Origin and Evolution of the Kiwifruit Canker Pandemic

Honour C. McCann[1,*,†], Li Li[2,†], Yifei Liu[3], Dawei Li[2], Hui Pan[2], Caihong Zhong[2], Erik H.A. Rikkerink[4], Matthew D. Templeton[4,5], Christina Straub[1], Elena Colombi[1], Paul B. Rainey[1,6,7,*,‡], and Hongwen Huang[2,3,*,‡]

[1]New Zealand Institute for Advanced Study, Massey University, Auckland, New Zealand

[2]Key Laboratory of Plant Germplasm Enhancement and Specialty Agriculture, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan, China

[3]Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China

[4]The New Zealand Institute for Plant and Food Research Limited, Auckland, New Zealand

[5]School of Biological Sciences, University of Auckland, New Zealand

[6]Department of Microbial Population Biology, Max Planck Institute for Evolutionary Biology, Plön, Germany

[7]École Supérieure de Physique et de Chimie Industrielles de la Ville de Paris (ESPCI ParisTech), CNRS UMR 8231 PSL Research University, Paris, France

†These authors contributed equally to this work.

‡Cosenior authors.

*Corresponding authors: E-mails: h.mccann@massey.ac.nz; rainey@evolbio.mpg.de; huanghw@scbg.ac.cn.

Accepted: March 13, 2017

Data deposition: Accession numbers (SRA and GenBank) are provided for all new sequence data analyzed in Table S1. Data for these 50 genomes will be publicly available immediately upon publication of this paper.

Abstract

Recurring epidemics of kiwifruit (*Actinidia* spp.) bleeding canker disease are caused by *Pseudomonas syringae* pv. *actinidiae* (*Psa*). In order to strengthen understanding of population structure, phylogeography, and evolutionary dynamics, we isolated *Pseudomonas* from cultivated and wild kiwifruit across six provinces in China. Based on the analysis of 80 sequenced *Psa* genomes, we show that China is the origin of the pandemic lineage but that strain diversity in China is confined to just a single clade. In contrast, Korea and Japan harbor strains from multiple clades. Distinct independent transmission events marked introduction of the pandemic lineage into New Zealand, Chile, Europe, Korea, and Japan. Despite high similarity within the core genome and minimal impact of within-clade recombination, we observed extensive variation even within the single clade from which the global pandemic arose.

Key words: pathogen evolution, genomic epidemiology, bacterial plant pathogen, plant-microbe interactions, disease emergence.

Introduction

A pandemic of kiwifruit (*Actinidia* spp.) bleeding canker disease caused by *Pseudomonas syringae* pv. *actinidiae* (*Psa*) emerged in 2008 with severe consequences for production in Europe, Asia, New Zealand, and Chile (Balestra et al. 2010; Abelleira et al. 2011; Everett et al. 2011; Vanneste et al. 2011; Koh et al. 2012; Zhao et al. 2013; Sawada et al. 2015). Earlier disease epidemics in China, South Korea, and Japan had regional impacts, however as infections were often lethal and the pathogen rapidly disseminated, *Psa* was predicted to pose a major threat to global kiwifruit production (Serizawa et al. 1989; Koh et al. 2002). Despite recognition of this threat—one subsequently realized in 2008—little was done to advance understanding of population structure, particularly across regions of eastern Asia that mark the native home of the genus *Actinidia*.

The origins of agricultural diseases and their link with plant domestication is shrouded by time. Kiwifruit (*Actinidia* spp.) is a rare exception because domestication occurred during the last century (Ferguson and Huang 2007; Ferguson 2011). Kiwifruit production and trade in plant material for commercial and breeding purposes has recently increased in Asia, Europe, New Zealand, and Chile (Shim and Ha 1999; Huang et al. 2004; Testolin and Ferguson 2009; Cruzat 2014;

© The Author(s) 2017. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Ferguson 2015), preceding the emergence of disease in some cases by less than a decade.

The first reports of a destructive bacterial canker disease in green-fleshed kiwifruit (*A. chinensis* var. *deliciosa*) came from Shizuoka, Japan (Takikawa et al. 1989; Fang et al. 1990). The causal agent was described as *Pseudomonas syringae* pv. *actinidiae* (*Psa*) (Takikawa et al. 1989). An outbreak of disease with symptoms similar to those produced by *Psa* was reported to have occurred in 1983–1984 in Hunan, China, though no positive identification was made or isolates stored at that time (Fang et al. 1990). *Psa* was also isolated from infected green kiwifruit in Korea shortly thereafter (Koh et al. 1994). The cultivation of more recently developed gold-fruiting cultivars derived from *A. chinensis* var. *chinensis* (e.g., "Hort16A") began only in the 2000s and an outbreak of global proportions soon followed. The first published notices of the latest outbreak on gold kiwifruit issued from Italy in 2008, with reports from neighboring European countries, New Zealand, Asia, and Chile occurring soon after (Balestra et al. 2010; Abelleira et al. 2011; Everett et al. 2011; Vanneste et al. 2011; European and Mediterranean Plant Protection Organization 2011; Koh et al. 2012; Sawada et al. 2015). Whole genome sequencing showed the most recent global outbreak of disease was caused by a new clade of *Psa* (previously referred to as *Psa*-V and now referred to as *Psa*-3), whereas earlier disease incidents in Japan and Korea were caused by strains forming separate clades referred to as *Psa*-1 (previously *Psa*-J) and *Psa*-2 (previously *Psa*-K), respectively (Marcelletti et al. 2011; Mazzaglia et al. 2012; Butler et al. 2013; McCann et al. 2013). These clades are marked by substantial variation in their complement of type III secreted effectors, which are required for virulence in *P. syringae*. Despite the surprising level of within-pathovar differences in virulence gene repertoires occurring subsequent to the divergence of these three clades, strains from each clade are capable of infecting and growing to high levels in both *A. chinensis* var. *deliciosa* and *A. chinensis* var. *chinensis* (McCann et al. 2013).

The severity of the latest global outbreak is largely predicated on the expansion in cultivation of clonally propagated highly susceptible *A. chinensis* var. *chinensis* cultivars, with trade in plant material and pollen likely providing opportunities for transmission between distant geographic regions. Identifying the source from which *Psa* emerged to cause separate outbreaks remains an important question. Intriguingly, despite the divergence in both the core and flexible genome, these distinct clades nevertheless exhibit evidence of recombination with each other and unknown donors (McCann et al. 2013). This suggests each clade emerged from a recombining source population. Definitive evidence for the location, extent of diversity, and evolutionary processes operating within this population remain elusive. Early reports suggested China may be the source of the latest global outbreak (Butler et al. 2013; McCann et al. 2013). Although the strains of *Psa* available at that time did not provide unambiguous and well-supported evidence of a Chinese origin, this speculation was based on the fact that kiwifruit are native to China; it is the provenance of the plant material selected for commercial and breeding purposes in China, New Zealand, Italy, and other kiwifruit growing regions; there is extensive trade in plant material between all of these regions; and one Chinese isolate was found to carry an integrative and conjugative element (ICE) that was also found in New Zealand *Psa*-3 isolates (Butler et al. 2013).

In order to strengthen understanding of the population structure, phylogeography, and evolutionary dynamics of *Psa*, we isolated *Psa* from cultivated kiwifruit across six provinces in China and obtained additional isolates from South Korea and Japan. Genome sequencing of 50 isolates and the inclusion of an additional 30 previously sequenced isolates show that whereas China is the origin of the pandemic lineage of *Psa*, only a single clade is currently present in China, whereas strains from multiple clades are present in both Korea and Japan. Strains from the pandemic lineage are closely related to each other, displaying reduced pairwise nucleotide diversity relative to that exhibited between strains in other clades, indicating a more recent origin. Distinct transmission events were responsible for introduction of the pandemic lineage of *Psa* into New Zealand, Chile, and Europe. Two independent transmission events occurred between China and Korea, and two Japanese isolates from 2014 cluster with New Zealand *Psa*. Despite high similarity at the level of the core genome and negligible impact of within-lineage recombination, there has been substantial gene gain and loss even within the single clade from which the global pandemic arose.

Materials and Methods

Bacterial Strains and Sequencing

Sampling regions were chosen to include areas with extensive kiwifruit cultivation and/or reservoirs of wild *Actinidia*, and included areas where modern kiwifruit cultivars originated (Huang 2014). Plants sampled were classified as asymptomatic, atypical or symptomatic for *Psa* infection. Sampling of wild plants occurred after first surveying the disease status of all identifiable *Actinidia* spp. in the site: no visible symptoms of disease were found among all wild *Actinidia* spp. observed. A total of 280 plants were sampled (148 cultivated, 132 wild) across eight provinces (Anhui, Chongqing, Guizhou, Henan, Hubei, Hunan, Shaanxi, Sichuan). Bacterial strain isolations were performed from same-day sampled leaf and stem tissue by homogenizing leaf or stem tissue in 800 µl 10 mM $MgSO_4$ and plating the homogenate on *Pseudomonas* selective media (King's B supplemented with cetrimide, fucidin, and cephalosporin, Oxoid). Isolations were performed in hotel rooms; movement to the next sampling location (including mobile laboratory relocation) resulted in variable incubation temperatures. Plates were incubated 48 h between 25

and 30 °C. Subsequent to the return to the laboratory (Wuhan Botanical Garden, Hubei), single colonies were restreaked and tested for oxidase activity and used to inoculate liquid overnight cultures in KB. Strains were then stored at −80 °C in 15% glycerol and the remainder of the liquid culture was reserved for genomic DNA isolation by freezing the pelleted bacterial cells at −20 °C. Genomic DNA extractions were performed using the Promega Wizard 96-well genomic DNA purification system.

Initial strain identification was performed by sequencing the citrate synthase gene (*cts*, aka *gltA* [Sarkar and Guttman 2004]). 1419 total single colony isolates were recovered, of which 819 were identified. Because no *Psa* was present among the identified isolates from Anhui and Henan, these provinces are not represented in our data. No *Psa* was isolated from asymptomatic plants or plants displaying atypical symptoms of infection. Subsequent to strain identification, 2 × 125 bp paired-end sequencing was performed using the Illumina HiSeq 2500 platform using 500 bp libraries (Novogene, Guangzhou, China). Additional 2 × 250 bp paired-end sequencing was performed at New Zealand Genomics Limited (Auckland, New Zealand) using the MiSeq platform.

Variant Calling and Recombination Analyses

The completely sequenced genome of *Pseudomonas syringae* pv. *actinidiae* NZ13 was used as a reference for variant calling. A near complete version of this genome was used as a reference in our previous publication and subsequently finished by Templeton et al. (2015), where it is referred to as ICMP 18884 (McCann et al. 2013; Templeton et al. 2015). Variant calling was performed on all *P. syringae* pv. *actinidiae* isolates for which read data was available.

Read data was corrected using the SPAdes correction module and Illumina adapter sequences were removed with Trimmomatic allowing two seed mismatches, with a palindrome and simple clip threshold of 30 and 10, respectively (Bankevich et al. 2012; Bolger et al. 2014). Quality-based trimming was also performed using a sliding window approach to clip the first ten bases of each read as well as leading and trailing bases with quality scores under 20, filtering out all reads with a length under 50 (Bolger et al. 2014). PhiX and other common sequence contaminants were filtered using the Univec Database (Mukherjee et al. 2015).

Reads were mapped to the complete reference genome *Psa* NZ13 with Bowtie2 (default settings for paired-end read-mapping) and duplicates removed with SAMtools (Li 2011; Langmead and Salzberg 2012). Freebayes was used to call variants, retaining variants if they had a minimum alternate allele count of ten reads and fraction of 95% of reads supporting the alternate call (–ploidy 1—min-alternate-fraction 0.95—min-alternate-count 10—report-monomorphic; Garrison and Marth 2012). The average coverage was calculated with SAMtools and used as a guide to exclude overrepresented SNPs (defined here as threefold higher coverage than the average) which may be caused by mapping to repetitive regions. BCFtools filtering and masking was used to generate final reference alignments including SNPs falling within the quality and coverage thresholds described above and excluding SNPs within 3 bp of an insertion or deletion (indel) event or indels separated by 2 or fewer base pairs. Invariant sites with a minimum coverage of ten reads were also retained in the alignment, areas of low (less than ten reads) or no coverage are represented as gaps relative to the reference.

Freebayes variant calling includes indels and multiple nucleotide insertions as well as single nucleotide insertions; however only SNPs were retained for downstream phylogenetic analyses. An implementation of ClonalFrame suitable for use with whole genomes was employed to identify recombinant regions using default settings and a maximum likelihood starting tree generated by RaxML (Stamatakis 2014; Didelot and Wilson 2015). All substitutions occurring within regions identified by ClonalFrameML as being introduced due to recombination were removed from the alignments. The reference alignments were manually curated to exclude substitutions in positions mapping to mobile elements such as plasmids, integrative and conjugative elements, and transposons.

Phylogenetic Analysis

The maximum likelihood phylogenetic tree of 80 *Psa* strains comprising new Chinese isolates and strains reflecting the diversity of all known clades was built with RAxML (version 7.2.8) using a 1,062,844 bp core genome alignment excluding all positions for which one or more genomes lacked coverage of ten reads or higher (Stamatakis 2014). Removal of 3,122 recombinant positions produced a 1,059,722 bp core genome alignment including 2,953 variant sites. A starting tree was used to compute 100 bootstrap replicates (-m GTRGAMMA -p $RANDOM -b $RANDOM -# 100), in turn used to draw bipartitions on the best scoring tree (-m GTRCAT -p $RANDOM -f b). Phylogenies were visualized with FigTree (v.1.4.2 http://tree.bio.ed.ac.uk/software/figtree/). Membership within each phylogenetic clade corresponds to a minimum average nucleotide identity of 99.70%. The average nucleotide identity was determined using a BLAST-based approach in JspeciesWS (ANIb), using a subset of 32 *Psa* genome assemblies spanning all clades (Richter et al. 2016). In order to fully resolve the relationships between more closely related recent outbreak strains, a phylogeny was constructed using only the 62 *Psa*-3 strains. This was determined using a 4,853,155 bp core genome alignment (excluding 258 recombinant SNPs), comprising invariant sites and 1,948 nonrecombinant SNPs and invariant sites. Trees were built with the generalized time-reversible model and gamma distribution of site-specific rate variation (GTR + Γ) and 100 bootstrap

replicates. *Psa* C16 was used to root the tree as this was shown to be the most divergent member of *Psa*-3 when including strains from multiple clades. Nodes shown have minimum bootstrap support values of 50.

Identification of the Pangenome

Genomes were assembled with SPAdes (v.3.6.2) using the filtered, trimmed and corrected reads and the –careful flag to reduce mismatches and indels (Bankevich et al. 2012). Subsequent to quality improvement with Pilon, assemblies were annotated using Prokka (Seemann 2014; Walker et al. 2014). The pangenome of *Psa*-3 was calculated using the ROARY pipeline (v.3.6.1, A.J. Page et al. 2015). The assembly quality (contig number, N50, coverage) of two assemblies (NZ60 and J39) was judged too weak for inclusion in final calculations of the pan genome. Thus orthologs present in 99% (59 out of a total of 60) genomes were considered core in order to account for assembly errors; presence in 57–59, 9–57, and 1–8 were considered soft-core, shell and cloud genomes, respectively. BLASTn-based confirmation (*E* value cutoff 10^{-5}) was used to confirm the identity of predicted virulence or pandemic-clade-restricted genes in genome assemblies. Integrative and conjugative elements (ICEs) were identified using *Psa* NZ13 ICE as a query for BLASTn (*E* value cutoff 10^{-5}). In cases where hits spanned multiple contigs, sequences were concatenated according to the position relative to the query. Identical ICEs share the same color in figure 1*A*.

Pathogenicity Assays

Growth assays were performed using both stab inoculation as in McCann et al. (2013) an initial inoculum of 10^8 cfu/ml and four replicate plants at day 0 and day 6 at all subsequent sampling time points. Bacterial density in inoculated tissue was assessed by serial dilution plating of homogenized tissue. Statistical significance between each treatment at each time point was assessed using two-tailed *t*-tests with uneven variance.

Results

The Phylogeography of *Psa*

The genomes of 50 *P. syringae* pv. *actinidiae* (*Psa*) isolated from symptomatic kiwifruit in China, Korea and New Zealand between 2010 and 2015 were sequenced (see supplementary table S1, Supplemental Material online). Combined with 30 *Psa* genomes from earlier outbreaks and different geographic regions (e.g., Italy and Chile), our samples represent the main *Psa* genotypes from the countries producing 90% of kiwifruit production worldwide. The completed reference genome of *Psa* NZ13 (ICMP 18884, CP011972.1, CP011973.1) comprises a 6,580,291bp chromosome and 74,423 bp plasmid (McCann et al. 2013; Templeton et al. 2015). Read mapping and variant calling with reference to the *Psa* NZ13 chromosome produced a 1,059,722 bp nonrecombinant core genome for all 80 genomes, including 2,953 nonrecombinant SNPs. A maximum likelihood phylogenetic analysis showed the four clades of *Psa* known to cause bleeding canker disease in kiwifruit were represented among the 80 strains (fig. 1). The first clade (*Psa*-1) includes the pathotype strain of *Psa* isolated and described during the first recorded epidemic of bleeding canker disease in Japan (1984–1988). The second clade (*Psa*-2) includes isolates from an epidemic in South Korea (1997–1998), and the third clade (*Psa*-3) includes isolates that define the global pandemic lineage (2008 to present). A fourth clade (*Psa*-5) is represented by a single strain, as no additional sequences or isolates were available (Fujikawa and Sawada 2016). The average between and within-clade pairwise identity is 98.93% and 99.73%, respectively (table 1). All *Psa* isolated from kiwifruit across six different provinces in China are members of the same clade: *Psa*-3. A subset of Chinese strains group with the *Psa* isolated during the global outbreak in Italy, Portugal, New Zealand, and Chile. This subset is referred to as the pandemic lineage of *Psa*-3.

In order to obtain greater resolution of the relationships between the new Chinese and pandemic isolates, we identified the 4,853,413 bp core genome of all 62 strains in *Psa*-3. The core genome includes both variant and invariant sites and excludes regions either unique to or deleted from one or more strains. To minimize the possibility of recombination affecting the reconstruction of evolutionary relationships and genetic distance within *Psa*-3, ClonalFrameML was employed to identify and remove 258 SNPs with a high probability of being introduced by recombination rather than mutation, retaining 1,948 nonrecombinant SNPs. The within-clade ratio of recombination to mutation (R/theta) is reduced in *Psa*-3 ($6.75 \times 10^{-2} \pm 3.24 \times 10^{-5}$) relative to between clade rates ($1.27 \pm 5.16 \times 10^{-4}$), and the mean divergence of imported DNA within *Psa*-3 is $8.54 \times 10^{-3} \pm 5.18 \times 10^{-7}$ compared with $5.68 \times 10^{-3} \pm 1.04 \times 10^{-8}$ between clades. Although recombination has occurred within *Psa*-3, it is less frequent and has introduced fewer polymorphisms relative to mutation: when accounting for polymorphisms present in recombinant regions identified by ClonalFrameML and/or present on transposons, plasmids, and other mobile elements, 7-fold more polymorphisms were introduced by mutation relative to recombination (table 2). Recombination has a more pronounced impact between clades, where substitutions are slightly more likely to have been introduced by recombination than by mutation (table 2).

The Source of Pandemic *Psa*

Data show greater diversity among the Chinese *Psa*-3 population than had been previously identified (fig. 2). Interestingly, clades defining *Psa*-1 and *Psa*-3 exhibit similar levels of

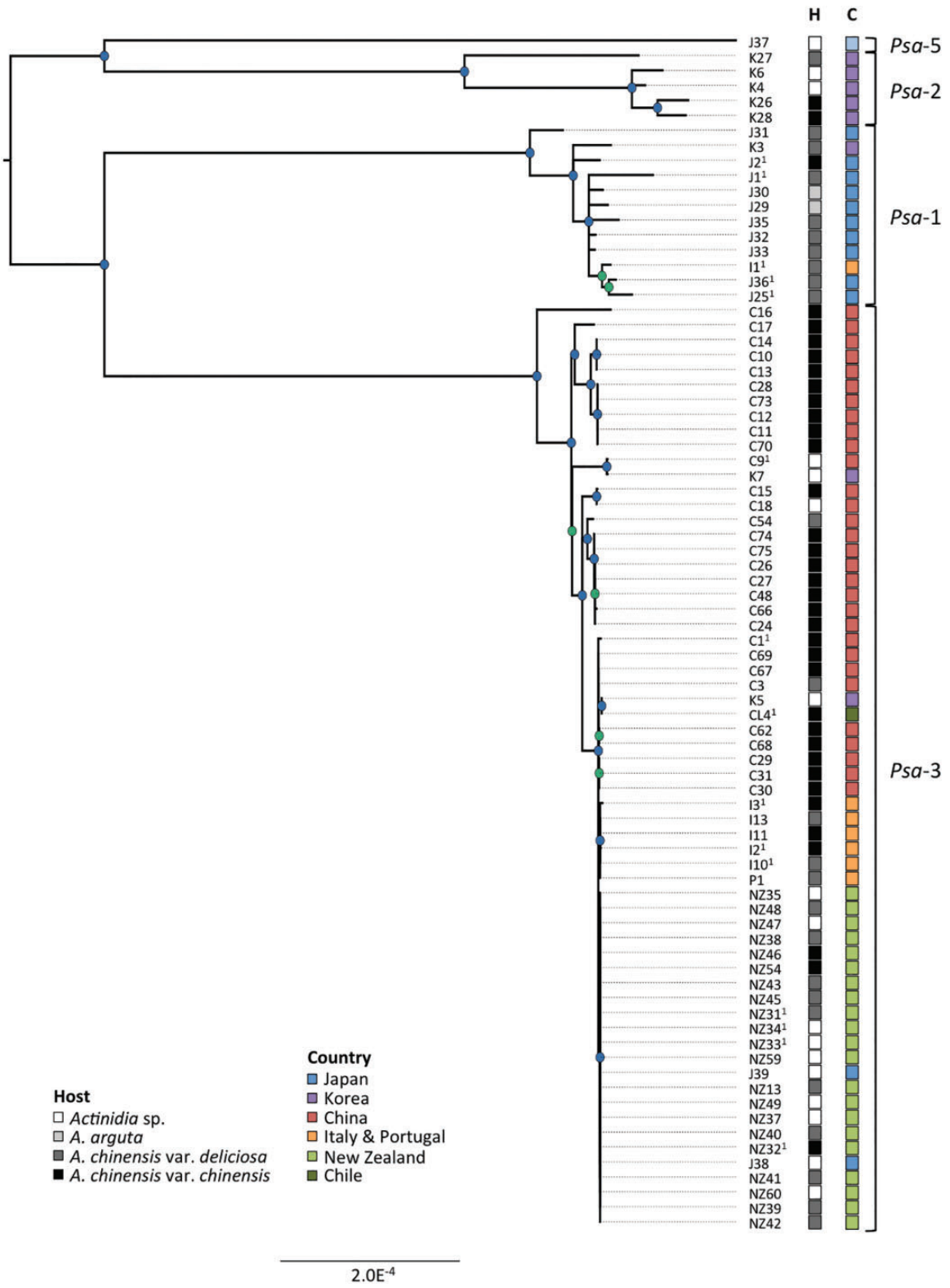


Fig. 1.—Phylogeny of *Psa*. RaxML Maximum likelihood tree based on 1,059,722bp nonrecombinant core genome alignment including 2,953 variant sites. FigTree was used for phylogeny visualization. All nodes displayed have bootstrap support values above 50% (50–75% in green, 76–100% in blue). Host (H) and country (C) of isolation is displayed in the first and second column, respectively. Clade designations are shown at right.

diversity (table 1). These clades share a common ancestor: assuming they are evolving at a similar rate, they may have been present in Japan and China for a similar duration. The strains isolated during the latest pandemic in Italy (I2, I3, I10, I11, I13), Portugal (P1), New Zealand (NZ13, NZ31–35, NZ37–43, NZ45–49, NZ54), Chile (CL4), Japan (J38, J39), and Korea (K5) during the latest kiwifruit canker pandemic cluster with nine Chinese isolates (C1, C3, C29–31, C62, C67–69) (fig. 2). This pandemic lineage exhibits little diversity at the level of the core genome, having undergone clonal expansion only very recently. The New Zealand isolates form a monophyletic group and share a common ancestor, indicating a single transmission event of *Psa* into New Zealand. Two recently isolated Japanese pandemic *Psa*-3 isolated in 2014 group within the New Zealand isolates, suggesting a transmission event between Japan and New Zealand (fig. 2).

Italian and Portuguese pandemic strains also form a separate group, indicative of a single transmission event from China to Italy. China is undoubtedly the source of the strains responsible for the pandemic of kiwifruit canker disease, but the precise region from which the pandemic arose remains unclear. Isolates from four different provinces in Western China (Guizhou, Shaanxi, Sichuan, and Chongqing) are represented among the pandemic lineage, indicating extensive regional transmission within China after emergence of the pandemic. Yet each province harboring pandemic isolates also harbors basally diverging *Psa*-3 isolates (fig. 3).

There is no phylogeographic signal among the more divergent Chinese strains. A group of isolates from Sichuan group together, but this may be due to oversampling rather than an association with this region. The data indicate there was likely extensive regional transmission of *Psa* both prior and subsequent to the emergence of the pandemic in China. Korea harbors representatives of both the divergent and pandemic *Psa*-3 strains. K5 groups with the Chilean *Psa*-3 strain in the pandemic lineage, whereas K7 groups with the more divergent Chinese isolates indicating that a transmission event from strains outside the pandemic lineage may have occurred. This pool of diversity therefore represents a reservoir from which novel strains are likely to emerge in the future.

The reduced level of diversity within the core genome of pandemic *Psa*-3 demonstrates these strains have been circulating for a shorter period of time relative to those responsible for earlier outbreaks in both Japan and Korea. In order to estimate the divergence time of the pandemic lineages we performed linear regression of root-to-tip distances against sampling dates using the RAxML phylogenies determined from the nonrecombinant core genome of *Psa*-3. No temporal signal was identified in the data. There were poor correlations between substitution accumulation and sampling dates, indicating the sampling period is too short for sufficient substitutions to occur. There may also be variation in the substitution rate within even a single clade. Forty-four unique nonrecombinant SNPs were identified among the 21 pandemic *Psa*-3 genomes sampled over 5 years in New Zealand [an average of 2.10 per genome] producing an estimated rate of 8.7×10^{-8} substitutions per site per year. The relatively slow substitution rate and the strong bottleneck effect experienced during infections hinders efforts to reconstruct patterns of transmission, as the global dissemination of a pandemic strain may occur extremely rapidly (Grad and Lipsitch 2014; Biek et al. 2015). The estimated divergence time of *Psa* broadly considered is likely older than the pandemic and epidemic events with which they are associated: the earliest report of disease cause by *Psa*-1 occurred in 1984 and the first report of infection from the latest pandemic was issued in 2008.

Table 1
Average Percent Identity within and between *Psa* Lineages

Lineage	*Psa*-1	*Psa*-2	*Psa*-3	*Psa*-5
Psa-1	99.70			
Psa-2	98.97	99.76		
Psa-3	99.06	98.91	99.73	
Psa-5	98.87	98.96	98.83	ND[a]

[a]Not determined for *Psa*-5 as only a single strain has been sequenced from this lineage. ANIb values determined using representative strains for *Psa*-1 (J31, K3, J2, J1, J30, J29, J35, J32, J33, I1, J36, J25), *Psa*-2 (K27, K6, K4, K26, K28), *Psa*-3 (C16, C17, C10, C11, C70, K7, C15, C54, C74, C69, K5, C62, I13, NZ13), and *Psa*-5 (J37).

Diversification among *Psa*-3 Isolates

2,206 SNPs mapping to the core genome of *Psa*-3 were identified; 258 of these mapped to recombinant regions identified by ClonalFrameML and/or plasmid, prophage, integrative and conjugative elements, transposons, and other mobile genetic elements (table 2 and fig. 2). The highest density of polymorphism was detected in and around the ~100kb integrative and conjugative element (ICE) in *Psa* NZ13 (5.41–5.51Mb,

Table 2
Origin of SNPs in Core Genomes

	Intergenic	Synon	Nonsyn	Extension	Trunc	Total
Psa-3 Mutation	412	583	927	2	24	1,948
Psa-3 Recombination	35	136	85	1	1	258
All *Psa* Mutation	457	1,493	982	3	18	2,953
All *Psa* Recombination	355	2,212	558	2	5	3,132[a]

[a]For ten positions in the alignment, two separate recombination events were predicted to occur.

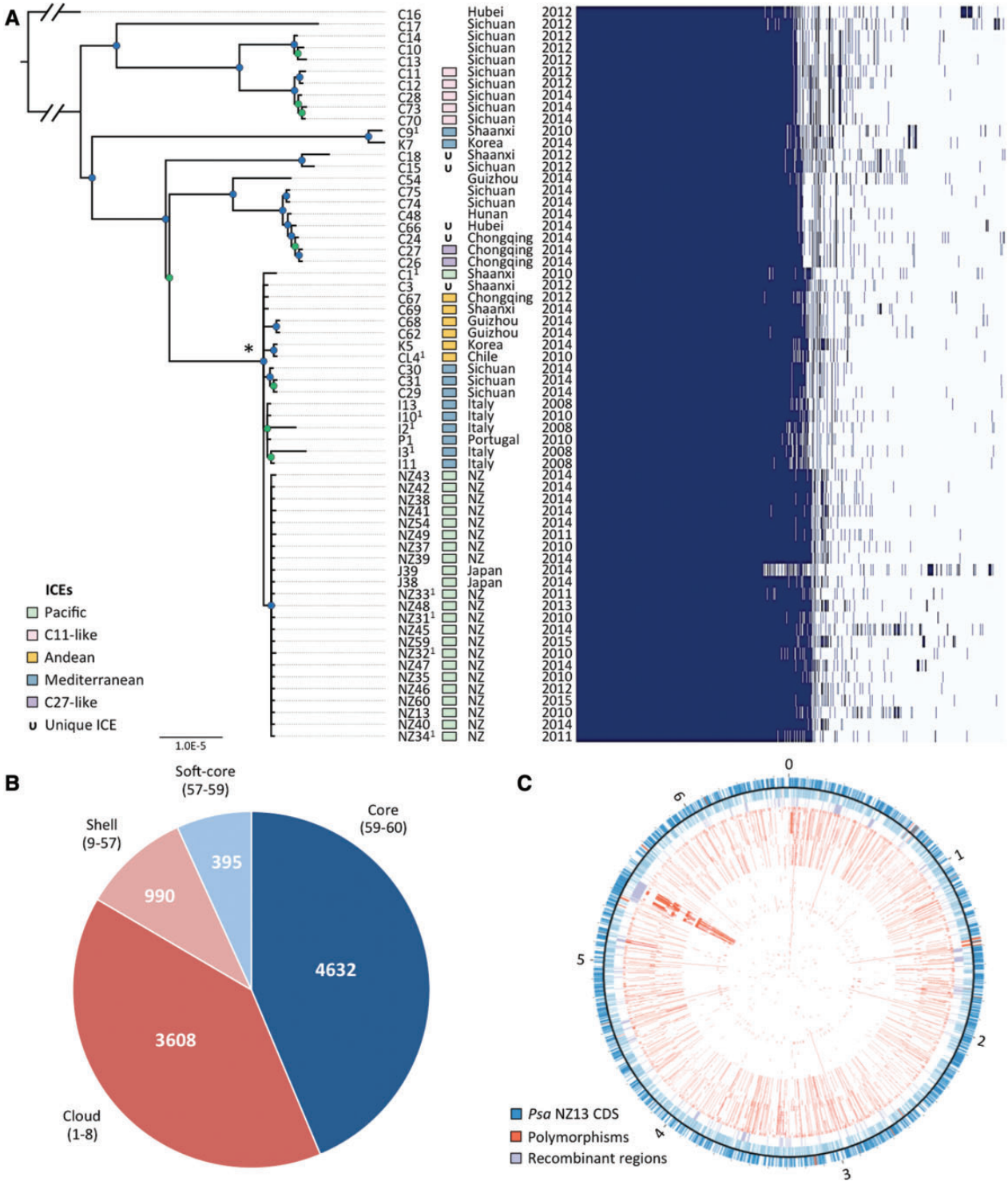


Fig. 2.—Phylogeny of *Psa*-3. (*A*) RaxML Maximum likelihood tree based on 4,853,155bp nonrecombinant core genome alignment including 1,948 variant sites. FigTree was used for phylogeny visualization. All nodes displayed have bootstrap support values above 55% (55–80% in green, 81–100% in blue). The asterisk denotes the pandemic lineage. Integrative and conjugative element (ICE) presence; province (China) or country and year of isolation shown at right, followed by presence/absence matrix of all core and accessory genes. (*B*) The core and flexible genome of *Psa*-3. Parenthetical numbers represent the number of strains in the core, soft-core, shell, and cloud genomes. (*C*) Polymorphisms and recombinant regions mapped onto *Psa* NZ13 reference genome using CIRCOS (Krzywinski et al. 2009). *Psa* NZ13 CDS are displayed in the first and second ring (blue), with annotated Type 3 secretion system and effectors highlighted (red). Inner rings display polymorphisms in *Psa*-3 genomes ordered from most to least divergent relative to *Psa* NZ13. The most polymorphic region corresponds to the location of the integrative and conjugative element (ICE) in *Psa* NZ13.

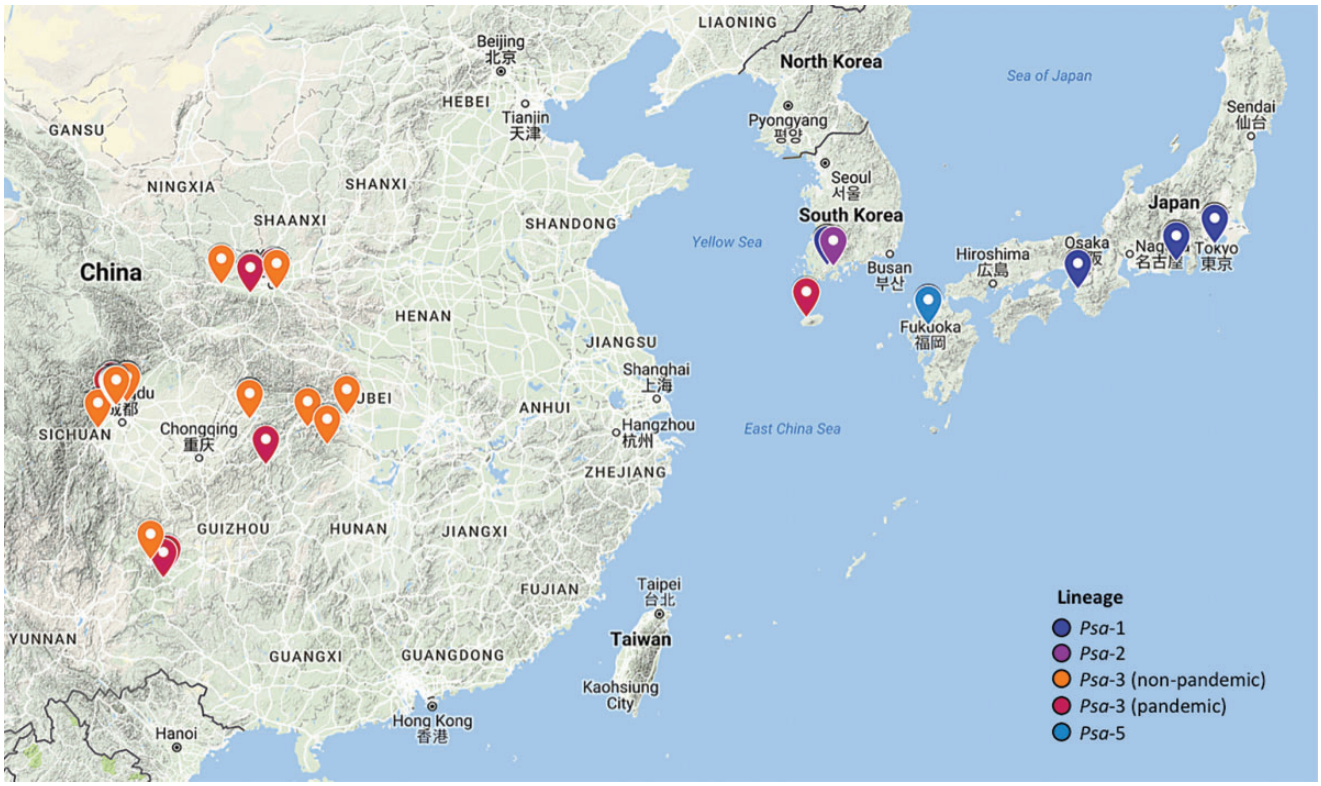


Fig. 3.—*Psa* isolation locations in East Asia. Markers correspond to location of isolation. Color corresponds to the isolate membership in each clade according to the phylogeny displayed in fig. 1. Only isolates for which the specific location of isolation is known are displayed here.

fig. 2). Of the 1,948 SNPs mapping to the nonrecombinant nonmobile core genome, 58.1% (1,132) are strain specific. Most strain-specific SNPs are found in the two most divergent members of the *Psa*-3 clade: *Psa* C16 and C17, with 736 and 157 strain-specific SNPs, respectively. The remaining isolates have an average of 4.0 strain-specific SNPs, ranging from 0 to 44 SNPs per strains. There are 816 SNPs shared between two or more *Psa*-3 strains. The pandemic clade differs from the more divergent Chinese strains by 72 shared SNPs. Within the pandemic lineage there are 125 strain-specific SNPs, an average of 3.1 unique SNPs per strain (ranging from 0 to 27 SNPs) and an additional 29 SNPs shared among pandemic strains. Protein-coding sequence accounts for 88.4% of the nonrecombinant, gap-free core genome of this clade. We observed that 78.9% (1,536/1,948) of mutations occurred in protein coding sequence, significantly different from the expectation (1,722/1,948) in the absence of selection (Pearson's χ^2 test: $P < 0.0001$, $\chi^2 = 173.46$). This suggests there is selection against mutations occurring in protein coding sequences. Of the 953 substitutions introduced by mutation in *Psa*-3, 927 resulted in amino acid substitutions, two resulted in extensions and 24 resulted in premature truncations.

Multiple synonymous and nonsynonymous mutations were identified in 269 genes (see supplementary table S3, Supplemental Material online). The accumulation of multiple independent mutations in the same gene may be a function of gene length, mutational hotspots or directional selection. A range of hypothetical proteins, membrane proteins, transporters, porins, and type III and IV secretion system proteins acquired between two and seven mutations. The fitness impact of these mutations—and the 38 amino-acid changing mutations in the ancestor of the pandemic lineage—is unknown. It is possible these patterns are the outcome of selective pressures imposed during bacterial residence within a similar host niche, but considering most of these strains have only recently diverged, it is likely neutral and deleterious mutations have not yet been purged by selection (Kryazhimskiy and Plotkin 2008).

Two substitutions are shared exclusively by the European pandemic strains (AKT28710.1 G1150A and AKT33438.1 T651C) and one silent substitution in a gene encoding an acyltransferase superfamily protein (AKT31915.1 C273T) is shared among the European pandemic and six of nine Chinese pandemic strains (C3, C29–31, C67, C69). As these six Chinese pandemic strains were isolated from Shaanxi, Sichuan, and Chongqing, they do not provide any insight into the precise geographic origins of the European pandemic *Psa*-3; transmission from China to Italy is likely concomitant with dissemination of the pandemic lineage across China. Six

Table 3
SNPs Shared between All Pandemic NZ and Japanese Isolates

Protein ID (NZ13)	Product	Codon
AKT31947.1	ion channel protein Tsx	71 (silent)
AKT32845.1	bifunctional glutamine-synthetase adenylyltransferase	W977R
AKT30494.1	chromosome segregation protein SMC	H694Q
AKT29651.1	cytidylate kinase	V173L
AKT32264.1	peptidase PmbA	M418K
Intergenic		362,522 (G→T)

conserved and diagnostic polymorphisms are present in the pandemic New Zealand and Japanese isolates (table 3). One of these is a silent substitution in an ion channel protein (AKT31947.1 A213G), another is an intergenic (T→G) mutation at position 362,522 of the reference *Psa* NZ13 chromosome and the remaining four are nonsynonymous substitutions in an adenylyltransferase (AKT32845.1, W977R); chromosome segregation protein (AKT30494.1, H694Q); cytidylate kinase (AKT29651.1, V173L), and peptidase protein (AKT32264.1, M418K).

The type III secretion system is known to be required for virulence in *P. syringae*. A 44,620 bp deletion event in *Psa* C17 resulted in the loss of 42 genes encoding the structural apparatus and conserved type III secreted effectors in *Psa* C17. This strain is highly compromised in its ability to grow in *A. chinensis* var. *deliciosa* "Hayward", attaining 1.2×10^7 cfu/g 3 days post inoculation (dpi) and declining to 8.8×10^4 cfu/g at 14 dpi (see supplementary fig. S1, Supplemental Material online). This is a marked reduction compared with *Psa* NZ13, which attains 3.0×10^9 and 4.2×10^7 cfu/g 3 and 14 dpi, respectively. *Psa* C17 nevertheless multiplies between day 0 and day 3, indicating that even in the absence of type III-mediated host defense disruption, *Psa* may still proliferate in host tissues. The loss of the TTSS does not inhibit the growth of *Psa* C17 as strongly in the more susceptible *A. chinensis* var. *chinensis* "Hort16A" cultivar.

Two potentially significant deletion events occurred in the ancestor of the pandemic lineage: a frameshift caused by a mutation and single base pair deletion in a glucan succinyltransferase (*opgC*) and a 6,456 bp deletion in the *wss* operon (see supplementary fig. S2, Supplemental Material online). Osmoregulated periplasmic glucans (OPGs, in particular *opgG* and *opgH*) are required for motility, biofilm formation and virulence in various plant pathogenic bacteria and fungi (Page et al. 2001; Klosterman et al. 2011; Wu et al. 2014). Homologs of *opgGH* remain intact in the pandemic lineage, yet the premature stop mutation in *opgC* likely results in the loss of glucan succinylation. The soft-rot pathogen *Dickeya dadantii* expresses OpgC in high osmolarity conditions, resulting in the substitution of OPGs by O-succinyl residues (Bontemps-Gallo et al. 2016). The most striking difference between the pandemic lineage and more divergent Chinese *Psa*-3 strains is the deletion of multiple genes involved in cellulose production and acetylation (see supplementary fig. S2, Supplemental Material online) (Spiers et al. 2002). In *P. fluorescens* SBW25 deletion of the cellulose-encoding Wss operon significantly compromises ability to colonize plant surfaces and in particular the phyllosphere of sugar beet (*Beta vulgaris*) seedlings (Gal et al. 2003). It is possible that loss of this locus aids movement through the vascular system and/or dissemination between plants, by limiting capacity for surface colonization.

Dynamic Genome Evolution of *Psa*-3

Despite the high similarity within the core genome, extensive variation is evident in the pangenome of *Psa*-3 (fig. 2). The core (4,632 genes in 99–100% of strains, and 395 genes in 95–99% of strains, or 57–60 genomes) comprises 50.5% of the total pangenome (9,625 genes). 990 genes are present in 15–95% of strains (9–57 genomes), the so-called "shell genes" (fig. 2). The flexible genome is comprised of the "shell" and "cloud" genes; the latter describes genes present in up to 15% of strains (one to eight genomes in this case). Cloud genes contribute most to the flexible genome: 3,608 genes are present in one to eight strains. This is a considerable amount of variation in a pathogen described as clonal and monomorphic, particular since our analysis includes only strains from one clade of a monophyletic pathovar group (McCann et al. 2013). A comparative analysis of 64 genomes spanning the entire *P. syringae* species complex estimated the core genome to be ~48% of an average *P. syringae* genome, and 20.6% of the total pangenome (Nowell et al. 2016). Estimated core genome sizes for other bacterial plant pathogens range from ~17% of the entire *Ralstonia solanacearum* pangenome ($n = 19$ genomes, (Ailloud et al. 2015)) to ~59% of the *Erwinia amylovora* pangenome ($n = 12$ genomes, (Mann et al. 2013)). These estimates are sensitive to sequencing quality, assembly and annotation errors, vary according to phylogenetic breadth and depth of sampling, and will certainly change with increasing sequence availability. For example, the omission of two low-quality assemblies (J39 and NZ60) results in a core and soft-core genome differing by 293 genes and a reduction of the cloud by 342.

Despite a relatively low rate of mutation and limited within-clade homologous recombination, the amount of heterologous recombination demonstrates that the genomes of these pathogens are highly labile. Mobile genetic elements like bacteriophage, transposons, and integrases make a dramatic contribution to the flexible genome. Integrative and conjugative elements (ICEs) are mobile and are involved in the transfer of copper resistance in *Psa* (Colombi et al.

2017). Prodigious capacity for lateral gene transfer creates extreme discordance between ICE type, host phylogeny and host geography making these regions unsuitable markers of host evolution and origin. Three divergent ICEs have been previously described from the global pandemic lineage (McCann et al. 2013). Within *Psa*-3, ICEs were found in 53 of 62 isolates (nine of the divergent Chinese isolates were devoid of any such element) (fig. 2). No phylogeographic signal is evident. For example, strains from Sichuan, Shaanxi, Korea, Italy, and Portugal share an identical ICE. Even within a single Chinese province, multiple ICEs exist (Shaanxi and Sichuan isolates harbor four and three different ICEs, respectively). Moreover, ICE host range is not limited to *Psa* alone: the ICE present in *Psa* C1 and all New Zealand *Psa*-3 exists in essentially identical form in a strain of *P. syringae* pv. *avellenae* CRAPAV013 isolated from hazelnut in 1991 in Latina, Italy (it exhibits 98% pairwise identity, differing from the New Zealand ICE by a transposon, 66 bp deletion and a mere 6 SNPs).

Discussion

We have described an endemic population of *Psa* infecting cultivated kiwifruit in China. All *Psa* isolated within China are members of *Psa*-3. The pandemic strains isolated in Italy, Portugal, Chile, and New Zealand are derived from a single lineage within this clade along with a subset of Chinese isolates, indicating that the pandemic ultimately emerged from the diverse Chinese population of clade 3 strains. Italian pandemic strains share a SNP with six of nine Chinese pandemic *Psa* strains, indicating there was likely a direct transmission event from China to Italy prior to 2008. The New Zealand isolates share six clade-defining mutations, indicating that a separate and single transmission event was responsible for the outbreak of disease there. Identification of the transmission pathway introducing *Psa* into New Zealand is dependent on obtaining a sample of *Psa* sharing some or all of the mutations characteristic of New Zealand *Psa* from either the overseas source population or from infected plant material arriving into New Zealand from an overseas location. Japanese pandemic strains cluster with New Zealand strains, and share all six clade-defining mutations. This suggests that pandemic *Psa*-3 was either introduced into Japan via New Zealand, or from the same as-yet unknown region in China from which transmission to New Zealand occurred. *Psa*-3 was first identified as causing disease in four prefectures across Japan in April 2014 (Sawada et al. 2015). Japan imported pollen and plant material from both China and New Zealand prior and subsequent to *Psa*-3 detection in both those countries, though the amount of pollen imported from New Zealand in 2012 (349 kg) and 2013 (190 kg) far outweighed the amount imported from China (1 kg in both 2012 and 2013) (Japanese Ministry of Agriculture, Forestry and Fisheries 2016). Although resolving the origin of transmission events is of much interest the rate of evolution places a lower bound on the ability of genomic epidemiology to deliver useful data. Divergence dating methods have been successfully applied to infer epidemiological processes for numerous bacterial mammalian pathogens, though our present estimate of *Psa*'s genome-wide evolutionary rate places it among the slowest evolving bacteria included in a recent survey of 36 whole genome data sets (Duchêne et al. 2016).

Our phylogeographic study of a single clade giving rise to a pandemic in *P. syringae* has revealed far greater diversity than was previously appreciated. Extensive diversity between *Psa* isolates collected from *Actinidia* spp. was observed in the same province. The amount of diversity present within *Psa*-3 indicates this population was present and circulating in China before the pandemic began. The emergence of the pandemic lineage moreover has not resulted in the displacement of more ancestral strains: both pandemic and divergent *Psa*-3 were isolated from four out of six provinces.

Strains from three different clades have been isolated in both Korea and Japan, whereas China harbors strains from only a single clade. The most basal clades of canker-causing *Psa* are comprised of Korean strains isolated between 1997 and 2014 (*Psa*-2) and a member of the recently identified *Psa*-5. One early isolate (*Psa* K3, 1997) groups with the Japanese isolates in *Psa*-1; *Psa* K7 (2014) groups with the more diverse Chinese isolates in *Psa*-3, and *Psa* K5 (2014) is a member of the *Psa*-3 pandemic lineage. Korea therefore harbors a more diverse population of *Psa* than China, with strains from three distinct clades of *Psa* (1, 2, pandemic and nonpandemic 3). The novel group of *Psa* recently identified in Japan (*Psa*-5) appears to share an ancestor with the Korean *Psa*-2 strains. With the recent dissemination of pandemic *Psa*-3 and the historical presence of *Psa*-1 in Japan, the Japanese population of *Psa* is comprised of three distinct clades of *Psa* (1, 5 and pandemic *Psa*-3). The coexistence of multiple clades in both Japan and Korea suggests that the source population of all *Psa* resides there rather than in China. The potential transmission of a nonpandemic *Psa*-3 strain from China to Korea and the identification of a new clade in Japan supports our earlier assertion that variants will continue to emerge to cause local epidemics and global pandemics in the future (McCann et al. 2013).

The apparent presence of just a single clade in China may reflect the use of clonal monocultures of popular cultivars. Most *Psa*-3 isolates from China were procured from *A. chinensis* var. *chinensis* "Hongyang" and *A. chinensis* var. *deliciosa* "Hayward." Clonal monocultures are typical in Korea and Japan as well however, and China has the highest diversity of both wild and cultivated kiwifruit: cultivars sampled are few generations removed from selections made from the wild across an enormous geographic area (Shaanxi, Guangxi, Fujian, Hubei, Henan, Hunan, Jiangsu).

Considering that the divergence time of this monophyletic pathovar predates the commercialization of kiwifruit by hundreds if not thousands of years, *Psa* is likely associated with a nondomesticated host(s) in the wild (Fujikawa and Sawada 2016; McCann et al. 2013). Both *A. chinensis* var. *deliciosa* or *A. chinensis* var. *chinensis* are found in natural ecosystems and have overlapping habitat ranges with cultivated kiwifruit in many areas. However, despite isolating 746 *Pseudomonas* strains from both wild and cultivated kiwifruit during this sampling program, we did not identify *Psa* among any of the 188 *Pseudomonas* spp. isolated from 98 wild *A. chinensis* var. *deliciosa* or *A. chinensis* var. *chinensis* sampled across China (see supplementary table S2, Supplemental Material online). Very few *Actinidia* spp. have ranges extending to South Korea and Japan—those that do include *A. arguta*, *A. kolomikta*, *A. polygama*, and *A. rufa*. *A. arguta* is broadly distributed across both Korea and Japan. Early work by Ushiyama et al. (1992) found that *Psa* could be isolated from symptomatic *A. arguta* plants in Japan. The possibility that this wild relative of kiwifruit harbors diverse strains of *Psa* that may emerge to cause future outbreaks is currently under investigation. Alternately, a host shift from another domesticated crop may have occurred after the expansion of kiwifruit cultivation.

Numerous epidemiological studies of human pathogens have demonstrated environmental or zoonotic origins, but there are fewer studies of bacterial plant pathogens (Almeida and Nunney 2015; Andam et al. 2016; Cauchemez et al. 2016; Clarke et al. 2015; Mather et al. 2013; McCann et al. 2013; Monteil et al. 2016; Quibod et al. 2016; Schwartz et al. 2015; Shapiro et al. 2016; Stukenbrock and Bataillon 2012; Vinatzer et al. 2014; Wagner et al. 2014; Arango Isaza et al. 2016; Depotter et al. 2017; Islam et al. 2016; Menardo et al. 2016). Where ecological and genetic factors restrict pathogens to a small number of plant hosts greater progress has been made, but for facultative pathogens such as *P. syringae* that colonize multiple hosts and are widely distributed among both plant and nonplant habitats, the environmental reservoirs of disease and factors affecting their evolutionary emergence are difficult to unravel (Singh et al. 2011; Monteil et al. 2016).

The emergence of *Psa* over the last three decades—concomitant with domestication of kiwifruit—offers a rare opportunity to understand the relationship between wild populations of both plants and microbes and the ecological and evolutionary factors driving the origins of disease, including the role of agriculture. It is now possible to exclude China as the native home to the source population, but the precise location remains unclear. Nonetheless, it is likely, given the extent of diversity among *Psa* isolates and the time-line to domestication, that ancestral populations exist in nonagricultural plant communities. Attention now turns to Korea and Japan and the interplay between genetic and ecological factors that have shaped *Psa* evolution.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

We gratefully acknowledge the assistance of the following guides, teachers, and graduate assistants who helped us identify sample locations in China: Junjie Gong, Yancang Wang, Shengju Zhang, Zupeng Wang, Yangtao Guo, Meiyan Chen, Kuntong Li, Moucai Wang, Jiaming He, Yonglin Zhao, Zhongshu Yu, Yan Lv, Mingfei Yao, Shihua Pu, Tingwen Huang, Qiuling Hu, Caizhi He, and Jiaqing Peng. Derk Wachsmuth at Max Planck Institute for computing cluster support. James Connell for assistance with biosecurity regulations. Members of the Rainey and Huang labs for discussion. Joel Vanneste for contributing strains. This work was funded by grants from the New Zealand Ministry for Business, Innovation and Employment (C11X1205), Canada Natural Sciences and Engineering Research Council (NSERC PDF), Chinese Academy of Sciences President's International Fellowship Initiative (Grant no. 2015PB063), China Scholarship Council (Grant no. 201504910013), National Natural Science Foundation of China (Grant no. 31572092), Science and Technology Service Network Initiative Foundation of The Chinese Academy of Sciences (Grant no. KFJ-EW-STS-076), Protection and Utilization of Crop Germplasm Resources Foundation of Ministry of Agriculture (Grant no. 2015NWB027). We thank three anonymous reviewers for their comments and suggestions.

Literature Cited

Abelleira A, et al. 2011. First report of bacterial canker of kiwifruit caused by *Pseudomonas syringae* pv. *actinidiae* in Spain. Plant Dis. 95:1583–1583.

Ailloud F, et al. 2015. Comparative genomic analysis of *Ralstonia solanacearum* reveals candidate genes for host specificity. BMC Genomics 16:270.

Almeida RPP, Nunney L. 2015. How do plant diseases caused by *Xylella fastidiosa* emerge?. Plant Dis. 99:1457–1467.

Andam CP, Worby CJ, Chang Q, Campana MG. 2016. Microbial genomics of ancient plagues and outbreaks. Trends Microbiol. 24(12):978–990.

Arango Isaza RE, et al. 2016. Combating a global threat to a clonal crop: banana black sigatoka pathogen *Pseudocercospora fijiensis* (synonym *Mycosphaerella fijiensis*) genomes reveal clues for disease control. PLoS Genet. 12(8):e1005876–e1005836.

Balestra GM, Renzi M, Mazzaglia A. 2010. First report of bacterial canker of *Actinidia deliciosa* caused by *Pseudomonas syringae* pv. *actinidiae* in Portugal. New Dis Rep. 22:10.

Bankevich A, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19:455–477.

Biek R, Pybus OG, Lloyd-Smith JO, Didelot X. 2015. Measurably evolving pathogens in the genomic era. Trends Ecol Evol. 30:306–313.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:1–7.

Bontemps-Gallo S, et al. 2016. The *opgC* gene is required for OPGs succinylation and is osmoregulated through RcsCDB and EnvZ/OmpR in the phytopathogen *Dickeya dadantii*. Sci Rep. 6:19619.
Butler MI, 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.
Cauchemez S, et al. 2016. Unraveling the drivers of MERS-CoV transmission. Proc Natl Acad Sci U S A. 113:9081–9086.
Clarke CR, et al. 2015. Genome-enabled phylogeographic investigation of the quarantine pathogen *Ralstonia solanacearum* race 3 biovar 2 and screening for sources of resistance against its core effectors. Phytopathology 105:597–607.
Colombi E, et al. 2017. Evolution of copper resistance in the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae* through acquisition of integrative conjugative elements and plasmids. Env Microbiol. doi: 10.1111/1462-2920.13662.
Cruzat C. 2014. The kiwifruit in Chile and in the world. Rev Brasil Fruticult 36(1):112–123.
Depotter J, Seidl M, van den Berg G, Thomma B, Wood T. 2017. A previously established divergent lineage of the hybrid fungal pathogen *Verticillium longisporum* emerges as stem striping pathogen in British oilseed rape. bioRxiv doi: 10.1101/102541.
Didelot X, Wilson DJ. 2015. ClonalFrameML: efficient inference of recombination in whole bacterial genomes. PLoS Comput Biol. 11:e1004041–e1004018.
Duchêne S, et al. 2016. Genome-scale rates of evolutionary change in bacteria. Microbial Genomics 2(11). doi: 10.1099/mgen.0.000094.
European and Mediterranean Plant Protection Organization. 2011. EPPO Reporting Service 3:4. http://archives.eppo.int/EPPOReporting/2011/Rse-1103.pdf.
Everett KR, et al. 2011. First report of *Pseudomonas syringae* pv. *actinidiae* causing kiwifruit bacterial canker in New Zealand. Aust Plant Dis Notes 6:67–71.
Fang Y, Xiaoxiang Z, Tao WY. 1990. Preliminary studies on kiwifruit disease in Hunan province. Sichuan Fruit Sci Technol. 18:28–29.
Ferguson AR. 2011. Kiwifruit: evolution of a crop. Acta Hort. 913:31–42.
Ferguson AR. 2015. Kiwifruit in the world—2014. Acta Hort. 1096:33–46.
Ferguson AR, Huang H. 2007. Genetic resources of kiwifruit: domestication and breeding. Hort Rev. 33:1–121.
Fujikawa T, Sawada H. 2016. Genome analysis of the kiwifruit canker pathogen *Pseudomonas syringae* pv. *actinidiae* biovar 5. Sci Rep. 6:21399–21311.
Gal M, Preston GM, Massey RC, Spiers AJ, Rainey PB. 2003. Genes encoding a cellulosic polymer contribute toward the ecological success of *Pseudomonas fluorescens* SBW25 on plant surfaces. Mol Ecol. 12:3109–3121.
Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. arXiv:1207.3907v2.
Grad YH, Lipsitch M. 2014. Epidemiologic data and pathogen genome sequences: a powerful synergy for public health. Genome Biol. 15:538.
Huang H, Wang Y, Zhang Z, Jiang Z, Wang S. 2004. *Actinidia* germplasm resources and kiwifruit industry in China. HortScience 39:1165–1172.
Huang H. 2014. Domestication and Commercialization of *Actinidia*. In: The genus Actinidia: a world monograph. Wang H, editor. Beijing: Science Press. p. 189–202.
Islam MT, et al. 2016. Emergence of wheat blast in Bangladesh was caused by a South American lineage of *Magnaporthe oryzae*. BMC Biol. 14(1):84.
Japanese Ministry of Agriculture, Forestry and Fisheries. 2016. Pest Risk Analysis Report on *Pseudomonas syringae* pv. *actinidiae* biovar 3. Yokohama Plant Protection Station, Research Division.
Klosterman SJ, et al. 2011. Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens. PLoS Pathog. 7:e1002137–e1002119.
Koh YJ, Cha JB, Chung JH, Lee HD. 1994. Outbreak and spread of bacterial canker in kiwifruit. Korean J Plant Pathol. 10:68–72.
Koh YJ, Jung JS, HUR JS. 2002. Current status of occurrence of major diseases on kiwifruits and their control in Korea. Acta Hort. 610:437–443.
Koh YJ, et al. 2012. Occurrence of a new type of *Pseudomonas syringae* pv. *actinidiae* strain of bacterial canker on kiwifruit in Korea. Plant Pathol J. 28:423–427.
Kryazhimskiy S, Plotkin JB. 2008. The population genetics of dN/dS. PLoS Genet. 4(12):e1000304.
Krzywinski M, Schein J, Birol I, Connors J. 2009. Circos: an information aesthetic for comparative genomics. Genome Res. 19:1639–1645.
Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359.
Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27:2987–2993.
Mann RA, et al. 2013. Comparative genomics of 12 strains of *Erwinia amylovora* identifies a pan-genome with a large conserved core. PLoS One 8(2):e55644.
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.
Mather AE, Reid S, Maskell DJ, Parkhill J. 2013. Distinguishable epidemics of multidrug-resistant *Salmonella* Typhimurium DT104 in different hosts. Science 341:1514–1517.
Mazzaglia A, 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.
McCann HC, 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:e1003503.
Menardo F, et al. 2016. Hybridization of powdery mildew strains gives rise to pathogens on novel agricultural crop species. Nat Genet. 48(2):201–205.
Monteil CL, et al. 2016. Population-genomic insights into emergence, crop-adaptation, and dissemination of *Pseudomonas syringae* pathogens. Microbial Genomics 2(10). doi: 10.1099/mgen.0.000089.
Mukherjee S, Huntemann M, Ivanova N, Kyrpides NC, Pati A. 2015. Large-scale contamination of microbial isolate genomes by Illumina PhiX control. Standards Genomic Sci. 10:18.
Nowell RW, Laue BE, Sharp PM, Green S. 2016. Comparative genomics reveals genes significantly associated with woody hosts in the plant pathogen *Pseudomonas syringae*. Mol Plant Pathol. 17(9):1409–1424.
Page AJ, et al. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693.
Page F, et al. 2001. Osmoregulated periplasmic glucan synthesis is required for *Erwinia chrysanthemi* pathogenicity. J Bacteriol. 183:3134–3141.
Quibod IL, et al. 2016. Effector diversification contributes to *Xanthomonas oryzae* pv. *oryzae* phenotypic adaptation in a semi-isolated environment. Sci Rep. 6:34137.
Richter M, Rosselo-Mora R, Glockner FO, Peplies J. 2016. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32:929–931.
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.

Sawada H, et al. 2015. Characterization of biovar 3 strains of *Pseudomonas syringae* pv. *actinidiae* isolated in Japan. Ann Phytopathol Soc Jpn. 81:111–126.
Schwartz AR, et al. 2015. Phylogenomics of *Xanthomonas* field strains infecting pepper and tomato reveals diversity in effector repertoires and identifies determinants of host specificity. Front Microbiol. 6:208–217.
Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.
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 bactericides. Annals of the Phytopathol Soc Jpn. 55:427–436.
Shapiro LR, et al. 2016. Horizontal gene acquisitions, mobile element proliferation, and genome decay in the host-restricted plant pathogen *Erwinia tracheiphila*. Genome Biol Evol. 8:649–664.
Shim KK, Ha YM. 1999. Kiwifruit production and research in Korea. Acta Hort. 498:127–132.
Singh RP, et al. 2011. The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. Annu Rev Phytopathol. 49:465–481.
Spiers AJ, Kahn SG, Bohannon J, Travisano M, Rainey PB. 2002. Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. I. Genetic and phenotypic bases of wrinkly spreader fitness. Genetics 161:33–46.
Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.
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.
Takikawa Y, Serizawa S, Ichikawa T, Tsuyumu S, Goto M. 1989. *Pseudomonas syringae* pv. *actinidiae* pv. nov.: The causal bacterium of canker of kiwifruit in Japan. Jpn J Phytopathol. 55:437–444.
Templeton MD, Warren BA, Andersen MT, Rikkerink EHA, Fineran PC. 2015. Complete DNA sequence of *Pseudomonas syringae* pv. *actinidiae*, the causal agent of kiwifruit canker disease. Genome Announc. 3:e01054–e01015.
Testolin R, Ferguson AR. 2009. Kiwifruit (*Actinidia* spp.) production and marketing in Italy. N Z J Crop Hort. Sci. 37:1–32.
Ushiyama K, et al. 1992. Matatabi as a source of transmission of kiwifruit disease. Japan Plant Pathol Report 58:426–430.
Vanneste JL, et al. 2011. First report of *Pseudomonas syringae* pv. *actinidiae*, the causal agent of bacterial canker of kiwifruit in France. Plant Dis. 95:1311–1311.
Vinatzer BA, Monteil CL, Clarke CR. 2014. Harnessing population genomics to understand how bacterial pathogens emerge, adapt to crop hosts, and disseminate. Annu Rev Phytopathol. 52:19–43.
Wagner DM, et al. 2014. *Yersinia pestis* and the Plague of Justinian 541–543 AD: a genomic analysis. Lancet Infect Dis. 14:319–326.
Walker BJ, et al. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963–e112914.
Wu X, et al. 2014. Deciphering the components that coordinately regulate virulence factors of the soft rot pathogen *Dickeya dadantii*. Mol Plant Microbe Interact. 27:1119–1131.
Zhao ZB, Gao XN, Huang QL, Huang LL, Qin HQ. 2013. Identification and characterization of the causal agent of bacterial canker of kiwifruit in the Shaanxi province of China. J Plant Pathol. 95:155–162.

Associate editor: Rachel O'Neill