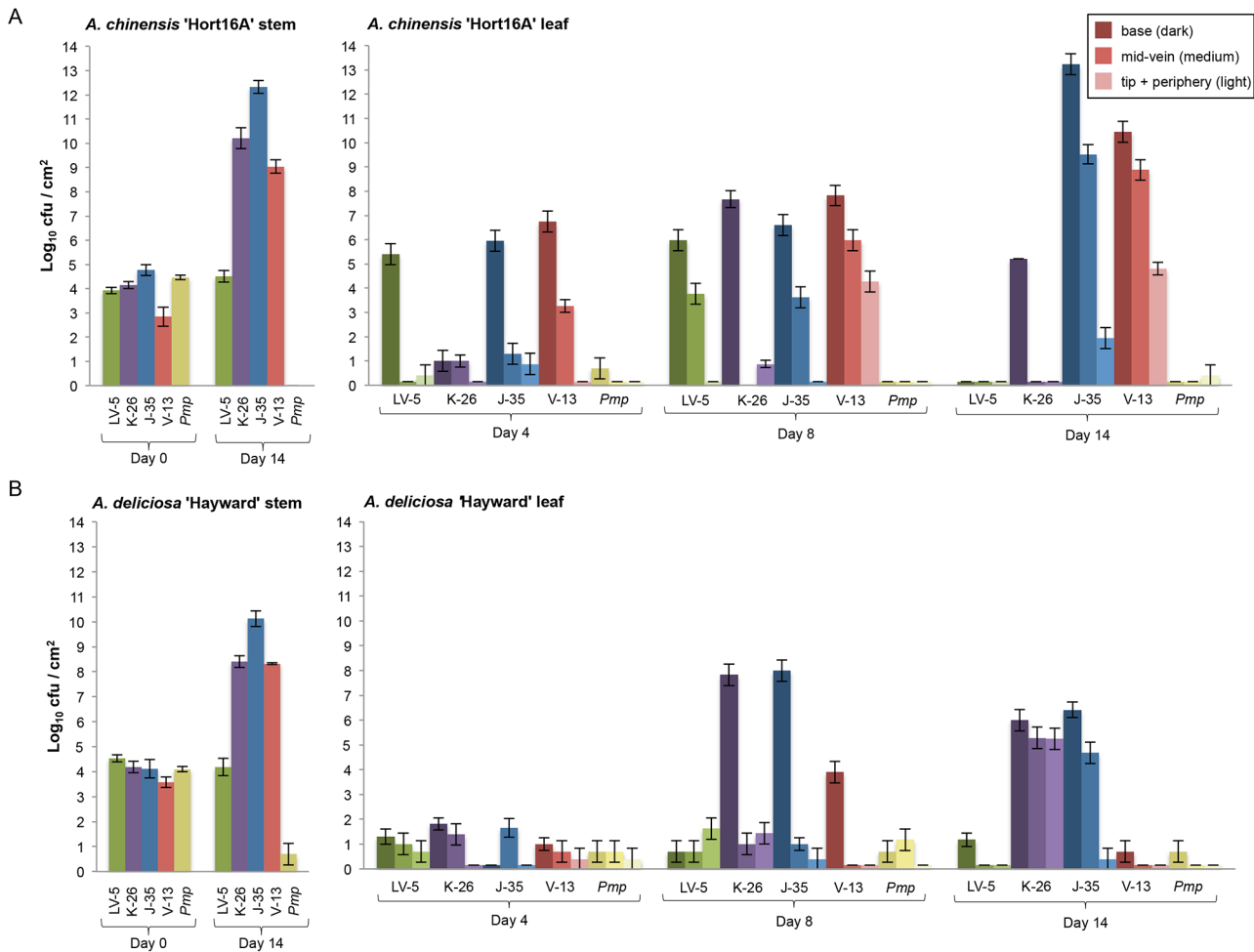
The sequenced *Psa* strains carry 51 known T3SEs (Figures 7 and S5). Seventeen T3SEs are found in all *Psa* genomes, while strains in the outbreak clade carry an average of 39 effectors, and strains in the low-virulent (LV) clade carry an average of 28. Many *Psa* T3SEs are not predicted to be translocated or functional due to the presence of frameshift mutations and transposon insertions. The effector complement is thus highly dynamic even within a single pathovar. T3SE gain, loss and adaptation are likely to have contributed to the enhanced virulence of the outbreak lineage. A large proportion of the variation in effector complement has occurred in a single pathogenicity island encompassing a locus with features similar to that described by Alfano et al, which we have therefore referred to as the exchangeable effector locus (EEL) [90]. Table S4 presents a comparison between effectors at the EEL across several isolates and suggests multiple re-arrangements, insertions and deletions have occurred in and around this island. The outbreak EEL is the largest with 16 T3SEs and the non-ribosomal peptide synthase (NRPS) cluster (Table S4). The pathotype strain *Psa* J-35 EEL is missing the NRPS cluster and has 13 T3SEs, while the Korean and low-virulent clades have fewer T3SEs in this region. The EEL appears to be absent from *Pth* altogether, although *hopAY1* is present. Three effectors in the Japanese and outbreak EEL (*hopBB1-1*, *hopBB1-2*, *hopX3*) share both a chaperone (*shcF*) and their predicted proteins share significant N-terminal sequence similarity to HopF2, as has been observed for AvrRpm2.

The disruption and loss of effectors may play a role in modulating virulence of the outbreak clade since this can result

in loss of host recognition. HopA1, for example, has been inactivated in multiple strains of *Psa* (Figure 8). A deletion of *shcA* and the 5' fragment of *hopA1* has occurred in *Psa* K-28; the fusion of *shcA* with an unknown gene downstream of *hopA1* has resulted in the deletion of *hopA1* in *Psa* J-35; and lastly, a transposon-mediated translocation has moved the 5' region of *hopA1* 600 kb away in *Psa* NZ V-13. In *Psa* K-26 *hopA1* contains six non-synonymous mutations; five of these are also present in the fragments of *hopA1* found in *Psa* NZ V-13. Six codons were also deleted in this strain; surprisingly, in the same position as the sixth substitution in *Psa* K-28. Only the NZ LV clade appears to have a fully functional version of HopA1, suggesting the inactivation of this effector contributes to vascular infection, potentially by abolishing resistance gene-mediated recognition. HopA1 is known to disrupt an *Arabidopsis* protein complex formed by EDS1, a central regulator of immunity, and RPS6, a NB-LRR resistance gene, activating effector-triggered immunity (ETI) [91]. Interestingly, HopA1-triggered immunity has also been shown to block the HopM1-mediated degradation of an *Arabidopsis* protein (AtMIN7) linked with vesicle transport and anti-microbial compound secretion [92].

HopM1 is broadly distributed among *P. syringae* pathovars and makes an important contribution to pathogen virulence by degrading AtMIN7. Inactivation of this key host target severely compromises both PAMP and effector triggered immunity (PTI, ETI) [92]. The transposon-mediated disruption of HopA1 in the outbreak strain may thus abolish ETI and allow HopM1 to interact with an *Actinidia* ortholog of AtMIN7. Transposon movement is doubly implicated in the enhanced virulence and transmissibility of the outbreak strain. The transposon-mediated introduction of HopH1 may have suppressed HopA1-induced ETI, as it has been shown to do in *Nicotiana* [93].

A small set of effectors appears to have been gained by the outbreak clade (and are largely absent from the J, K, and LV clades). This includes *hopH1* and *hopZ5*: both effectors are adjacent to each other on a small transposon, suggesting introduction by a lateral transfer event. *hopZ5* is a new member of the YopJ acetyltransferase family and shares greatest similarity to homologs from *Acidovorax* and *Xanthomonas*, although both these lack the putative myristoylation site (Figure S6). Members of the YopJ acetyltransferase family are believed to impede protein kinases by



**Figure 9. Pathogenicity assay of *Psa* and *Pmp* strains on kiwifruit.** The growth of the canker-causing *Psa* J-35 (blue), NZ V-13 (red), and K-26 (purple) isolates was assayed on the ‘Hort16A’ (A) and ‘Hayward’ (B) cultivars of kiwifruit, along with the low-virulent NZ LV-5 (green) and a strain of *P. syringae* pv. *morsprunorum* (*Pmp*, yellow), that causes canker disease in *Prunus* spp. The average bacterial density (cfu ± SE) was assayed in the stem tissue at day 0 immediately following stab inoculation, as well as in the base of the first leaf above the inoculation site (no *Psa* or *Pmp* observed, data not shown). The bacterial density was quantified in the base of the first leaf above the inoculation site (dark colored bar), the center of the leaf along the mid-vein (medium colored bar), and at the leaf tip and periphery (light colored bar) 4, 8 and 14 days after inoculation. A mock inoculation with MgSO<sub>4</sub> buffer was also performed, no *Psa* or *Pmp* growth was observed (not shown). doi:10.1371/journal.ppat.1003503.g009

trans-acetylating key serine or threonine residues in the kinase activation loop [94]. Recently HopZ1a has also been shown to interfere with microtubule assembly in the host, perhaps by binding to and acetylating tubulin – or indirectly via its effect on kinases as detailed above [95]. Another HopZ1 family member has been shown to interact with a protein involved in isoflavone biosynthesis, resulting in increased host susceptibility [96].

**Differences in virulence between *Psa* strains**

The identification of the effector repertoires and outbreak-specific genes may be linked with differences in host-specificity demonstrated by the growth of *Psa* strains on different kiwifruit cultivars. Growth assays of representative isolates from the low-virulent clade (NZ LV-5) and canker-causing Korean (K-26), Japanese (J-35), and outbreak (NZ V-13) were performed in *A. deliciosa* cv. ‘Hayward’ and *A. chinensis* ‘Hort16A’ in order to understand dynamics of *in planta* growth. Assays involved regular quantification of bacterial density at both the site of initial inoculation (stem) and at three spatially distinct regions of the first

leaf above the inoculation point (Figure S1B). This latter measure provides a quantitative indication of the capacity of a given strain to spread *in planta*. *Pmp*, which causes canker in *Prunus* spp., was included as a negative control.

Differences among treatments were tested using a three-way ANOVA within each cultivar, testing for effects of isolate (NZ V-13, NZ LV-5, J-35, K-26 and *Pmp*), time (days 0, 4, 8 and 14) and sample location (stem (point of initial inoculation), base of first leaf, middle of first leaf and periphery of first leaf), with Fisher’s protected LSD post-hoc test at α = 0.05 (Supplementary Table S5). On ‘Hort16A’ J-35 reached highest cell densities in the leaf at day 14, however, NZ V-13 spread more rapidly reaching the leaf tip by day eight (Figure 9A). K-26 showed reduced capacity for both growth and *in planta* spread. This capacity was further reduced in NZ LV-5. *Pmp* failed to spread from the point of inoculation, or persist within stem tissue, providing a compelling example of the effectiveness of non-host resistance in vascular plant tissue.

On cultivar ‘Hayward’ all strains showed a reduction in both growth and *in planta* spread (relative to ‘Hort16A’). This was

particularly striking for NZ V-13 which attained moderate leaf population densities by day eight, failed to spread beyond the leaf base, and was barely detectable by day 14 (Figure 9B).

Overall these data provide evidence of cultivar-specific differences in the growth dynamics of *Psa* from different clades. Noticeable is the enhanced growth and spread of NZ V-13 on Hort16A relative to its performance on 'Hayward'. This apparent tradeoff is suggestive of adaptation of this *Psa* lineage to *A. chinensis*. J-35, however, was equally capable of colonizing both 'Hort16A' and 'Hayward', although showed a reduced capacity for *in planta* spread. This suggests that differences in the gene content between NZ V-13 and J-35 may be central to understanding the enhanced transmission of 2008 outbreak strains both within and between plants [28].

While differences in the growth of all three canker-causing lineages were evident, similarities were also apparent. For example, all three, despite substantial difference in accessory gene context, were capable of spread and persistence within both kiwifruit cultivars. In light of the predicted existence of an Asian source population this finding has special relevance. The fact that three strains with such different genomes can each grow *in planta* raises the possibility that the capacity to cause disease in kiwifruit resides primarily in the core – as opposed to accessory – genome. If so, then it is possible that numerous different strains from the source population may be able to cause disease in kiwifruit. This highlights the importance of understanding the source population and minimizing the chance of future transmission events from this reservoir. It also suggests that strategies to develop durable resistance should not focus solely on effectors or other typically targeted components of the accessory genome.

## Conclusions

The origins of crop diseases are linked to domestication of plants [3]. Since most crops were brought under domestication many centuries ago, opportunities to understand the emergence of disease are limited. Kiwifruit is an exception: it is one of the few plants to have been domesticated in the 20<sup>th</sup> Century [12].

Our in-depth genomic and population analysis of more than 30 *Psa* strains from diseased kiwifruit vines – strains representing each of three major disease outbreaks obtained from different geographic regions over the course of three decades – has captured initial stages in the emergence of a pathogen population concomitant with domestication of its host. Particularly striking is the genetic structure of the *Psa* population. Rather than a single genetically homogeneous lineage that establishes in one region before transmission to another, each outbreak of *Psa* represents a distinct lineage with its own unique repertoire of accessory genes including effectors and toxins. There can be little doubt that each lineage represents an independent sampling from a single source population. This is evident in the shared ancestry (they form a single monophyletic group), the marked signatures of within pathovar gene conversion, and lack of diversity within each disease-causing lineage.

The geographic location of the predicted source population and the extent of diversity encompassed awaits elucidation, nonetheless, given the recent timing of kiwifruit domestication it is likely that the source population is associated with wild *Actinidia* species in Asia – a finding supported by the divergence of strain C-9 [27]. The diversity of genotype and virulence is likely substantive given knowledge of the heterogeneous characteristics of the lineages studied thus far. For example, the LV isolates from New Zealand cause leaf spot but minimal damage to vines, whereas *Psa-V* isolates cause cankers and stem wilt, are highly transmissible, and have devastating effects on vine health. Interestingly, the source

population encompasses *Pth*, the tea pathogen. Given that both hosts, tea and kiwifruit, have their origins in Asia, this finding is perhaps not surprising. Whether existence of two different pathovars within a single lineage represents an instance of host-shift or divergence from within the source population is unclear, but nonetheless highlights the importance of defining the nature of the source population.

The concept of a source population from which disease arises by sampling events has implications for both disease control and plant breeding. From the perspective of disease control there are two interrelated issues. The first is opportunity for transmission from the source population, the second is opportunity for a given transmission event to result in establishment of a new population. Every effort needs to be made to ensure that routes of transmission from the source population to commercial orchards – and between commercial orchards – are minimized. This includes protocols to prevent transmission of Japanese and Korean lineages, which our growth data show could be equally problematic, were they to arrive in countries such as New Zealand from which these lineages are currently absent. At the same time, there is need to move from orchards comprised of single clone varieties in order to reduce opportunity for plant-to-plant transmission and the establishment of high pathogen titers. It is possible that densely planted orchards of *A. chinensis* in Italy were central to the establishment and global dissemination of *Psa-V*.

Challenges to plant breeding are potentially more significant. Producing kiwifruit vines resistant to *Psa-V* is an important goal, however, resistance is unlikely to be durable if breeding programs exclusively target this lineage. The extensive diversity of effectors and associated virulence genes evident among the currently known *Psa* lineages, combined with capacity for lateral transmission, means that new *Psa* variants – either from the source population or produced *de novo* by recombination and horizontal gene transfer – are to be expected. This study has identified 17 effectors conserved across all *Psa* lineages that might usefully be the focus of a breeding program, while metabolic and physiological factors should not be overlooked. In this regard it may be of more than passing interest that the only significant genomic change that has thus far occurred among the extant *Psa-V* lineage involves three ICEs, which despite highly diverse sets of core genes contain an identical cargo of accessory genes predicted to manipulate host metabolism.

Human activities have long shaped the evolution of microbes [97] Ensuing feedback effects on the evolution of human populations – particularly in the context of human disease – are widely recognized [71,98,99]. Less well understood are anthropomorphic impacts on the evolution of plant pathogens [3]. Although our study provides just the first glimpse of a plant disease as it emerges in the face of domestication, continued analysis of the population processes and genetic phenomena underpinning evolution of *Psa* stand to enlighten both critical events driving the evolution of virulence and broader issues associated with the evolutionary interaction between microbes, their hosts, and the communities of which they are part.

## Supporting Information

**Dataset S1 Artemis input file for depicting positions of SNPs.** Read the readme file. Artemis is required to view the files: <http://www.sanger.ac.uk/resources/software/artemis/>. (ZIP)

**Figure S1 Inoculation and sampling areas for pathogenicity assay.** A sample image displaying stab inoculation on kiwifruit plantlets (A). A 2 µL drop of inoculum was suspended on the wound site created with a needle dipped in inoculum. Leaf



tissue samples for the quantification of bacterial density were taken as shown in B from the base (A), middle (B), tip (C) and periphery (D).

(TIF)

**Figure S2 Pulsed-field gel electrophoresis profiles of *Psa* NZ V-13 and J-35 plasmids.** Lanes 3 to 6 display *Psa* V-13 (32 µg) plasmid digested with the restriction enzymes *Bam*H1 (lane 3), *Hind*III (lane 4), *Eco*RI (lane 5) and *Not*I (lane 6). *Psa* J-35 (19 µg) plasmid digested with *Not*I is shown in lane 7. Kb+, Hyperladder I and Midrange II ladders are shown in lanes 1, 9 and 10. Gel electrophoresis conditions are stipulated in the methods section.

(TIF)

**Figure S3 Paired-end read depth and coverage of the *Psa* NZ V-13 chromosome, plasmid and Pacific Island.** Hagfish plot showing the read depth across the *Psa* NZ V-13 chromosome and plasmid (A) and the Pacific Island (B).

(TIF)

**Figure S4 Analysis of the location of the Pacific and Mediterranean islands in *Psa* NZ V-13 and *Psa* I-12.** (A) Analysis of the insertion site of the ICEs from *Psa* NZ V-13 and *Psa* I-12, and the ability of these elements to excise and circularise *in vitro*. PCR was carried out using primer sets designed to detect circularisation (lanes 2, 8, 15 and 21), excision (lanes 3, 9, 16 and 22) and to identify the *att* site the ICE was inserted into (lanes 4–7, 10–13, 17–20 and 23–26). Primer combinations and sequences are shown in (B) and (C) respectively.

(TIF)

**Figure S5 Type 3 secreted effector repertoires of *Psa* strains.** T3SE presence (black), absence (white) or presence with disruption, truncation or incomplete sequence (grey) is displayed for all sequenced genomes. Strains are colored according to their phylogenetic classification (Figure 2A). T3SE presence on a predicted genomic island or transposon is indicated in the bottom row. The exchangeable effector locus is designated in yellow.

(TIF)

**Figure S6 Phylogeny of HopZ effectors.** Phylogeny constructed using the Geneious tree builder based on HopZ protein alignments generated by ClustalW.

(TIF)

**Figure S7 Pathogenicity assay of *Psa* and *Pmp* strains on kiwifruit.** The growth of the canker-causing *Psa* J-35 (blue),

NZ V-13 (red), and K-26 (purple) isolates was assayed on the ‘Hort16A’ (A) and ‘Hayward’ (B) cultivars of kiwifruit, along with the low-virulent NZ LV-5 (green) and a strain of *P. syringae* pv. *morsprunorum* (*Pmp*, yellow) that causes canker disease in *Prunus* spp. The average bacterial density (cfu ± SE) was assayed in the stem tissue at day 0 immediately following stab inoculation, as well as in the base of the first leaf above the inoculation site (no *Psa* or *Pmp* observed, data not shown). The bacterial density was quantified in the base of the first leaf above the inoculation site (dark colored bar), the center of the leaf along the mid-vein (medium colored bar), and at the leaf tip and periphery (light colored bar) 4, 8 and 14 days after inoculation.

(TIF)

**Protocol S1 AWK-script findUniqueBlast.awk.**

(ZIP)

**Table S1 Comparison of *Psa* with other completed *P. syringae* genomes.**

(DOCX)

**Table S2 Core *Psa* genes under positive selection.**

(DOCX)

**Table S3 Outbreak clade-specific genes.**

(DOCX)

**Table S4 Comparison of the effector complement at the *Psa* EEL locus.**

(DOCX)

**Table S5 Pathogenicity assay analysis of variance.**

(DOCX)

## Acknowledgments

We thank Heath O’Brien (University of Toronto) and Richard Newcomb (PFR) for critically reviewing the manuscript. We are grateful to Tianchi Wang, Jonathan Crawford and Stephen Hoyte (all PFR) for generation and maintenance of kiwifruit plants. We thank Bevan Weir (Landcare Research) for advice on isolates and Cecilia Deng (PFR) for helping to generate Figure 5.

## Author Contributions

Conceived and designed the experiments: HCM EHAR MDT PBR DSG. Performed the experiments: HCM EHAR FB MF AL APG MTA BH JRG PWW CS JV PBR MDT. Analyzed the data: HCM EHAR FB MF AL APG BH DSG MWW PBR MDT. Wrote the paper: PBR HCM MDT EHAR FB.

## References

- Strange RN, Scott PR (2005) Plant disease: A threat to global food security. *Annu Rev Phytopathol* 43: 83–116.
- Wulff BBH, Horvath DM, Ward ER (2011) Improving immunity in crops: new tactics in an old game. *Curr Opin Plant Biol* 14: 468–476.
- Morris CE, Bardin M, Kinkel LL, Moury B, Nicot PC, et al. (2009) Expanding the paradigms of plant pathogen life history and evolution of parasitic fitness beyond agricultural boundaries. *PLoS Pathog* 5: e1000693.
- Kniskern JM, Barrett LG, Bergelson J (2011) Maladaptation in wild populations of the generalist plant pathogen *Pseudomonas syringae*. *Evolution* 65: 818–830.
- Stukenbrock EH, Banke S, Javan-Nikkah M, McDonald BA (2007) Origin and domestication of the fungal wheat pathogen *Mycosphaerella graminicola* via sympatric speciation. *Mol Biol Evol* 24: 398–411.
- Stukenbrock EH, Bataillon T (2012) A population genomics perspective on the emergence and adaptation of new plant pathogens in agro-ecosystems. *PLoS Pathog* 8: e1002893.
- Stukenbrock EH, Bataillon T, Duthéil JY, Hansen TT, Li RQ, et al. (2011) The making of a new pathogen: Insights from comparative population genomics of the domesticated wheat pathogen *Mycosphaerella graminicola* and its wild sister species. *Genome Res* 21: 2157–2166.
- Stukenbrock EH, Christiansen FB, Hansen TT, Duthéil JY, Schierup MH (2012) Fusion of two divergent fungal individuals led to the recent emergence of a unique widespread pathogen species. *Proc Natl Acad Sci U S A* 109: 10954–10959.
- Stukenbrock EH, Jorgensen FG, Zala M, Hansen TT, McDonald BA, et al. (2010) Whole-genome and chromosome evolution associated with host adaptation and speciation of the wheat pathogen *Mycosphaerella graminicola*. *PLoS Genet* 6: e1001189.
- Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. *Annu Rev Phytopathol* 46: 75–100.
- Li X, Li J, Doejarto DD (2009) Advances in the study of the systematics of *Actinidia* Lindley. *Frontiers of Biology in China* 4: 55–61.
- Ferguson AR, Huang H (2007) Genetic resources of kiwifruit: domestication and breeding. *Hortic Rev* 33: 1–121.
- Brook PJ (1990) Diseases of kiwifruit. In: Warrington IJW, G.C., editor. *Kiwifruit: science and management*. Auckland: Ray Richards. pp. 420–428.
- Serizawa S, Ichikawa T, Takikawa Y, Tsuyumu S, Goto M (1989) Occurrence of bacterial canker of kiwifruit in Japan: description of symptoms, isolation of the pathogen and screening of bacteriocides. *Annals Phytopathol Soc of Japan* 55: 427–436.
- Fang Y, Zhu X, Wang Y (1990) Preliminary studies on kiwifruit diseases in Hunan province. *Sichuan Fruit Science and Technology* 18: 28–29.
- Koh Y, Cha B, Chung H, Lee D (1994) Outbreak and spread of bacterial canker in kiwifruit. *Korean J Plant Pathol* 10: 68–72.
- Scortichini M (1994) Occurrence of *Pseudomonas syringae* pv. *actinidiae* on Kiwifruit in Italy. *Plant Pathol* 43: 1035–1038.

18. Koh YJ, Kim GH, Jung JS, Lee YS, Hur JS (2010) Outbreak of bacterial canker on Hort16A (*Actinidia chinensis* Planchon) caused by *Pseudomonas syringae* pv. *actinidiae* in Korea. *N Z J Crop Hortic Sci* 38: 275–282.
19. Scortichini M, Marcelletti S, Ferrante P, Petriccione M, Firrao G (2012) *Pseudomonas syringae* pv. *actinidiae*: a re-emerging, multi-faceted, pandemic pathogen. *Mol Plant Pathol* 13: 631–640.
20. Balestra GM, Mazzaglia A, Quattucci A, Renzi M, Rossetti A (2009) Occurrence of *Pseudomonas syringae* pv. *actinidiae* in Jin Tao kiwi plants in Italy. *Phytopathol Mediterr* 48: 299–301.
21. Ferrante P, Scortichini M (2010) Molecular and phenotypic features of *Pseudomonas syringae* pv. *actinidiae* isolated during recent epidemics of bacterial canker on yellow kiwifruit (*Actinidia chinensis*) in central Italy. *Plant Pathol* 59: 954–962.
22. Mazzaglia A, Studholme DJ, Taratufolo MC, Cai RM, Almeida NF, et al. (2012) *Pseudomonas syringae* pv. *actinidiae* (PSA) isolates from recent bacterial canker of Kiwifruit outbreaks belong to the same genetic lineage. *PLoS One* 7: e36518.
23. Chapman JR, Taylor RK, Weir BS, Romberg MK, Vanneste JL, et al. (2012) Phylogenetic relationships among global populations of *Pseudomonas syringae* pv. *actinidiae*. *Phytopathology* 102: 1034–1044.
24. Vanneste JL, Poliakoff F, Audusseau C, Cornish DA, Paillard S, et al. (2011) First report of *Pseudomonas syringae* pv. *actinidiae*, the causal agent of bacterial canker of kiwifruit in France. *Plant Dis* 95: 1311–1312.
25. Everett K, Taylor R, Romberg M, Rees-George J, Fullerton R, et al. (2011) First report of *Pseudomonas syringae* pv. *actinidiae* causing kiwifruit bacterial canker in New Zealand. *Australasian Plant Dis Notes* 6: 67–71.
26. Marcelletti S, Ferrante P, Petriccione M, Firrao G, Scortichini M (2011) *Pseudomonas syringae* pv. *actinidiae* draft genomes comparison reveal strain-specific features involved in adaptation and virulence to *Actinidia* species. *PLoS One* 6: e27297.
27. Butler MI, Stockwell PA, Black MA, Day RC, Lamont IL, et al. (2013) *Pseudomonas syringae* pv. *actinidiae* from recent outbreaks of kiwifruit bacterial canker belong to different clones that originated in China. *PLoS One* 8: e57464.
28. Vanneste JL, Yu J, Cornish DA, Tanner DJ, Windner R, et al. (2013) Identification, virulence and distribution of two biovars of *Pseudomonas syringae* pv. *actinidiae* in New Zealand. *Plant Dis* 97: 708–719.
29. Balestra GM, Taratufolo MC, Vinatzer BA, Mazzaglia A (2013) A multiplex PCR assay for detection of *Pseudomonas syringae* pv. *actinidiae* and differentiation of populations with different geographic origin. *Plant Dis* 97: 472–478.
30. Holt KE, Baker S, Weill FX, Holmes EC, Kitchen A, et al. (2012) *Shigella sonnei* genome sequencing and phylogenetic analysis indicate recent global dissemination from Europe. *Nat Genet* 44: 1056–1059.
31. Morelli G, Song YJ, Mazzoni CJ, Eppinger M, Roumagnac P, et al. (2010) *Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity. *Nat Genet* 42: 1140–1143.
32. Mutreja A, Kim DW, Thomson NR, Connor TR, Lee JH, et al. (2011) Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature* 477: 462–465.
33. Okoro CK, Kingsley RA, Connor TR, Harris SR, Parry CM, et al. (2012) Intracontinental spread of human invasive *Salmonella Typhimurium* pathovariants in sub-Saharan Africa. *Nat Genet* 44: 1215–1221.
34. Fraser C, Alm EJ, Polz MF, Spratt BG, Hanage WP (2009) The bacterial species challenge: making sense of genetic and ecological diversity. *Science* 323: 741–746.
35. Kassen R, Rainey PB (2004) The ecology and genetics of microbial diversity. *Annu Rev Microbiol* 58: 207–231.
36. King E, Ward M, Raney D (1954) Two simple media for the demonstration of pyocyanin and fluorescin. *J Lab Clin Med* 44: 301–307.
37. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W (2011) Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27: 578–579.
38. Sambrook J, Russell DW, editors (2001) Molecular cloning: a laboratory manual. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory.
39. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 23: 254–267.
40. Sawyer SA (1989) Statistical tests for detecting gene conversion. *Mol Biol Evol* 6: 526–538.
41. Chen F, Mackey AJ, Stoeckert CJ, Roos DS (2006) OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Res* 34: D363–D368.
42. Domazet-Lošo M, Haubold B (2011) Alignment-free detection of local similarity among viral and bacterial genomes. *Bioinformatics* 27: 1466–1472.
43. Loytynoja A, Goldman N (2005) An algorithm for progressive multiple alignment of sequences with insertions. *Proc Natl Acad Sci U S A* 102: 10557–10562.
44. Loytynoja A, Goldman N (2008) Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science* 320: 1632–1635.
45. Loytynoja A, Goldman N (2010) webPRANK: a phylogeny-aware multiple sequence aligner with interactive alignment browser. *BMC Bioinformatics* 11: 579. doi:10.1186/1471-2105-11-579.
46. Slater GS, Birney E (2005) Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6: 31. doi:10.1186/1471-2105-6-31
47. Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24: 1586–1591.
48. Bielawski JP, Yang Z (2005) Maximum likelihood methods for detecting adaptive protein evolution. In: Nielsen R, editor. *Statistical Methods in Molecular Evolution*. New York: Springer-Verlag. pp. 103–124.
49. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, et al. (2010) PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 26: 1608–1615.
50. Lindeberg M, Cartinhouer S, Myers CR, Schechter LM, Schneider DJ, et al. (2006) Closing the circle on the discovery of genes encoding Hrp regulon members and type III secretion system effectors in the genomes of three model *Pseudomonas syringae* strains. *Mol Plant-Microbe Interact* 19: 1151–1158.
51. Langille MGJ, Hsiao WWL, Brinkman FSL (2010) Detecting genomic islands using bioinformatics approaches. *Nat Rev Microbiol* 8: 372–382.
52. Feil H, Feil WS, Chain P, Larimer F, DiBartolo G, et al. (2005) Comparison of the complete genome sequences of *Pseudomonas syringae* pv. *syringae* B728a and pv. *tomato* DC3000. *Proc Natl Acad Sci U S A* 102: 11064–11069.
53. Redondo-Nieto M, Barret M, Morrissey JP, Germaine K, Martinez-Granero F, et al. (2012) Genome sequence of the biocontrol strain *Pseudomonas fluorescens* F113. *J Bacteriol* 194: 1273–1274.
54. Silby MW, Cerdeño-Tarraga AM, Vernikos GS, Giddens SR, Jackson RW, et al. (2009) Genomic and genetic analyses of diversity and plant interactions of *Pseudomonas fluorescens*. *Genome Biol* 10: R51.
55. Paulsen IT, Press C, Ravel J, Kobayashi D, Myers GS, et al. (2005) Complete genome sequence of the plant commensal *Pseudomonas fluorescens* Pf-5. *Nat Biotechnol* 23: 873–878.
56. Vodovar N, Vallet D, Cruveiller S, Rouy Z, Barbe V, et al. (2006) Complete genome sequence of the entomopathogenic and metabolically versatile soil bacterium *Pseudomonas entomophila*. *Nat Biotechnol* 24: 673–679.
57. Matilla MA, Pizarro-Tobias P, Roca A, Fernandez M, Duque E, et al. (2011) Complete genome of the plant growth-promoting rhizobacterium *Pseudomonas putida* BIRD-1. *J Bacteriol* 193: 1290.
58. Payne RW, Murray DA, Harding SA, Baird DB, Soutar DM (2011) An introduction to GenStat for windows. Hemel Hempstead, UK: VSN International.
59. Rees-George J, Vanneste JL, Cornish DA, Pusharajah IPS, Yu J, et al. (2010) Detection of *Pseudomonas syringae* pv. *actinidiae* using polymerase chain reaction (PCR) primers based on the 16S-23S rDNA intertranscribed spacer region and comparison with PCR primers based on other gene regions. *Plant Pathol* 59: 453–464.
60. Buell CR, Joardar V, Lindeberg M, Selengut J, Paulsen IT, et al. (2003) The complete genome sequence of the Arabidopsis and tomato pathogen *Pseudomonas syringae* pv. *tomato* DC3000. *Proc Natl Acad Sci U S A* 100: 10181–10186.
61. Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theor Popul Biol* 7: 256–276.
62. Cai RM, Lewis J, Yan SC, Liu HJ, Clarke CR, et al. (2011) The plant pathogen *Pseudomonas syringae* pv. *tomato* is genetically monomorphic and under strong selection to evade tomato immunity. *PLoS Pathog* 7: e1002130.
63. Baltrus DA, Nishimura MT, Dougherty KM, Biswas S, Mukhtar MS, et al. (2012) The molecular basis of host specialization in bean pathovars of *Pseudomonas syringae*. *Mol Plant-Microbe Interact* 25: 877–888.
64. Bruen TC, Philippe H, Bryant D (2006) A simple and robust statistical test for detecting the presence of recombination. *Genetics* 172: 2665–2681.
65. Bandelt HJ, Dress AWM (1992) Split decomposition: a new and useful approach to phylogenetic analysis of distance data. *Mol Phylogenet Evol* 1: 242–252.
66. Boller T, He SY (2009) Innate immunity in plants: an arms race between pattern recognition receptors in plants and effectors in microbial pathogens. *Science* 324: 742–744.
67. Korber B (2000) HIV signature and sequence variation analysis. In: Learn AGRaGH, editor. *Computational analysis of HIV molecular sequences*. Dordrecht: Kluwer Academic Publishers. pp. 55–72.
68. Nei M, Gojori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 3: 418–426.
69. Sarkar SF, Guttman DS (2004) Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. *Appl Environ Microbiol* 70: 1999–2012.
70. Feil EJ, Spratt BG (2001) Recombination and the population structures of bacterial pathogens. *Annu Rev Microbiol* 55: 561–590.
71. Lowder BV, Guinane CM, Ben Zakour NL, Weinert LA, Conway-Morris A, et al. (2009) Recent human-to-poultry host jump, adaptation, and pandemic spread of *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* 106: 19545–19550.
72. Baltrus DA, Nishimura MT, Romanchuk A, Chang JH, Mukhtar MS, et al. (2011) Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathog* 7: e1002132.
73. McCann HC, Nahal H, Thakur S, Guttman DS (2012) Identification of innate immunity effectors using molecular signatures of natural selection. *Proc Natl Acad Sci U S A* 109: 4215–4220.
74. Brunings AM, Gabriel DW (2003) *Xanthomonas citri*: breaking the surface. *Mol Plant Pathol* 4: 141–157.
75. da Silva ACR, Ferro JA, Reinach FC, Farah CS, Furlan LR, et al. (2002) Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* 417: 459–463.

76. Green S, Studholme DJ, Laue BE, Dorati F, Lovell H, et al. (2010) Comparative genome analysis provides insights into the evolution and adaptation of *Pseudomonas syringae* pv. *aesculi* on *Aesculus hippocastanum*. PLoS One 5: e10224.
77. Roos IMM, Hattin MJ (1987) Systemic invasion of plum leaves and shoots by *Pseudomonas syringae* pv. *syringae* introduced into petioles. Phytopathology 77: 1253–1257.
78. Felix G, Duran JD, Volko S, Boller T (1999) Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. The Plant Journal 18: 265–276.
79. Sengupta M, Austin S (2011) Prevalence and significance of plasmid maintenance functions in the virulence plasmids of pathogenic bacteria. Infect Immun 79: 2502–2509.
80. Ciraulo MB, Santos DS, Rodrigues A, de Oliveira MV, Rodrigues T, et al. (2010) Transcriptome analysis of the phyto bacterium *Xylella fastidiosa* growing under xylem-based chemical conditions. J Biomed Biotechnol article ID 781365. DOI:10.1155/2010/781365.
81. Jackson RW, Vinatzer B, Arnold DL, Dorus S, Murillo J (2011) The influence of the accessory genome on bacterial pathogen evolution. Mob Genet Elements 1: 55–65.
82. Pitman AR, Jackson RW, Mansfield JW, Kaitell V, Thwaites R, et al. (2005) Exposure to host resistance mechanisms drives evolution of bacterial virulence in plants. Curr Biol 15: 2230–2235.
83. Bendtsen JD, Kiemer L, Fausboll A, Brunak S (2005) Non-classical protein secretion in bacteria. BMC Microbiology 5: 58. doi:10.1186/1471-2180-5-58.
84. Arnold R, Brandmaier S, Kleine F, Tischler P, Heinz E, et al. (2009) Sequence-based prediction of type III secreted proteins. PLoS Pathog 5: e1000376.
85. Schechter LM, Roberts KA, Jamir Y, Alfano JR, Collmer A (2004) *Pseudomonas syringae* type III secretion system targeting signals and novel effectors studied with a *Cya* translocation reporter. J Bacteriol 186: 543–555.
86. Mellgren EM, Kloek AP, Kunkel BN (2009) Mqo, a tricarboxylic acid cycle enzyme, is required for virulence of *Pseudomonas syringae* pv. *tomato* strain DC3000 on *Arabidopsis thaliana*. J Bacteriol 191: 3132–3141.
87. Guo W, Cai LL, Zou HS, Ma WX, Liu XL, et al. (2012) Ketoglutarate transport protein KgtP is secreted through the type III secretion system and contributes to virulence in *Xanthomonas oryzae* pv. *oryzae*. Appl Environ Microbiol 78: 5672–5681.
88. Taghbalout A, Rothfield L (2007) RNaseE and the other constituents of the RNA degradosome are components of the bacterial cytoskeleton. Proc Natl Acad Sci U S A 104: 1667–1672.
89. Campos A, da Costa G, Coelho AV, Fevereiro P (2009) Identification of bacterial protein markers and enolase as a plant response protein in the infection of *Olea europaea* subsp. *europaea* by *Pseudomonas savastanoi* pv. *savastanoi*. Eur J Plant Pathol 125: 603–616.
90. Alfano JR, Charkowski AO, Deng W-L, Badel JL, Petnicki-Ocwieja T, et al. (2000) The *Pseudomonas syringae* Hrp pathogenicity island has a tripartite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in plants. Proc Natl Acad Sci U S A 97: 4856–4861.
91. Bhattacharjee S, Halane MK, Kim SH, Gassmann W (2011) Pathogen effectors target Arabidopsis EDS1 and alter its interactions with immune regulators. Science 334: 1405–1408.
92. Nomura K, Meccey C, Lee YN, Imboden LA, Chang JH, et al. (2011) Effector-triggered immunity blocks pathogen degradation of an immunity-associated vesicle traffic regulator in Arabidopsis. Proc Natl Acad Sci U S A 108: 10774–10779.
93. Guo M, Tian F, Wamboldt Y, Alfano JR (2009) The majority of the type III effector inventory of *Pseudomonas syringae* pv. *tomato* DC3000 can suppress plant immunity. Mol Plant-Microbe Interact 22: 1069–1080.
94. Mukherjee S, Keitany G, Li Y, Wang Y, Ball HL, et al. (2006) *Yersinia* YopJ acetylates and inhibits kinase activation by blocking phosphorylation. Science 312: 1211–1214.
95. Lee AHY, Hurley B, Felsensteiner C, Yea C, Ckurshumova W, et al. (2012) A bacterial acetyltransferase destroys plant microtubule networks and blocks secretion. PLoS Pathog 8: e1002523.
96. Zhou H, Lin J, Johnson A, Morgan RL, Zhong W, et al. (2011) *Pseudomonas syringae* type III effector HopZ1 targets a host enzyme to suppress isoflavone biosynthesis and promote infection in soybean. Cell Host Microbe 9: 177–186.
97. Cohen ML (2000) Changing patterns of infectious disease. Nature 406: 762–767.
98. Ferguson NM, Donnelly CA, Anderson RM (2001) The foot-and-mouth epidemic in Great Britain: Pattern of spread and impact of interventions. Science 292: 1155–1160.
99. Nathanson N, Wilesmith J, Griot C (1997) Bovine spongiform encephalopathy (BSE): Causes and consequences of a common source epidemic. Am J Epidemiol 145: 959–969.