

Microbial Genomics

Population-genomic insights into emergence, crop-adaptation, and dissemination of *Pseudomonas syringae* pathogens

--Manuscript Draft--

Manuscript Number:	MGEN-D-16-00057R1
Full Title:	Population-genomic insights into emergence, crop-adaptation, and dissemination of <i>Pseudomonas syringae</i> pathogens
Article Type:	Research Paper
Section/Category:	Microbial evolution and epidemiology: Population Genomics
Corresponding Author:	Boris A Vinatzer, PhD Virginia Tech Blacksburg, UNITED STATES
First Author:	Caroline L Monteil, PhD
Order of Authors:	Caroline L Monteil, PhD Koji Yahara, PhD David J Studholme, PhD Leonardos Mageiros, PhD Guillaume Méric, PhD Bryan Swingle, PhD Cindy E Morris, PhD Boris A Vinatzer, PhD Samuel K Sheppard, PhD
Abstract:	Many bacterial pathogens are well characterized but, in some cases, relatively little is known about the populations from which they emerged. This limits understanding of the molecular mechanisms underlying disease. The crop pathogen <i>Pseudomonas syringae</i> sensu lato has been widely isolated from the environment, including wild plants and components of the water cycle, and causes disease in several economically important crops. Here, we compared genome sequences of 45 <i>P. syringae</i> crop pathogen outbreak strains with 69 closely related environmental isolates. Phylogenetic reconstruction revealed that crop pathogens emerged many times independently from environmental populations. Unexpectedly, differences in gene content between environmental populations and outbreak strains were minimal with most virulence genes present in both. However, a genome-wide association study identified a small number of genes, including the type III effector genes hopQ1 and hopD1, to be associated with crop pathogens, but not with environmental populations, suggesting that this small group of genes may play an important role in crop disease emergence. Intriguingly, genome-wide analysis of homologous recombination revealed that the locus Psyr 0346, predicted to encode a protein that confers antibiotic resistance, has been frequently exchanged among lineages and thus may contribute to pathogen fitness. Finally, we found that isolates from diseased crops and from components of the water cycle, collected during the same crop disease epidemic, form a single population. This provides the strongest evidence yet that precipitation and irrigation water are an overlooked inoculum source for disease epidemics caused by <i>P. syringae</i> .



MICROBIAL GENOMICS

Research paper

1 Population-genomic insights into emergence, crop- 2 adaptation, and dissemination of *Pseudomonas* 3 *syringae* pathogens

4

5 Caroline L. Monteil^{1,2,3,4}, Koji Yahara^{1,5}, David J. Studholme⁶, Leonards Mageiros¹, Guillaume Méric⁸, Bryan
6 Swingle⁷, Cindy E. Morris³, Boris A. Vinatzer^{2,*†}, Samuel K. Sheppard^{8,9,*†}

7

8 ¹ Institute of Life Science, College of Medicine, Swansea University, Swansea, United Kingdom

9 ² Department of Plant Pathology, Physiology, and Weed Science, Virginia Tech, Blacksburg, VA, USA

10 ³ INRA, UR0407 Pathologie Végétale, Montfavet cedex, France

11 ⁴ CEA/CNRS/Aix-Marseille Université, UMR 7265 Biologie Végétale et Microbiologie Environnementales,
12 Laboratoire de Bioénergétique Cellulaire, Institut de Biosciences et Biotechnology Aix-Marseille (BIAM), 13108
13 Saint-Paul-lès-Durance, France.

14 ⁵ National Institute of Infectious Diseases, Tokyo, Japan

15 ⁶ Biosciences, University of Exeter, Exeter, Devon, UK

16 ⁷ School of Integrative Plant Science, Section of Plant Pathology and Plant-Microbe Biology, Cornell University,
17 Ithaca, NY, USA; United States Department of Agriculture-Agricultural Research Service, Ithaca, NY, USA

18 ⁸ The Milner Centre for Evolution, Department of Biology and Biochemistry, University of Bath, Bath, BA2 7AY,
19 UK

20 ⁹ Department of Zoology, University of Oxford, Oxford, UK.

21

22 *These authors equally contributed to this work.

23

24 * Corresponding authors

25 E-mail: s.k.sheppard@bath.ac.uk and vinatzer@vt.edu

26

27

28 ABSTRACT

29 Many bacterial pathogens are well characterized but, in some cases, little is known about the
30 populations from which they emerged. This limits understanding of the molecular mechanisms
31 underlying disease. The crop pathogen *Pseudomonas syringae* *sensu lato* has been widely isolated
32 from the environment, including wild plants and components of the water cycle, and causes disease

33 in several economically important crops. Here, we compared genome sequences of 45 *P. syringae*
34 crop pathogen outbreak strains with 69 closely related environmental isolates. Phylogenetic
35 reconstruction revealed that crop pathogens emerged many times independently from
36 environmental populations. Unexpectedly, differences in gene content between environmental
37 populations and outbreak strains were minimal with most virulence genes present in both. However,
38 a genome-wide association study identified a small number of genes, including the type III effector
39 genes *hopQ1* and *hopD1*, to be associated with crop pathogens, but not with environmental
40 populations, suggesting that this small group of genes may play an important role in crop disease
41 emergence. Intriguingly, genome-wide analysis of homologous recombination revealed that the
42 locus Psyr 0346, predicted to encode a protein that confers antibiotic resistance, has been
43 frequently exchanged among lineages and thus may contribute to pathogen fitness. Finally, we
44 found that isolates from diseased crops and from components of the water cycle, collected during
45 the same crop disease epidemic, form a single population. This provides the strongest evidence yet
46 that precipitation and irrigation water are an overlooked inoculum source for disease epidemics
47 caused by *P. syringae*.

48

49

50 **DATA SUMMARY**

51

52 We confirm all supporting data, code and protocols have been provided within the article or through
53 supplementary data files. Sequencing reads have been submitted to the NCBI Small Read Archive as
54 Bioproject PRJNA320409 with biosample accession numbers SAMN04942971 to SAMN04943055 and
55 SAMN05301579 to SAMN05301583 and can be accessed at the following link:

56 http://www.ncbi.nlm.nih.gov/bioproject?LinkName=sra_bioproject&from_uid=2500232

57 ☒

58

59 **IMPACT STATEMENT**

60

61 Just like human diseases, new crop diseases emerge without warning and sometimes spread rapidly
62 around the globe causing devastation. Where these pathogens originally came from is often
63 unknown. The bacterial species *Pseudomonas syringae* consists of a group of genetically diverse
64 bacteria including strains that are important crop pathogens as well as strains isolated from wild
65 plants and components of the water cycle, such as clouds, rain and fresh water. The existence of
66 these environmental strains, that are closely related to crop pathogens, suggests that crop
67 pathogenic *P. syringae* possibly emerged from a diverse pre-existing *P. syringae* population that was
68 present in the environment before the development of modern agriculture. Here we found evidence
69 for this hypothesis by sequencing and comparing the genomes of crop pathogenic and
70 environmental strains, we inferred their evolutionary relationships, and identified genes with
71 putative key roles in emergence of crop disease.

72

73

74 INTRODUCTION

75

76 Successful disease prevention and management rely on a detailed understanding of the ecological
77 and evolutionary processes driving disease emergence. In the case of bacterial crop diseases, a lot
78 has been learned about crop pathogen virulence genes and their function (Lindeberg et al., 2008,
79 Tampakaki et al., 2011, O'Brien et al., 2011) but little is known about the genetic basis of crop
80 disease emergence (Vinatzer et al., 2014) and the conditions that promote it (Stukenbrock and
81 McDonald, 2008). For diseases caused by host-restricted obligate pathogens such as *Puccinia*
82 *striiformis* f. sp. *Triticici* and *Puccinia graminis*. f. sp. *tritici*, the causal agents of stripe rust and stem
83 rust respectively (Chen, 2005, Singh et al., 2011), these issues can be addressed relatively easily
84 because dissemination patterns and ecology are restricted to one or a small number of plant hosts.
85 Where infection can be caused by isolates in multiple environmental sources, it can be more difficult
86 to pin point the source. For many human pathogens, the role of environmental reservoirs in disease
87 epidemiology has been well described (Grosso-Becerra et al., 2014, Struve and Kogfelt, 2004, Whiley
88 et al., 2013, Hazen et al., 2015), but for facultative saprophytic crop pathogens, with environmental
89 reservoirs, dissemination routes and interactions within multiple habitats are mostly
90 uncharacterized (Woolhouse et al., 2001, Johnson et al., 2015).

91 In the past ten years, multilocus sequence analysis (MLSA) studies have revealed
92 considerable genetic diversity among environmental isolates that are closely related to epidemic,
93 clonal crop-pathogenic lineages of the plant pathogen *Pseudomonas syringae* (*sensu lato*) (Morris et
94 al., 2008, Morris et al., 2010, Monteil et al., 2012, Monteil et al., 2014b). *P. syringae* is one of the
95 economically most important bacterial crop pathogens and a well characterized model species for
96 molecular plant-microbe interactions (Hirano and Upper, 2000, O'Brien et al., 2011). Environmental
97 isolates have been collected from wild plants as well as non-plant reservoirs including soil, plant
98 debris, and components of the water cycle including clouds, precipitation, and surface water (Morris
99 et al., 2013, Berge et al., 2014). Host range analysis revealed that some crop-pathogenic epidemic
100 clones (referred to as "crop pathogens" from here on) within *P. syringae*, such as the most common
101 lineage of the tomato pathogen *P. syringae* pathovar *tomato* (*Pto*), have a narrow host range limited
102 to tomato (Cai et al., 2011a, Cai et al., 2011b). Conversely, lineages such as the cantaloupe pathogen
103 *P. syringae* pathovar *aptata* (*Pap*) have a broad host range, infecting various plant families (Morris et
104 al., 2000, Berge et al., 2014). In the case of *Pto*, MLSA revealed the existence of closely related
105 isolates from natural freshwater sources and recombination events between these environmental
106 isolates, *Pto*, and other crop pathogens (Monteil et al., 2013). The environmental lineages were
107 found to be equipped with some of the same virulence genes as the crop pathogen *Pto*, in particular,
108 genes coding for type III-secreted (T3S) effectors, the best studied and most important class of
109 virulence genes in *P. syringae* (Lindeberg et al., 2008). Moreover, the environmental isolates had a
110 wider host range than *Pto* but were less virulent on tomato (Monteil et al., 2013). Taken together,
111 these results are consistent with the evolution of highly virulent crop pathogens with a relatively
112 narrow host range from a population of ancestors with a wider host range. This potentially occurs
113 through the acquisition of genomic elements that promote virulence on the crop hosts but reduce
114 virulence (or fitness) on other hosts. However, what these genomic elements might be, and whether
115 they were acquired by horizontal gene transfer, remains unknown.

116 The increasing availability of large genomic datasets provides new opportunities for
117 investigating pathotypes in multiple niches (Vinatzer et al., 2014). By comparing the genomes of *P.*

118 *syringae* crop pathogens and isolates from environmental reservoirs, it should be possible to identify
119 the genetic basis of disease emergence and the genomic regions that are horizontally transferred
120 between strains, in particular, between crop pathogens and their environmental relatives.
121 Therefore, we sequenced the genomes of 107 isolates of crop-pathogenic and environmental *P.*
122 *syringae* and analyze them together with 86 publically available *P. syringae* genomes. We
123 investigated two *P. syringae* phylogroups (Berge et al., 2014) with contrasting host ranges (Buell et
124 al., 2003) and disease etiology (Feil et al., 2005): phylogroup 1a, which includes *Pto* and other
125 related crop pathogens and environmental isolates; and a subset of phylogroup 2d, for which we
126 sampled exclusively closely related *Pap* isolates from diseased cantaloupe and the environment.
127 Importantly, each phylogroup also includes one intensively studied model pathogen isolate: *P.*
128 *syringae* pv. *tomato* (*Pto*) DC3000 in phylogroup 1a and *P. syringae* (*Psy*) B728a in
129 phylogroup 2d. For both of these isolates, virulence traits have been investigated for decades and
130 closed genome sequences are available (O'Brien et al., 2011). Phylogenetic reconstruction, core and
131 accessory genome analysis, and genome-wide association approaches were then used to
132 characterize the evolutionary relationships between crop pathogens and environmental relatives,
133 the population structure of these phylogroups, and the genetic basis of crop-adaptation (Sheppard
134 et al., 2013, Pascoe et al., 2015). The results provide new insight into crop pathogen emergence,
135 crop-adaptation, and pathogen dissemination.

136

137 METHODS

138

139 Isolates and sequencing

140

141 Genomes of 92 *P. syringae* isolates from phylogroups 1a and 2d (Berge et al., 2014) were chosen for
142 genome sequencing, whereby phylogroup 1a was sampled maximizing genetic diversity avoiding
143 multiple crop pathogen strains with identical MLSA sequences while for phylogroup 2d only one
144 subset was sampled with isolates that were identical at two MLSA loci. These datasets were
145 augmented with 12 genome sequences from the same phylogroups available in public databases,
146 including reference genomes from *Pto* strain DC3000 (Buell et al., 2003) and *Psy* strain B728a (Feil et
147 al., 2005). Most of the isolates sequenced in this study were described previously (Monteil et al.,
148 2012, Monteil et al., 2014b, Morris et al., 2008, Morris et al., 2010) with 36 isolates collected from
149 diseased crops (defined as cultivated lands) and 56 isolates collected from streams and rivers (11
150 isolates), precipitation (15 isolates), irrigation water (11 isolates) or epilithic biofilms (12 isolates)
151 and leaf litter (7 isolates). Table S1 contains a detailed list. Genomes representing other *P. syringae*
152 phylogroups were included in the analyses: three of them were sequenced in this study while 86
153 genomes were downloaded from public databases to give a total of 193 isolate genomes (Table S1).

154 For genomes sequenced in this study, DNA was extracted using the Gentra Puregene bacteria
155 kit (QIAGEN; cat. number: 158567), using manufacturer's instructions. DNA was quantified using a
156 Nanodrop spectrophotometer. The library preparation was performed with the Nextera XT DNA
157 sample prep kit from Illumina following manufacturer's instructions for denaturing and
158 normalization steps. High-throughput genome sequencing was performed using 151 X 151 PE Rapid
159 Run mode of a HiSeq 2500 sequencer (Illumina, San Diego, CA, USA) using Illumina Truseq
160 sequencing reagents. The quality of resulting sequencing reads was examined using FastQC
161 (Andrews, 2010). TrimGalore (Krueger, 2015) was used to trim sequence reads and remove poor-

162 quality data, using command-line options “-q 30 and --paired”. Illumina adapter sequences were
163 removed using CutAdapt (Martin, 2011). Cleaned sequencing reads were assembled using the *de*
164 *novo* assembly algorithm Velvet (Zerbino and Birney, 2008) (version 1.2.08). Value of k was
165 optimised by assembling over a range of values and choosing the assembly with maximal N₅₀. The
166 minimum output contig size was set to 200 bp with the scaffolding option switched off, all other
167 program settings were left unchanged. The average number of contigs and the standard error in 95
168 newly sequenced *P. syringae* genomes was 474 ± 32 for an average total assembled sequence size of
169 5,969,322 ± 18,142 bp. Genome sequence data, including raw sequence reads in FastQ format
170 deposited in the Sequence Read Archive (SRA), are available via BioProject accession PRJNA320409.

171

172 Core and accessory genome

173

174 Analyzing core and accessory genome variation, genealogies, and recombination patterns, we
175 investigated the evolutionary relationships linking outbreak strains to their relatives in the
176 environment. A reference pan genome approach (Meric et al., 2014) and gene-by-gene alignment
177 (Sheppard et al., 2012), was implemented using BIGSdb open source software (Jolley and Maiden,
178 2010). First, a reference gene list was assembled from four publicly available genomes, *Pto* DC3000
179 (Buell et al., 2003) and *Psy* B728a (Feil et al., 2005) *P. syringae* pv. *phaseolicola* 1448a (Joardar et al.,
180 2005) and the environmental strain CC1557 (Hockett et al., 2014)(Table S1). The total number of
181 genes in these isolates was 20,955 and after removal of 13,471 allelic variants that shared >70%
182 nucleotide identity across ≥50% of the genes length, the final reference pan-genome list contained
183 7,484 unique loci. Each locus was searched in the 193 genomes of all isolates using the BLAST
184 algorithm and setting parameters for locus match to a minimum of 70% sequence identity over a
185 minimum of 50% of the query sequence length. The average core genome nucleotide sequence
186 identity within *P. syringae* is considerably higher than the BLAST match criteria. Therefore these
187 BLAST parameters ensure relatively low stringency for identifying homologous genes as in existing
188 whole genome MLST methodology (Maiden et al., 2013, Meric et al., 2014, Sheppard et al., 2012,
189 Jolley and Maiden, 2010). Consistent with whole genome MLST (Berge et al., 2014), a matrix was
190 produced summarizing the presence/absence and allelic diversity of reference pan-genome genes,
191 based upon these BLAST parameters. Each gene of the reference pan-genome that was not, or only
192 partially, detected in a genome was indicated as missing or truncated, and this number was
193 calculated at each locus for all *P. syringae*. However, truncated gene sequences detected at the end
194 or beginning of a contig, were considered as present but were not counted as alleles. For each pair
195 of isolates, the number of shared genes and alleles (identical sequences at a given locus over the
196 whole sequence length) was calculated and the core genome for each species, and for the genus,
197 was defined as the complement of genes that were present in all isolates.

198

199 Population genetic structure

200

201 Phylogenetic trees were constructed from the alignment of the core genome. The core genome was
202 determined based on the 5619 coding sequences of the reference genome *Pto* DC3000 and its two
203 plasmids. Genes in the core genome were aligned individually using MAFFT (Katoh and Toh, 2008)
204 and concatenated to produce contiguous sequence alignments.

205 For the analysis of the entire *P. syringae* species complex (193 genomes), core loci for which
206 some of the sequences were truncated were kept in the analysis, which accounts for a total of 1889
207 genes. A recompiled version of FastTree 2.1.7 (Price et al., 2010) was used to reconstruct an
208 approximation of a maximum likelihood tree. With this configuration, the minimum branch length
209 was 1 substitution for every 2,000,000 bp (1000 times higher than the default FastTree parameters).
210 The software was run with the Jukes Cantor model of nucleotide evolution and gaps from truncated
211 sequence alignments were considered as missing data. The tree was visualized and annotated using
212 FigTree v1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree>).

213 For the analysis of the phylogroups 1a and 2d, trees were constructed from alignments
214 including only non-truncated sequences for all core loci (respectively 810 and 2147 loci).
215 Genealogies were inferred using ClonalFrame, a model-based approach for inference of
216 microevolution in bacteria that accounts for recombination events that can disrupt a phylogenetic
217 reconstruction (Didelot and Falush, 2007). This program differentiates mutation and recombination
218 events on each branch of the tree based on the density of polymorphisms. Clusters of
219 polymorphisms are likely to have arisen from recombination, and scattered polymorphisms are likely
220 to have arisen from mutation. The program was run with 20,000 burn-in iterations, followed by
221 50,000 and 100,000 sampling iterations for phylogroup 2d and 1a respectively, until convergence.
222 The consensus tree represents combined data from three independent runs with 75% consensus
223 required for inference of relatedness. Recombination events were defined as sequences with a
224 length of >50 bp with a probability of recombination 75% over the length, reaching 95% in at least
225 one site.

226 Associations of lineages with isolation sources were investigated applying the HierBAPS
227 clustering model (Cheng et al., 2013, Corander et al., 2004). This method allows revealing of nested
228 genetic population structures and any association of strain metadata with genetically divergent
229 clusters and the substructure within them. Therefore, we could specify the genetic boundaries
230 between *P. syringae* lineages at different resolutions and test if they are associated to a unique
231 isolation source or not. The same alignments used for each tree construction (whole *P. syringae*
232 diversity, 1a or 2d) were used to infer genetically divergent clusters increasing levels of resolution
233 from 1 to 4. The mixture partition was inferred setting the prior to 10 *k* panmictic subpopulations in
234 which individuals are uniformly distributed and the structure, which maximizes the posterior
235 distributions, is obtained using a stochastic search algorithm.

236

237 Genome wide association mapping

238

239 We sought to determine the genetic basis of crop pathogen emergence by comparing genomes of
240 crop pathogens and their relatives from other sources. We used a recently developed GWAS method
241 that is adapted for bacterial populations (Sheppard et al., 2013, Pascoe et al., 2015, Yahara et al.,
242 2016b) and excludes associations due to confounding population structure. The whole genome
243 sequence of each isolate was fragmented into unique overlapping 30-bp words. For each word, the
244 method examines the extent of association with the phenotype. To test significance of association of
245 each word after controlling for the effect of population structure and clonal inheritance of genetic
246 variants, here determined using Clonalframe (Didelot and Falush, 2007), the method computes P-
247 values by comparing the observed association score with a null distribution of the score calculated
248 through Monte Carlo simulations (Martins and Garland, 1991, Garland et al., 2005). To account for

249 multiple testing, only words with a probability below $5 \cdot 10^{-4}$ were considered significant. The
250 distribution of source-associated words for which homologs were found in reference genome *Pto*
251 DC3000 (Buell et al., 2003) was visualized using Artemis (Rutherford et al., 2000) and DNAPlotter
252 (Carver et al., 2009) (Fig. 3).

253

254 Inference of homologous recombination

255

256 Extent of homologous recombination was inferred between crop pathogens and related
257 environmental isolates and ancestors of phylogroups 1a and 2d, to investigate genetic fluxes at a
258 fine scale and the evolutionary origin of genes associated with disease outbreak populations. First,
259 chromosome painting (Lawson et al., 2012, Yahara et al., 2013) builds a co-ancestry matrix
260 summarizing the number of recombination-derived “chunks” of DNA from each donor to each
261 recipient isolate. Using this matrix, fineSTRUCTURE (Lawson et al., 2012) conducts model-based
262 clustering of individuals by a Bayesian MCMC (Markov chain Monte Carlo) approach that explores
263 the space of possible partitions. In parallel, we applied the Ordered Painting approach (Yahara et al.,
264 2014) to identify hot spots of recombination within phylogroups. In this method, the extent of
265 genealogical changes for a specific site due to recombination compared with the average genome
266 genealogy is represented by the distance statistic H_i representing recombination hotness at each
267 polymorphic site i (Yahara et al., 2016a). These atypical changes, for example above the top
268 percentile of the whole genome distribution, are expected to indicate hot spots of recombination.

269

270

271 **RESULTS**

272

273 Population structure of the *P. syringae* species complex

274

275 Before determining the population structure of *P. syringae sensu lato*, we determined whether the
276 193 *P. syringae* genome sequences (sequenced in this study or publically available) formed a clade
277 that is distinct from genomes of other *Pseudomonas* species. A genealogy of 629 available
278 *Pseudomonas* spp. genomes was constructed based upon 52 ribosomal protein gene sequences
279 (Table S1 and S2). The phylogenetic tree confirmed *P. syringae* as a largely monophyletic species
280 complex. The one exception was phylogroup 13 (isolate UB246 (Berge et al., 2014)), which clustered
281 with *P. fluorescens* based on population structure analysis inferred with BAPS, a Bayesian statistical
282 clustering method (Fig. S1).

283 To analyze evolutionary relationships with greater resolution within the *P. syringae* species
284 complex, a core genome of the 193 isolates (Table S2) was determined by aligning all genome
285 sequences with the annotated genes of the fully sequenced and annotated reference genome *Pto*
286 DC3000 (Buell et al., 2003). The *P. syringae* genealogy constructed from 1889 core gene loci (a total
287 of 108,393 bp for which 38,484 were variable sites) confirmed genealogies previously inferred by
288 MLSA and by previous core genome analysis (Berge et al., 2014). Importantly for the goal of this
289 study, it also revealed recent common ancestry of all lineages within phylogroups 1a and within

290 phylogroup 2d, respectively (Fig. 1a). To determine the extent to which these phylogroups reflected
291 distinct genetically divergent clusters within the *P. syringae* species complex, we performed a
292 hierarchical clustering analysis of the core genome using the HierBAPS method that estimates
293 nested population structures (Cheng et al., 2013). Based on clustering at the lowest level of
294 resolution, clade designations were congruent with most phylogroups, including phylogroups 2d and
295 1a.

296 To investigate the role of environmental reservoirs in the emergence of crop pathogens, we
297 then focused on phylogroup 1a, for which we had sequenced representative isolates of all crop
298 pathogens available to us and all available environmental relatives in order to include as much
299 genetic diversity as possible. Individual genealogies were built and Bayesian analysis of population
300 structures (Cheng et al., 2013) was performed. Since homologous recombination may have had an
301 impact on population genetic structure, clonal relationships of crop pathogens and their relatives
302 were first investigated using Clonalframe (Didelot and Falush, 2007), which accounts for
303 recombination when constructing the genealogy. At the lowest degree of resolution, the population
304 structure analysis of all core genes identified three genetic clusters, each containing both crop
305 pathogens and environmental relatives (Fig. 1b). The clonal frame analysis also revealed that some
306 monophyletic groups of lineages within these clusters correspond to one isolation source only. For
307 example, the *PtoDC3000* crop pathogen clusters only with other crop pathogens and the LAB0041
308 isolate from an alpine epithilic biofilm, clusters only with other environmental isolates. Importantly
309 though, several crop pathogens, such as the tomato pathogens *Pto* T1 and *Pto* JL1065 and the
310 snapdragon pathogen *Pan* 126, are interspersed with lineages from rain or surface water within the
311 same group. This shared ancestry between some crop pathogens and some environmental isolates,
312 together with the long external, short internal branches of the clonal frame tree (Fig. 1b), and a high
313 degree of reticulation in the NeighborNet network (Fig. S2a), are all consistent with a scenario of
314 multiple emergences of different crop pathogens and environmental lineages from recombining
315 ancestral populations.

316 For phylogroup 2d, we investigated the different components of the water cycle in pathogen
317 dissemination. Isolates from diseased crops, collected during a cantaloupe blight epidemic in France,
318 were sequenced together with isolates from precipitation, surface water, irrigation water, and
319 ground water (Morris et al., 2000, Morris et al., 2008). Specifically, we chose isolates that were
320 identical at two MLSA loci and we wanted to determine whether crop and environmental isolates
321 would cluster together, or separately, based on core genome sequences. ClonalFrame and
322 population structure analysis revealed that most of these isolates clustered together with a star-like
323 genealogy without any separation between crop isolates and environmental isolates. The genomes
324 of crop and environmental isolates in the first cluster were extremely similar to each other (Average
325 Nucleotide Identities ranging from 99.50 to 99.97% within each clade). To confirm these results,
326 sequencing reads were aligned against the *Psy* B728a genome and single nucleotide polymorphisms
327 (SNPs) were identified. This approach revealed that some of the crop isolates differed from their
328 most similar environmental relatives by as few as 3 SNPs per million bp (data not shown) suggesting
329 very recent exchanges of *P. syringae* between cantaloupes and water cycle component reservoirs.
330 Moreover, ClonalFrame showed that some core genes experienced recent homologous
331 recombination between crop and environmental isolates, which was supported by the Neighbor Net
332 shown in Fig. S2b.

333

334 Core and accessory genome variation and pathogen emergence in the 2d and 1a phylogroups

335

336 Genome-wide genetic differences between crop pathogens and environmental isolates were
337 investigated using a reference pan-genome approach (Meric et al., 2014). In short, four fully
338 assembled and annotated genome sequences from four different phylogroups were chosen as
339 references: *PtoDC3000* (Buell et al., 2003), *PsyB728a* (Feil et al., 2005), *Pph1448a* (Joardar et al.,
340 2005) and CC1557 (Hockett et al., 2014). Then the pan-genome of these four reference genomes,
341 defined as the total set of gene families present in the four genomes, was determined and found to
342 consist of 7,484 unique genes. Finally, the reference pan-genome was aligned against all the other
343 genomes. This analysis revealed that every genome in phylogroup 1a contained a set of 3576 genes
344 of the reference pan-genome thus representing the 1a core genome. Every genome of phylogroup
345 2d contained a set of 4147 core genes representing the 2d core genome. 3062 genes were present in
346 both phylogroups representing their combined core genome. At the remaining 4422 loci, genes were
347 either present or absent presenting the combined accessory genome (Fig. 2a). Phylogroup 2d had a
348 lower average p-distance between allele pairs, and on average fewer unique alleles per gene —
349 0.201 ± 0.001 compared to 0.480 ± 0.003 for phylogroup 1a (Student's t-test, $P < 0.001$, Fig. S3a &
350 S3b). Note that differences in the genetic diversity of phylogroups 2d and 1a reflect our sampling
351 strategy, maximizing genetic diversity for phylogroup 1a while prioritizing a single genetic lineage for
352 phylogroup 2d.

353 To determine the extent of genes associated with isolation sources, we compared gene
354 content of crop pathogens and environmental isolates. It is striking that none of the genes that are
355 core to crop pathogen populations are absent in environmental populations, and that none of the
356 genes that are core to environmental populations are absent in crop pathogen populations (Fig. 2a).
357 This suggests that there is weak ecological differentiation between crop pathogens and their
358 environmental relatives. Moreover, for both phylogroups the majority of accessory genes present in
359 environmental isolates are also present in crop pathogens and *vice versa*. This result reveals that
360 there is no strong barrier to gene flow between environmental isolates and crop pathogens.
361 Importantly, about 61% of T3S effector genes detected in crop pathogen populations were present
362 in environmental relatives as well (Table S4 and Fig. S4).

363 Pairwise genome comparisons showed that patterns of core genome allelic similarity (Fig.
364 2b) and accessory genome similarity, measured as similarity in regard to presence and absence of
365 accessory genes (Fig. 2c), reflected genealogies rather than isolation host or source (crop versus
366 environment). Extending this analysis beyond phylogroups 1a and 2d to *P. syringae* genome
367 sequences in phylogroups 2b and 3 confirmed the same trend (Fig. S5). Therefore, pairwise genome
368 comparisons confirmed the absence of strong barriers to gene flow between crop pathogens and
369 environmental relatives as well as between pathogens of different hosts.

370

371 Candidate genes associated to pathoadaptation

372

373 Although the previous analyses clearly showed that crop pathogens are not genetically isolated from
374 their environmental relatives, it is still possible that at least a small set of genes, or alleles, may be
375 more frequently associated with either crop pathogens or environmental isolates. This would
376 suggest that crop pathogens are adapted to a pathogenic life style (pathoadaptation) that is
377 different from the adaptation of environmental isolates to a life cycle in non-agricultural
378 environments. Therefore, genetic elements overrepresented either in crop pathogen strains, or in

379 environmental isolates, were determined using a genome-wide association study (GWAS) (Sheppard
380 et al., 2013, Pascoe et al., 2015), which identifies 30-bp DNA sequences (words) in the core and
381 accessory genome taking into account the clonal frame and vertical inheritance and core genes.
382 Based on the analysis of 67 and 37 isolates within the 1a and 2d phylogroups, 73,299 and 5,970
383 words, respectively, were identified that were over represented in crop pathogens. These mapped
384 to 571 genes in phylogroup 1a and 222 genes in phylogroup 2d, 74% of which are annotated with a
385 putative function (Table S3). Genes containing pathogen-associated elements were mapped to
386 reference genomes of the crop pathogens *Pto* DC3000 and *Psy* B728a, for 1a and 2d isolates
387 respectively, using Artemis (Rutherford et al., 2000) and DNAPlotter (Carver et al., 2009) (Fig. 3a).
388 Associated genes were dispersed across the genome with evidence of 7 (phylogroup 1a) and 5
389 (phylogroup 2d) hot-spots of strong pathogen association—with a p-value less than 5.10⁻⁶.

390 For phylogroup 1a isolates, only 7% of the mapped words were associated with known
391 virulence genes, with 22% of the total associated words mapping to a single 25-kb region consisting
392 in 21 adjacent genes of the *Pto* DC3000 genome (Fig. 3a). This included the T3S effector genes *hopD1*
393 and *hopQ1* (Lindeberg et al., 2008). These genes were present in all but two of crop pathogen
394 isolates and absent from all environmental isolates (Fig. 3b). Importantly, after aligning raw reads of
395 these two isolates against the *PtoDC300* genome (data not shown), relic fragments of *hopD1* and
396 *hopQ1* were even detected in the only two crop pathogens, *P. syringae* pv. *apii* BS252 and *P.*
397 *syringae* pv. *antirrhini* 126, that did not contain the intact genes. Other crop pathogen-associated
398 hotspots in the 1a phylogroup contained genes encoding putative proteins related to: (i) replication,
399 integration, recombination and repair of DNA; (ii) carbohydrate, lipid and amino acid transport; (iii),
400 energy production and conservation (Table S3).

401 Interestingly, no T3S effector genes were associated with crop pathogen isolates in
402 phylogroup 2d and, compared to the 1a phylogroup, the repertoire of T3S effector genes was
403 generally smaller in the 2d phylogroup (Fig. S4). Predicted functions of genes that were significantly
404 associated with crop pathogen isolates in 2d were: (i) DNA transcription and translation regulation;
405 (ii) uptake of sparse substrates linked to TonB-dependent transporters; (iii) the conversion of energy
406 into storage molecules; (iv) secondary metabolite production and export; (v) Type I secretion
407 systems (Table S3). It should be noted that sampling within phylogroup 2d focused on a subset of
408 very similar isolates to address questions about pathogen dissemination. Additional sampling would
409 be necessary for more robust inference of pathogen associations in this lineage.

410

411 Homologous recombination and pathoadaptation

412

413 In recombining bacteria, the acquisition of DNA from other lineages can confer novel functions, such
414 as those related to pathoadaptation (Ochman et al., 2000). ClonalFrame analysis and patterns of
415 reticulation using simple NeighborNet genealogical reconstructions (Fig. S2), suggested recent
416 homologous recombination among various crop pathogens and environmental isolates. A more
417 detailed analysis was thus carried out to investigate inter- and intra-phylogroup homologous
418 recombination and to quantify recombination landscapes across the genome.

419 Gene flow within and between phylogroups was quantified by characterizing DNA donated
420 and received among the strains using chromosome painting (Yahara et al., 2013). The number of
421 recombination-derived chunks of DNA, defined as genetic material donated from a nearest 'donor'
422 to a 'recipient' haplotype, was summarized into a co-ancestry matrix. A co-ancestry matrix with all

423 1a and 2d strains confirmed the barrier to gene flow between the two phylogroups for which the
424 number of DNA chunks was under 1 per genome on average. However, gene flow was observed
425 within each phylogroup (Fig. 4a and 4b). Both co-ancestry matrices not only showed admixture
426 between ancestors of all lineages but also showed that gene flow occurred between crop pathogens
427 and environmental isolates in both directions, like for *Pap* CC94 (Fig. 4b). However, in phylogroup 1a
428 isolates, gene flow was asymmetrical with some lineages being principally donors or recipients. For
429 example, *Pto* DC3000 was mostly a donor and not a recipient to isolates from both isolation sources
430 (Fig. 4a) in contrast to *Psy* B728a from phylogroup 2d, which received twice as many DNA chunks
431 from ancestors of other crop and environmental isolates compared to what the strain donated to
432 other strains (Fig. 4b).

433 In parallel to characterizing the direction of gene flow, recombination hot spots across the
434 genome were identified based upon a per-site estimate of intensity of recombination (H_i) (Yahara et
435 al., 2016a), which refers to a normalized value quantifying the extent of genealogical changes due to
436 recombination compared to the average genealogy (Fig. 5a and 5b). For phylogroups 2d and 1a
437 respectively, 144 and 244 recombining genes had H_i values in the upper 2.5% for at least one base
438 position (Table S5). Evidence for a role of recombination in pathoadaptation was seen in the 1a
439 phylogroup where 72 genes recombined in crop pathogens. A total of 8% were known virulence
440 genes, including genes coding for T3S effectors (*HopAA1-1*, *HopAH1* and *HopB1*), T4 pili, chemotaxis,
441 pioverdine production, levansucrase (*lsc-1*). The highest rates of recombination in crop pathogens
442 were found for those genes that were also identified as hot spots of recombination in environmental
443 isolates and were mostly associated with hypothetical proteins or a peptide ABC transporter
444 permease (*PSPTO* 265, 561, 2271, 2587 and 5552). In 2d, 116 genes recombined in crop isolates and
445 were mainly associated with metabolism, regulators (*lysR*, *lcrR*, *tetR*), transporters, the type 1
446 secretion system (involving *TolC* and *HlyD* proteins), while the 23 genes recombining in
447 environmental isolates were associated with other functions among which were several T3S
448 effectors and structural components (e.g. *hopM1*, *hrcN*, *hrcQ*, *hrcV*, and *hrcK1*). As observed within
449 the 1a phylogroup, hot spots of recombination in 2d isolates from crops were sometimes hot spots
450 in environmental isolates as well. Hot spots in 2d were associated with genes coding for hypothetical
451 proteins (PSYR 392 and 393), a tRNA-dihydouridine synthetase A (PSYR 1936) and a binding-protein
452 dependent transport system (PSYR 2903 and 2904). Importantly, 18 orthologous genes were found
453 to be recombining in both phylogroups, some of them coding for ABC transporters, proteins involved
454 in antibiotic resistance (*PSPTO* 3132, 3302 and 5191), chemotaxis, extracellular solute binding
455 (*PSPTO* 2962 and 3302), glutamate racemase (*murl*), porins (*oprD*) and a heavy metal translocating
456 ATPase (*cadA_1*) (Table S5).

457 Finally, we sought to identify those genes associated with crop pathogens (based on GWAS)
458 and hot spots of recombination, and determine whether they have been previously characterized as
459 virulence genes. Comparisons flagged the *PSPTO* 5191/PSYR 0346 gene, a member of the
460 AcrB/AcrD/AcrF family of membrane proteins, implicated in multiple antibiotic resistance in
461 *Salmonella typhimurium* (Piddock, 2006), that is not only recombining in both phylogroups but also
462 associated with pathoadaptation in 2d. Other genes associated with pathoadaptation in phylogroup
463 2d were also recombination hot spots in crop pathogens, such as PSYR 195, 336, 1992 and 3131
464 coding for a hypothetical protein, an outer membrane autotransporter barrel, a zinc-containing
465 alcohol dehydrogenase superfamily protein and a secretion protein HlyD, respectively. However,
466 only PSYR 1794 and PSYR 3151 coding for a non-ribosomal peptide and the protein E of a type II
467 secretion system respectively, were also known virulence genes (Lindeberg et al., 2008).

468 In phylogroup 1a, the T3E gene *hopAH1* was the only known virulence gene located within a
469 recombination hot spot and that had also been associated with pathoadaptation in GWAS. Other
470 crop pathogen-associated genes that are not known virulence genes corresponded to hot spots of
471 homologous recombination, including the ferric iron reductase protein *fhuF*, an acyltransferase
472 *PSPTO_0997* (COG accession COG1835) and a transcriptional regulator from the GntR family (SMART
473 accession smart00895).

474

475 DISCUSSION

476

477 The bacterial plant pathogen *P. syringae* is well known as a model organism to study the molecular
478 basis of plant - microbe interactions (Alfano and Collmer, 2004, Xin and He, 2013). Moreover, there
479 is no other bacterial plant pathogen species for which so much is known about genetic diversity
480 outside of agricultural environments (Morris et al., 2008, Morris et al., 2010, Monteil et al., 2012,
481 Berge et al., 2014), which has made *P. syringae* a model for studying crop pathogen emergence as
482 well (Mohr et al., 2008, Yan et al., 2008, Cai et al., 2011b, Cai et al., 2011a, Diallo et al., 2012, Morris
483 et al., 2013, Bartoli et al., 2015). Here, we applied for the first time a population genomics approach
484 to *P. syringae* crop pathogens and their close environmental relatives and gained new insight into
485 crop pathogen ancestry, emergence, crop-adaptation, and dissemination.

486 Phylogenetic reconstruction based on a small number of genes had already suggested that
487 epidemic *P. syringae* crop pathogens in phylogroup 1a, such as *Pto*, have close relatives in non-
488 agricultural environments (Monteil et al., 2013). After sequencing the genomes of representative
489 crop pathogens and their environmental relatives in phylogroup 1a and 2d, we have now shown that
490 several crop pathogens, including *Pto* T1 and *Pap* CC94, are more closely related to environmental
491 isolates than to other crop pathogens. Moreover, investigating patterns of recombination and
492 population structure at the whole genome level within each phylogroup, we show that several crop
493 pathogens emerged from an ancestral recombining population independently from each other.

494 A fundamental question is when these emergence events occurred. Accurate molecular
495 clock estimates are not possible within the sample frame of this study as it does not include a
496 longitudinal sample. However, almost identical pathogen isolates have been sampled from crops
497 dozens of years apart (Cai et al., 2011a, Clarke et al., 2015) and this is consistent with a mutation
498 rate as low as one mutation per million bp per year. Therefore, considering that some of the *P.*
499 *syringae* crop pathogen lineages in phylogroup 1a have diverged substantially from each other, more
500 than one mutation per 1000 bp in some MLSA loci, it is likely that their most recent ancestor existed
501 before humans started domesticating crop plants and before the advent of agriculture 5,000 to
502 10,000 years ago. Therefore, the inferred ancestral population may have existed in non-agricultural
503 plant communities and environmental reservoirs.

504 For phylogroup 2d, our analysis focused on a subset of isolates selected based on their
505 identity at two MLSA loci excluding most of the genetic diversity that is known to exist in this
506 phylogroup. This sample frame allowed comparison of contemporaneous isolates collected from
507 diseased cantaloupe, sampled during a cantaloupe blight epidemic in France, and their closest
508 relatives isolated from irrigation water, precipitation, and ground water. Based on whole-genome
509 analysis, isolates with almost identical genome sequences clustered together despite their different
510 isolation sources and with some isolates from cantaloupes collected as far apart as 350 kilometers.
511 This result is consistent with frequent migration events between cantaloupe production fields and

512 components of the water cycle and suggests that rain and irrigation water are involved in the
513 dissemination of crop-pathogenic *P. syringae* between geographically distant fields. Although *P.*
514 *syringae* had been reported in rain before (Morris et al., 2008, Constantinidou et al., 1990, Morris et
515 al., 2010, Monteil et al., 2014a), no strong genetic linkage between the presence of *P. syringae* in
516 rain or irrigation water on one hand and *P. syringae* isolated from diseased crops on the other was
517 possible without genome sequences. Additionally, in the specific case of phylogroup 2d isolates,
518 which have a wide host range (Morris et al., 2000), rain and irrigation water may also be the original
519 inoculum source of epidemics by transporting the pathogen from colonized wild and crop plants
520 over long distances to crop fields and starting new outbreaks. This has important implications for
521 crop disease prevention programs since it is common place to link new *P. syringae* disease outbreaks
522 to contaminated seed and nearby weeds but not to components of the water cycle or irrigation
523 water (McCarter et al., 1983, Gitaitis and Walcott, 2007).

524 To understand the molecular basis of crop disease emergence, it is necessary to determine
525 what differentiates epidemic crop pathogen isolates from their close relatives that are not epidemic
526 pathogens. Bulk shotgun sequencing and virulence analysis of a small number of environmental
527 isolates in phylogroup 1a had already revealed that these isolates contain well-known virulence
528 genes such as T3S effector genes and that some of the environmental isolates are almost as virulent
529 as *Pto* on tomato and other plant species (Monteil et al., 2013). By extending comparison of gene
530 content to multiple whole genomes, we show here that all environmental relatives in phylogroup 1a
531 are equipped with T3SSs and with repertoires of T3S effectors and other virulence genes similar to
532 those of crop pathogens. Moreover, extending our virulence assays to all environmental relatives in
533 phylogroup 1a, we confirmed that these isolates are all pathogenic on tomato, although less virulent
534 than *Pto* (data not shown). We thus conclude that although the analyzed environmental isolates
535 were originally mostly collected from water, they appear to be adapted to life in association with
536 plants and they possibly are pathogens of wild plants. These observations raise many questions
537 about the role of wild plants and crop plants in the emergence and diversification of virulence traits
538 in plant pathogenic populations. In particular, considering the genetic diversity of virulence-gene-
539 equipped environmental populations it is not clear why there are not more frequent emergence
540 events. For example, there is only a single *Pto* lineage that has spread successfully worldwide on
541 tomato (Cai et al., 2011a). It is therefore possible that there is a genetic barrier to emergence
542 whereby only rare combinations of virulence genes allow emergence of an epidemic clone.

543 To test this hypothesis, a GWAS (Sheppard et al., 2013) was performed to identify genomic
544 regions that show a statistically significant association with crop pathogens. Intriguingly, only two of
545 the 58 T3S effector genes, *hopD1* and *hopQ1*, were found to be pathogen-associated in phylogroup
546 1a and not a single T3S effector was found to be pathogen associated in phylogroup 2d. Like some
547 other T3S effectors, HopD1 interferes with the immune response triggered by other effectors (Block
548 et al., 2014) and HopQ1 interferes with immunity triggered by microbial-associated molecular
549 patterns, specifically the immune response triggered by the bacterial flagellum (Li et al., 2013b, Li et
550 al., 2013a, Hann et al., 2014). Moreover, just as for many other effectors, deletion of either *hopQ1* or
551 *hopD1* from *Pto* DC3000 has been shown to reduce bacterial growth on some plant genotypes under
552 laboratory conditions (Wei et al., 2007). Therefore, the fact that only *hopQ1* and *hopD1* were
553 identified as crop pathogen-associated in the GWAS suggests that the specific contribution to
554 virulence by these two effectors is in some way more relevant in the life cycle of epidemic crop
555 pathogens compared to the life cycle of bacteria associated with non-agricultural environments
556 while the other T3S effectors, that are more evenly shared by crop pathogens and their
557 environmental relatives, play a role in fitness in agricultural as well as non-agricultural environments.

558 In phylogroup 2d, although our sampling strategy may have inflated associations, the GWAS
559 revealed fewer statistically significant associations with source of isolation with lower p-values
560 compared to associations in phylogroup 1a. In particular, the distribution of the top GWAS hits
561 within the group of most closely related isolates (Fig. 3b) showed no clear association at all. This is
562 most likely a result of our sampling frame and is consistent with our earlier conclusion that the
563 sequenced 2d isolates represent a single population of *P. syringae* that regularly transfers between
564 cantaloupe, other plant hosts, and components of the water cycle.

565 Finally, our previous analysis of phylogroup 1a isolates (Monteil et al., 2013), and analysis of
566 the kiwifruit pathogen *P. syringae* pv. *actinidia*, suggested that crop pathogens emerge from
567 recombining *P. syringae* populations. While it has been suggested that this recombination mainly
568 occurs within *P. syringae* pathogen populations specific to a host plant species, such as kiwifruit
569 (McCann et al., 2013), we provide evidence that the recombining *P. syringae* populations may not be
570 host-specific and could include environmental *P. syringae* residing, for example, in wild plants and in
571 decaying plant material, where *P. syringae* populations reach densities as high as 10^6 colony forming
572 units per gram (Monteil et al., 2012). Importantly, using quantitative analysis of homologous
573 recombination, we demonstrate that recombination between crop pathogens and their
574 environmental relatives is as frequent as recombination between isolates within either niche (Fig. 4).
575 This result reveals that *P. syringae* crop pathogens can acquire new genes from environmental
576 populations and vice versa, possibly including potential virulence genes, which might exist at low
577 frequency in environmental populations. In fact, although *hopD1* and *hopQ1* were not found in the
578 environmental isolates analyzed here, they may exist at low frequencies in these populations. One
579 possible emergence scenario is that when a strain receives these genes in a recombination event
580 from a donor, its fitness on crops increases, and it emerges as a highly virulent crop pathogen.
581 Surprisingly though, *hopAH1* was the only T3S effector located in one of the identified hot spots of
582 homologous recombination while *hopD1* and *hopQ1* were not. Their exclusive presence in crop
583 pathogens that are phylogenetically distinct and their absence from closely related environmental
584 isolates, is nonetheless consistent with acquisition by horizontal gene transfer. In fact this conclusion
585 is supported by inferred phylogenies based on *hopD1* and on *hopQ1* (Fig. S6) that were incongruent
586 with the core genome tree suggesting that these genes were in fact subject to recombination.

587 Among the genes found to be located within hotspots of recombination in both analyzed
588 phylogroups, and more frequently present in crop pathogen isolates than environmental isolates,
589 was the gene with locus tag PSPTO 5191/PSYR 0346. Functional predictions using BLASTP of the
590 translated sequences identified strong signatures of a conserved domain specific of the acriflavin
591 resistance protein family (Pfam accession PF00873, e-value 4.03×10^{-143} ; TIGRFAM00915, 8.04×10^{-81} ,
592 >90% length of sequence). Therefore, this putative membrane protein could be part of an
593 aminoglycoside efflux system involved in either toxin production or resistance processes. This gene
594 has been implicated in multiple antibiotic resistances in human pathogens such as *Salmonella*
595 Typhimurium (Piddock, 2006) but its potential role in conferring resistance to antibiotics and/or
596 other agropesticides in *P. syringae* is unknown. It remains the case that recombination was found to
597 be most frequent between closely related isolates with little genetic exchange between 1a and 2d
598 isolates, consistent with well-known homology dependence of recombination (Hanage et al., 2009).

599 In conclusion, the extensive comparison of multiple *P. syringae* crop pathogens and their
600 environmental relatives using several population genomics approaches clearly showed that not only
601 does *P. syringae* frequently move between crop hosts and non-agricultural environments, but also *P.*
602 *syringae* genes move just as frequently between one crop pathogen to the other as between crop
603 pathogens and environmental relatives. This suggests that most virulence genes, including T3S

604 effectors, are equally important for fitness on crops and non-crop hosts and that pathogen
605 populations in environmental reservoirs could be important sources of virulence genes for crop
606 pathogens. Intriguingly, the frequency of a small number of virulence genes, and of some genes of
607 yet unknown function, is significantly higher in crop pathogens compared to their frequency in
608 environmental relatives. Therefore, these genes appear to play a particularly important role in crop
609 disease emergence and/or for fitness in agricultural settings. These findings provide a new basis for
610 an improved understanding of crop pathogen emergence and control.

611

612

613 ACKNOWLEDGEMENTS

614

615 We are grateful to Haiji Liu for assistance in DNA extraction, Zhongmeng Bao for assistance in
616 preparation of some DNA sequencing libraries, and Caroline Guilbaud for assistance in preparing,
617 verifying, and shipping strains. We thank Susan Murray who helped running one of the
618 bioinformatics analyses. The computational calculations were done at the Human Genome Center at
619 the Institute of Medical Science (University of Tokyo, Japan) and HPC Wales (UK). Caroline L. Monteil
620 received support from INRA and the European Union, in the framework of the Marie-Curie FP7
621 COFUND People Programme, through the award of an AgreenSkills' fellowship (under grant
622 agreement n° 267196). Research in Boris A. Vinatzer's laboratory and genome sequencing was
623 funded by the National Science Foundation of the USA (grants IOS-1354215 and DEB-1241068).
624 Funding for work in the Vinatzer laboratory was also provided in part by the Virginia Agricultural
625 Experiment Station and the Hatch Program of the National Institute of Food and Agriculture, U.S.
626 Department of Agriculture. Work carried out in the Sheppard laboratory was supported by the
627 Medical Research Council (MRC) grant MR/L015080/1, and the Wellcome Trust grant
628 088786/C/09/Z. GM was supported by a NISCHR Health Research Fellowship (HF-14-13).

629

630 AUTHOR CONTRIBUTIONS

631

632 SKS, CLM, and BAV conceived and designed the overall project. CLM performed experiments. CEM
633 contributed strains. BAV and BS contributed genome sequences. DJS generated genome assemblies.
634 CLM, LM, KY, DJS and GM analyzed the data. SKS, KY and DJS contributed analytic tools. CLM, BAV,
635 and SKS wrote the paper.

636

637

638 REFERENCES

639

640 **Alfano, J. R. and Collmer, A. (2004).** Type III secretion system effector proteins: Double agents in
641 bacterial disease and plant defense. *Annual Review of Phytopathology* **42**, 385-414.

- 642 **Andrews, S. (2010).** FastQC: A quality control tool for high throughput sequence data.
643 [http://www.bioinformatics.babraham.ac.uk/projects/fastqc/] *B Bioinformatics*.
- 644 **Bartoli, C., Lamichhane, J. R., Berge, O., Guilbaud, C., Varvaro, L., Balestra, G., Vinatzer, B. A. and**
- 645 Morris, C. E. (2015). A framework to gage the epidemic potential of plant pathogens in
646 environmental reservoirs: The example of kiwifruit canker. *Mol Plant Pathol* **16**, 137-149.
- 647 **Berge, O., Monteil, C. L., Bartoli, C., Chandeysson, C., Guilbaud, C., Sands, D. C. and Morris, C. E.**
648 (2014). A user's guide to a data base of the diversity of *Pseudomonas syringae* and its
649 application to classifying strains in this phylogenetic complex. *Plos One* **9**(9).
- 650 **Block, A., Toruno, T. Y., Elowsky, C. G., Zhang, C., Steinbrenner, J., Beynon, J. and Alfano, J. R.**
651 (2014). The *Pseudomonas syringae* type III effector HopD1 suppresses effector-triggered
652 immunity, localizes to the endoplasmic reticulum, and targets the *Arabidopsis* transcription
653 factor NTL9. *New Phytol* **201**(4), 1358-70.
- 654 **Buell, C. R., Joardar, V., Lindeberg, M., Selengut, J., Paulsen, I. T., Gwinn, M. L., Dobson, R. J.,**
- 655 Deboy, R. T., Durkin, A. S., Kolonay, J. F., & other authors (2003). The complete genome
656 sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. *tomato* DC
657 3000. *Proc Natl Acad Sci USA* **100**, 10181-10186.
- 658 **Cai, R., Lewis, J., Yan, S., Liu, H., Clarke, C. R., Campanile, F., Almeida, N. F., Studholme, D. J.,**
- 659 Lindeberg, M., Schneider, D., & other authors (2011a). The plant pathogen *Pseudomonas*
660 *syringae* pv. *tomato* is genetically monomorphic and under strong selection to evade tomato
661 immunity. *PLoS Pathogens* **7**(8), e1002130.
- 662 **Cai, R., Yan, S., Liu, H., Leman, S. and Vinatzer, B. A. (2011b).** Reconstructing host range evolution of
663 bacterial plant pathogens using *Pseudomonas syringae* pv. *tomato* and its close relatives as a
664 model. *Infection Genetics and Evolution* **11**(7), 1738-1751.
- 665 **Carver, T., Thomson, N., Bleasby, A., Berriman, M. and Parkhill, J. (2009).** DNAPlotter: circular and
666 linear interactive genome visualization. *Bioinformatics* **25**(1), 119-120.
- 667 **Chen, X. M. (2005).** Epidemiology and control of stripe rust *Puccinia striiformis* f. sp *tritici* on wheat.
668 *Can J Plant Pathol* **27**(3), 314-337.
- 669 **Cheng, L., Connor, T. R., Siren, J., Aanensen, D. M. and Corander, J. (2013).** Hierarchical and
670 Spatially Explicit Clustering of DNA Sequences with BAPS Software. *Molecular Biology and*
671 *Evolution* **30**(5), 1224-1228.
- 672 **Clarke, C. R., Studholme, D. J., Hayes, B., Runde, B., Weisberg, A., Cai, R., Wroblewski, T., Daunay,**
- 673 M.-C., Wicker, E., Castillo, J. A., & other authors (2015). Genome-Enabled Phylogeographic
674 Investigation of the Quarantine Pathogen *Ralstonia solanacearum* Race 3 Biovar 2 and
675 Screening for Sources of Resistance Against Its Core Effectors. *Phytopathology* **105**(5), 597-
676 607.
- 677 **Constantinidou, H. A., Hirano, S. S., Baker, L. S. and Upper, C. D. (1990).** Atmospheric dispersal of
678 ice nucleation active bacteria - The role of rain. *Phytopathology* **80**(10), 934-937.
- 679 **Corander, J., Marttinen, P., Siren, J. and Tang, J. (2008).** Enhanced Bayesian modelling in BAPS
680 software for learning genetic structures of populations. *BMC Bioinformatics* **9**, 539.

- 681 **Corander, J., Waldmann, P., Marttinen, P. and Sillanpaa, M. J. (2004).** BAPS 2: enhanced
682 possibilities for the analysis of genetic population structure. *Bioinformatics* **20**(15), 2363-
683 2369.
- 684 **Diallo, M. D., Monteil, C. L., Vinatzer, B. A., Clarke, C. R., Glaux, C., Guilbaud, C., Desbiez, C. and**
685 **Morris, C. E. (2012).** *Pseudomonas syringae* naturally lacking the canonical type III secretion
686 system are ubiquitous in nonagricultural habitats, are phylogenetically diverse and can be
687 pathogenic. *ISME J* **6**(7), 1325-1335.
- 688 **Didelot, X. and Falush, D. (2007).** Inference of bacterial microevolution using multilocus sequence
689 data. *Genetics* **175**(3), 1251-1266.
- 690 **Feil, H., Feil, W., Chain, P., Larimer, F., DiBartolo, G., Copeland, A., Lykidis, A., Trong, S., Nolan, M.,**
691 **Goltsman, E., & other authors (2005).** Comparison of the complete genome sequences of
692 *Pseudomonas syringae* pv. *syringae* B728a and pv. *tomato* DC3000. *Proc Natl Acad Sci USA*
693 **102**, 11064-11069.
- 694 **Garland, T., Bennett, A. F. and Rezende, E. L. (2005).** Phylogenetic approaches in comparative
695 physiology. *J Exp Biol* **208**(16), 3015-3035.
- 696 **Gitaitis, R. D. and Walcott, R. (2007).** The Epidemiology and Management of Seedborne Bacterial
697 Diseases. *Annu Rev Phytopathol* **45**, 371-397
- 698 **Grosso-Becerra, M., Santos-Medellín, C., González-Valdez, A., Méndez, J., Delgado, G., Morales-**
699 **Espinosa, R., Servín-González, L., Alcaraz, L. and Soberón-Chávez, G. (2014).** *Pseudomonas*
700 *aeruginosa* clinical and environmental isolates constitute a single population with high
701 phenotypic diversity. *BMC Genomics* **15**(318).
- 702 **Hanage, W. P., Fraser, C., Tang, J., Connor, T. R. and Corander, J. (2009).** Hyper-recombination,
703 diversity, and antibiotic resistance in *Pneumococcus*. *Science* **324**(5933), 1454-1457.
- 704 **Hann, D. R., Dominguez-Ferreras, A., Motyka, V., Dobrev, P. I., Schornack, S., Jehle, A., Felix, G.,**
705 **Chinchilla, D., Rathjen, J. P. and Boller, T. (2014).** The *Pseudomonas* type III effector HopQ1
706 activates cytokinin signaling and interferes with plant innate immunity. *New Phytol* **201**(2),
707 585-598.
- 708 **Hazen, T., Lafon, P., Garrett, N., Lowe, T., Silberger, D., Rowe, L., Frace, M., Parsons, M., Bopp, C.,**
709 **Rasko, D., & other authors (2015).** Insights into the environmental reservoir of pathogenic
710 *Vibrio parahaemolyticus* using comparative genomics. *Front Microbiol* **6**(204).
- 711 **Hirano, S. S. and Upper, C. D. (2000).** Bacteria in the leaf ecosystem with emphasis on *Pseudomonas*
712 *syringae* - a pathogen, ice nucleus, and epiphyte. *Microbiol Mol Biol Rev* **64**(3), 624-653.
- 713 **Hockett, K. L., Nishimura, M., Karsrud, E., Dougherty, K. M. and Baltrus, D. A. (2014).** *P. syringae*
714 CC1557: a highly virulent strain with an unusually small type III effector repertoire that
715 includes a novel effector. *Am Phytopath Society* 10.1094/MPMI-11-13-0354-R.
- 716 **Joardar, V., Lindeberg, M., Jackson, R. W., Selengut, J., Dodson, R., Brinkac, L. M., Daugherty, S. C.,**
717 **DeBoy, R., Durkin, A. S., Giglio, M. G., & other authors (2005).** Whole-genome sequence
718 analysis of *Pseudomonas syringae* pv. *phaseolicola* 1448A reveals divergence among
719 pathovars in genes involved in virulence and transposition. *J Bacteriol* **187**(18), 6488-6498.
- 720 **Johnson, P. T. J., De Roode, J. C. and Fenton, A. (2015).** Why infectious disease research needs
721 community ecology. *Science* **349**(6252).

- 722 **Jolley, K. A. and Maiden, M. C. J. (2010).** BIGSdb: Scalable analysis of bacterial genome variation at
723 the population level. *Bmc Bioinformatics* **11**, 595.
- 724 **Katoh, K. and Toh, H. (2008).** Recent developments in the MAFFT multiple sequence alignment
725 program. *Briefings in Bioinformatics* **9**(4), 286-298.
- 726 **Krueger, F. (2015).** Trim Galore!: A wrapper tool around Cutadapt and FastQC to consistently apply
727 quality and adapter trimming to FastQ files.
728 [http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/].
- 729 **Lawson, D. J., Hellenthal, G., Myers, S. and Falush, D. (2012).** Inference of population structure
730 using dense haplotype data. *Plos Genet* **8**(1).
- 731 **Li, W., Chiang, Y.-H. and Coaker, G. (2013a).** The HopQ1 effector's nucleoside hydrolase-like domain
732 is required for bacterial virulence in *Arabidopsis* and tomato, but not host recognition in
733 tobacco. *Plos One* **8**(3).
- 734 **Li, W., Yadeta, K. A., Elmore, J. M. and Coaker, G. (2013b).** The *Pseudomonas syringae* Effector
735 HopQ1 Promotes Bacterial Virulence and Interacts with Tomato 14-3-3 Proteins in a
736 Phosphorylation-Dependent Manner. *Plant Physiology* **161**(4), 2062-2074.
- 737 **Lindeberg, M., Myers, C. R., Collmer, A. and Schneider, D. J. (2008).** Roadmap to new virulence
738 determinants in *Pseudomonas syringae*: Insights from comparative genomics and genome
739 organization. *Mol Plant Microbe In* **21**(6), 685-700.
- 740 **Maiden, M. C. J., van Rensburg, M. J. J., Bray, J. E., Earle, S. G., Ford, S. A., Jolley, K. A. and
741 McCarthy, N. D. (2013).** MLST revisited: the gene-by-gene approach to bacterial genomics.
742 *Nature Reviews Microbiology* **11**(10), 728-736.
- 743 **Martin, M. (2011).** Cutadapt removes adapter sequences from high-throughput sequencing reads.
744 *EMBnet.journal* **17**(1).
- 745 **Martins, E. P. and Garland, T. (1991).** Phylogenetic analyses of the correlated evolution of
746 continuous characters: A simulation study. *Evolution* **45**(3), 534-557.
- 747 **McCann, H. C., Rikkerink, E. H. A., Bertels, F., Fiers, M., Lu, A., Rees-George, J., Andersen, M. T.,
748 Gleave, A. P., Haubold, B., Wohlers, M. W., & other authors (2013).** Genomic analysis of the
749 kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae* provides insight into the origins of
750 an emergent plant disease. *PLoS Pathogens* **9**(7), e1003503.
- 751 **McCartter, S. M., Jones, J. B., Gitaitis, R. D. and Smitley, D. R. (1983).** Survival of *Pseudomonas*
752 *syringae* pv *tomato* in association with tomato seed, soil, host tissue, and epiphytic weed
753 hosts in Georgia. *Phytopathology* **73**(10), 1393-1398.
- 754 **Meric, G., Yahara, K., Mageiros, L., Pascoe, B., Maiden, M. C. J., Jolley, K. A. and Sheppard, S. K.
755 (2014).** A reference pan-genome approach to comparative bacterial genomics: Identification
756 of novel epidemiological markers in pathogenic *Campylobacter*. *PLoS One* **9**(3).
- 757 **Mohr, T. J., Liu, H., Yan, S., Morris, C. E., Castillo, J. A., Jelenska, J. and Vinatzer, B. A. (2008).**
758 Naturally occurring nonpathogenic isolates of the plant pathogen *Pseudomonas syringae*
759 lack a type III secretion system and effector gene orthologues. *J Bacteriol* **190**(8), 2858-2870.

- 760 **Monteil, C. L., Bardin, M. and Morris , C. E. (2014a).** Features of air masses associated with the
761 deposition of *Pseudomonas syringae* and *Botrytis cinerea* by rain and snowfall. *ISME J*
762 doi:10.1038/ismej.2014.55, 1-15.
- 763 **Monteil, C. L., Cai, R., Liu, H., Mechan Llontop, M. E., Leman, S., Studholme, D. J., Morris, C. E. and**
764 **Vinatzer, B. A. (2013).** Non-agricultural reservoirs contribute to emergence and evolution of
765 *Pseudomonas syringae* crop pathogens. *New Phytol* **199**(3), 800–811.
- 766 **Monteil, C. L., Guilbaud, C., Glaux, C., Lafolie, F., Soubeyrand, S. and Morris, C. E. (2012).**
767 Emigration of the plant pathogen *Pseudomonas syringae* from leaf litter contributes to its
768 population dynamics in alpine snowpack. *Environ Microbiol* **14**(8), 2099-112.
- 769 **Monteil, C. L., Lafolie, F., Laurent, J., Clement, J. C., Simler, R., Travi, Y. and Morris, C. E. (2014b).**
770 Soil water flow is a source of the plant pathogen *Pseudomonas syringae* in subalpine
771 headwaters *Environ Microbiol* **16**(7), 2038–2052.
- 772 **Morris, C. E., Glaux, C., Latour, X., Gardan, L., Samson, R. and Pitrat, M. (2000).** The relationship of
773 host range, physiology, and genotype to virulence on cantaloupe in *Pseudomonas syringae*
774 from cantaloupe blight epidemics in France. *Phytopathology* **90**(6), 636-646.
- 775 **Morris, C. E., Monteil, C. L. and Berge, O. (2013).** The life history of *Pseudomonas syringae*: Linking
776 agriculture and Earth system processes. *Annu Rev Phytopathol* **51**, 85-104.
- 777 **Morris, C. E., Sands, D. C., Vanneste, J. L., Montarry, J., Oakley, B., Guilbaud, C. and Glaux, C.**
778 **(2010).** Inferring the evolutionary history of the plant pathogen *Pseudomonas syringae* from
779 its biogeography in headwaters of rivers in North America, Europe, and New Zealand. *mBio*
780 **1**(3), 1 - 10.
- 781 **Morris, C. E., Sands, D. C., Vinatzer, B. A., Glaux, C., Guilbaud, C., Buffiere, A., Yan, S. C.,**
782 **Dominguez, H. and Thompson, B. M. (2008).** The life history of the plant pathogen
783 *Pseudomonas syringae* is linked to the water cycle. *ISME J* **2**(3), 321-334.
- 784 **O'Brien, H. E., Thakur, S. and Guttman, D. S. (2011).** Evolution of plant pathogenesis in
785 *Pseudomonas syringae*: A genomics perspective. *Annu Rev Phytopathol* **49**, 269-289.
- 786 **Ochman, H., Lawrence, J. G. and Groisman, E. A. (2000).** Lateral gene transfer and the nature of
787 bacterial innovation. *Nature* **405**(6784), 299-304.
- 788 **Pascoe, B., Meric, G., Murray, S., Mageiros, L., Yahara, K., Bowen, R., Jones, N. H., Jeeves, R. E.,**
789 **Lappin-Scott, H. M., H., A., & other authors (2015).** Enhanced biofilm formation evolves
790 from divergent genetic backgrounds in host generalist *Campylobacter jejuni*. *Environ*
791 *Microbiol* 10.1111/1462-2920.13051.
- 792 **Piddock, L. V. (2006).** Multidrug-resistance efflux pumps — not just for resistance. *Nat Rev Microbiol*
793 **4**, 629-636.
- 794 **Price, M. N., Dehal, P. and Arkin, A. P. (2010).** FastTree 2 - Approximately maximum-likelihood trees
795 for large alignments. *PLoS One* **3**, e9490.
- 796 **Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A. and Barrell, B. (2000).**
797 Artemis: Sequence visualization and annotation. *Bioinformatics* **16**(10), 944-945.
- 798 **Sheppard, S. K., Didelot, X., Meric, G., Torralbo, A., Jolley, K. A., Kelly, D. J., Bentley, S. D., Maiden,**
799 **M. C. J., Parkhill, J. and Falush, D. (2013).** Genome-wide association study identifies vitamin

- 800 B-5 biosynthesis as a host specificity factor in *Campylobacter*. *Proc Natl Acad Sci USA*
801 **110**(29), 11923-11927.
- 802 **Sheppard, S. K., Jolley, K. A. and Maiden, M. C. J. (2012)**. A gene-by-gene approach to bacterial
803 population genomics: Whole genome MLST of *Campylobacter*. *Genes* **3**(2), 261-277.
- 804 **Singh, R. P., Hodson, D. P., Huerta-Espino, J., Jin, Y., Bhavani, S., Njau, P., Herrera-Foessel, S.,**
805 **Singh, P. K., Singh, S. and Govindan, V. (2011)**. The emergence of Ug99 races of the stem
806 rust fungus is a threat to world wheat production. *Annu Rev Phytopathol* **49**, 465-481.
- 807 **Struve, C. and Kogfelt, K. (2004)**. Pathogenic potential of environmental *Klebsiella pneumoniae*
808 isolates. *Environ Microbiol* **6**(6), 584-590.
- 809 **Stukenbrock, E. H. and McDonald, B. A. (2008)**. The origins of plant pathogens in agro-ecosystems.
810 *Annu Rev Phytopathol* **46**, 75-100.
- 811 **Tampakaki, A. P., Skandalis, N., Gazi, A. D., Bastaki, M. N., Sarris, P. F., Charova, S. N., Kokkinidis,**
812 **M. and Panopoulos, N. J. (2011)**. Playing the "Harp": evolution of our understanding of
813

hrp/hrc genes. Annu Rev Phytopathol **48**, 347-370.
- 814 **Vinatzer, B. A., Monteil, C. L. and Clarke, C. R. (2014)**. Harnessing population genomics to
815 understand how bacterial pathogens emerge, adapt to crop hosts and disseminate. *Annu*
816 *Rev Phytopathol* **52**.
- 817 **Wei, C.-F., Kvittko, B. H., Shimizu, R., Crabbill, E., Alfano, J. R., Lin, N.-C., Martin, G. B., Huang, H.-C.**
818 **and Collmer, A. (2007)**. A *Pseudomonas syringae* pv. *tomato* DC3000 mutant lacking the
819 type III effector HopQ1-1 is able to cause disease in the model plant *Nicotiana benthamiana*.
820 *Plant Journal* **51**(1), 32-46.
- 821 **Whiley, H., van den Akker, B., Giglio, S. and R., B. (2013)**. The role of environmental reservoirs in
822 human campylobacteriosis. *int J Envirion Res Public Helth* **10**(11), 5886-5907.
- 823 **Woolhouse, M. E. J., Taylor, L. H. and Haydon, D. T. (2001)**. Population biology of multihost
824 pathogens. *Science* **292**(5519), 1109-1112.
- 825 **Xin, X.-F. and He, S. Y. (2013)**. *Pseudomonas syringae* pv. *tomato* DC3000: A model pathogen for
826 probing disease susceptibility and hormone signaling in plants. *Annu Rev Phytopathol* **51**,
827 473-498.
- 828 **Yahara, K., Didelot, X., Ansari, M. A., Sheppard, S. K. and Falush, D. (2014)**. Efficient inference of
829 recombination hot regions in bacterial genomes. *Mol Biol Evol* **31**(6), 1593-1605.
- 830 **Yahara, K., Didelot, X., Jolley, K. A., Kobayashi, I., Maiden, M. C. J., Sheppard, S. K. and Falush, D.**
831 **(2016a)**. The landscape of realized homologous recombination in pathogenic bacteria. *Mol*
832 *Biol Evol* **33**(2), 456-471.
- 833 **Yahara, K., Furuta, Y., Oshima, K., Yoshida, M., Azuma, T., Hattori, M., Uchiyama, I. and Kobayashi,**
834 **I. (2013)**. Chromosome painting in silico in a bacterial species reveals fine population
835 structure. *Molecular Biology and Evolution* **30**(6), 1454-1464.
- 836 **Yahara, K., G., M. r., Taylor, A., de Vries, S., Murray, S., Pascoe, B., Mageiros, L., Torralbo, A., Vidal,**
837 **A., Ridley, A., & other authors (2016b)**. Genome-wide association of functional traits linked
838 with *Campylobacter jejuni* survival from farm to fork. *PeerJ Preprints* **4**, e2300v1.

839 Yan, S., Liu, H., Mohr, T. J., Jenrette, J., Chiodini, R., Zaccardelli, M., Setubal, J. C. and Vinatzer, B.
840 A. (2008). Role of recombination in the evolution of the model plant pathogen *Pseudomonas*
841 *syringae* pv. *tomato* DC3000, a very atypical tomato strain. *Applied and Environmental*
842 *Microbiology* **74**(10), 3171-3181.

843 Zerbino, D. and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn
844 graphs. *Genome Res* **18**(5), 821-829.

845

846

847 **DATA BIBLIOGRAPHY**

848

849 Studholme, D., Monteil C., Swingle B. and Vinatzer, B. A. NCBI, *Pseudomonas syringae* BioProject
850 (ID 320409). (2016)

851

852

853

854 **FIGURES AND TABLES**

855

856 **Fig. 1. Population structure of 193 isolates from the *P. syringae* species complex.** (a) An unrooted
857 phylogenetic tree was constructed from 1889 genes detected in all isolates using an approximation
858 of the maximum-likelihood algorithm and the GTR model for nucleotide substitution. The scale
859 represents the number of substitutions per site. Isolates are colored according to monophyletic
860 groups described by Berge et al. (2014). Panel (b) represents genealogies inferred by ClonalFrame
861 (Didelot and Falush, 2007) for phylogroups 1a and 2d based only on those genes with no evidence
862 of homologous recombination (810 and 2293 genes, respectively). The trees are drawn to scale,
863 with branch lengths corresponding to the number of substitutions per site. Dark and light grey
864 circles symbolize isolation sources: crop pathogens and environmental strains, respectively. Solid
865 and dotted branches symbolize subpopulations within each phylogroup determined with the first
866 level of HierBAPS hierarchy (Corander et al., 2008, Cheng et al., 2013).

867

868 **Fig. 2. Core genome similarity and accessory genome variation within and between phylogroups**
869 **2d and 1a.** (a) Overlap between the core and accessory genomes calculated for the 1a and 2d
870 phylogroups. Core genes (Core), accessory genes (Acc) and absent genes (Abs) were defined as
871 genes detected in 100% of the isolates, detected in less than 100% but more than 0% of the
872 isolates, and absent in 100% of the isolates, respectively. The radius of each circle is proportional
873 to the number of detected genes. Dark gray circles represent genes from strains isolated from
874 diseased crops while light gray circles represent genes from environmental isolates. White circles
875 represent overlapping genes from both populations. Matrices show pairwise comparison of core
876 genome similarity (b) and accessory genome variation (c) between 104 isolates ordered according
877 to the phylogenetic tree presented in Fig. 1. Entire matrices with all 193 *P. syringae* isolates are

878 shown in Fig. S5. Dark grey boxes symbolize crop pathogens while isolates from other sources are
879 symbolized by light grey boxes. Heat-map colors ranging from white, through yellow, red to black
880 represent values from the lowest to the highest number of shared alleles or genes, in the first and
881 second heatmap respectively. The minimum number of shared alleles in the core genome ranged
882 from 0 to 3010 while the minimum number of shared accessory genes ranged from 1105 to 2508.

883

884 **Fig. 3. Distribution of genes associated with *P. syringae* crop pathogen strains within reference**
885 **genomes and within populations.** (a) 30-bp long words that were found to be significantly
886 associated with crop pathogens in either phylogroup 1a or 2d were mapped on the genome of the
887 reference crop pathogens *Pto* DC3000 and *Psy* B728a, respectively. A total of 73,299 and 5,970
888 crop pathogen-associated words in *Pto* DC3000 and *Psy* B728a were distributed in 571 and 222
889 genes, respectively. The list of these genes and the distribution of mapped words is given in the
890 Table S3. For each chromosome and plasmid, the first grey circle represents virulence genes (listed
891 in Lindeberg *et al.* (2008)). The four next colored circles correspond to the genes in which words
892 were mapped from the highest to the lowest p-value (blue, green, orange and red corresponding
893 to p-value cutoffs of 5.10^{-4} , 5.10^{-5} , 5.10^{-6} and 5.10^{-7} respectively). The 25kb-region containing
894 *hopQ1* and *hopD1* described in the text is indicated with a * (b) Heatmaps show the presence of
895 genes for which at least one word was significantly associated with crop pathogens with a
896 probability inferior or equal to 5.10^{-7} . Black boxes denote that at least one word was mapped for
897 the corresponding isolate and gene, while white boxes denote that not a single word was present
898 with this probability. Isolates were organized following the core genome trees built with
899 ClonalFrame (Didelot and Falush, 2007) in Fig. 1b. The trees are drawn to scale, with branch
900 lengths proportional to the number of substitutions per site. Dark and light grey circles symbolize
901 crop pathogen strains and environmental isolates, respectively.

902

903 **Fig. 4. Inference of genetic fluxes within *P. syringae* populations.** Co-ancestry matrices were
904 determined by Chromosome Painting and fineSTRUCTURE for phylogroups 1a (a) and 2d (b). The
905 expected number of “chunks” imported from a donor genome (column) to a recipient genome
906 (row) is given by the color of each cell of the matrices. The trees are drawn to scale, with branch
907 lengths proportional to the number of substitutions per site. Arrows in the first panel point to *Pto*
908 DC3000, while arrows in the second panel point to *Pap* CC94 and *Psy* B728a at the top and bottom
909 of the tree respectively. Dark and light grey circles symbolize crop pathogen strains and
910 environmental isolates, respectively.

911

912 **Fig. 5. Inference of homologous recombination hot spots within *P. syringae* populations.**
913 Homologous recombination hot spots were inferred in phylogroups 1a (a) and 2d (b) as described
914 in Yahara *et al.* (2014). A total of 177,790 SNPs and 141,997 SNPs were used respectively
915 representing the chromosome of each reference genome. For each group, the extent of
916 recombination was estimated from the whole phylogroup, from the crop pathogen strains only
917 and, finally, considering only isolates from environmental reservoirs. The X-axis indicates the
918 position in the reference genomes *Pto* DC3000 (Buell *et al.*, 2003) and *Psy* B728a (Feil *et al.*, 2005).
919 The Y-axis indicates the empirical distribution of the distance statistic H_i representing the intensity
920 of normalized recombination. The solid line represents the average value of H_i in a genome The
921 dotted lines represent the top and bottom 2.5 percentiles. Grey lines represent virulence genes

922 (Lindeberg et al., 2008). Turquoise lines represent genes associated with being a crop pathogen
923 using the GWAS approach (Table S3). Orange lines represent genes associated with hot spots of
924 recombination. The genes in the regions showing more intense recombination are given in Table
925 S5.

926

927 **SUPPLEMENTARY FIGURE LEGENDS**

928

929 **Figure S1. Approximate Maximum-Likelihood tree of the entire *Pseudomonas* genus**
930 **reconstructed from concatenated ribosomal protein gene sequences (rMLST scheme proposed**
931 **by Jolley et al. (Jolley et al., 2012)).** The analysis involved 629 sequences from *Pseudomonas*
932 genomes with 53 tagged ribosomal protein genes as described in (Jolley et al., 2012). The tree was
933 rooted with the Gamma proteobacterium *Francisella* strain TX077308. In addition to all *P. syringae*
934 genomes used in this study, data of finished and draft genomes of 436 bacterial isolates from the
935 *Pseudomonas* genus were downloaded from the NCBI database (<http://www.ncbi.nlm.nih.gov>) and
936 uploaded to a BIGSdb database (available at <http://rmlst.org/>). Each color/number corresponds to
937 a cluster determined by the first level of HierBAPS and show how species are structured within the
938 genus. List of isolates and their species, cluster, and accession number can be found in Table S2.
939 While some groups look monophyletic (i.e. *P. aeruginosa*), others are polyphyletic, like *P.*
940 *fluorescens*.

941

942 **Figure S2. NeighborNet showing the effect of intraspecific homologous recombination on tree**
943 **clade structure for *P. syringae* crop pathogen strains and environmental isolates.** The split
944 networks were built using the Neighbor-net method implemented in SplitTree4 (Huson and Bryant,
945 2006) and the concatenated sequences of 117 and 129 recombining core genes identified with
946 ClonalFrame for the 1a (a) and 2d (b) phylogroups, respectively. The trees are drawn to scale, with
947 branch lengths measured in the number of substitutions per site. Dark and light grey circles
948 symbolize isolation sources (diseased crops and environment, respectively).

949

950 **Figure S3. Genetic variation in sampled 1a and 2d strains.** (a) Nucleotide divergence of *P. syringae*
951 strains based on their core genomes and estimated from the individual alignment of genes within
952 each population (“Dis”; for crop pathogens, and “Env”; for environmental relatives). The mean of
953 nucleotide divergence with its standard error is presented next to each boxplot. Upper case values
954 and lower case values are assigned to a different comparison. Values associated with the same
955 letter are not significantly different (MWU test, P<0.05). It is important to note that the 2d
956 phylogroup is mainly composed of clones but some strains significantly raised the average p-
957 distance. (b) Allelic variation of core genes shared by crop pathogens and relatives isolated from
958 the water cycle for 1a and 2d phylogroups. The core genome was estimated from the alignment of
959 the reference strains *Pto* DC3000 and *Psy* B728a for each population, respectively. The ratio of the
960 number of unique allelic variants and the number of isolates was calculated for each of the 3,576
961 and 4,147 core genes of the two phylogroups (diseased plants or water cycle). Grey points above
962 the proportionality line (solid black line) represent genes that have more unique alleles per isolate
963 in crop pathogen populations than in environmental populations and vice versa.

964

965 **Figure S4. Frequency of genes coding for effector proteins in crop pathogens and environmental**
966 **isolates.** Sequences of effector genes were downloaded from the PPI website
967 (<http://pseudomonas-syringae.org>) and are given in Table S4. We kept multiple alleles of the same
968 gene when BLAST showed considerable differences in frequencies of occurrence within groups
969 (e.g. HopF2, AvrRpm1). A gene was considered present if its original sequence aligned over a
970 minimum length of 50% with 70% identity. Left and right barplots represent frequencies for 1a and
971 2d phylogroups, respectively. Dark gray bars represent crop pathogen strains while light gray bars
972 represent isolates from water, rain, snow, river or irrigation water. Stars next to the bars highlight
973 genes totally absent in the environment while present in crop pathogens. Genes present in all
974 isolates of both populations in at least one of the two groups are marked with a black border.
975 Numbers next to the bars show the ratio between the number of unique alleles and the number of
976 isolates.

977

978 **Figure S5. Pairwise core genome similarity and accessory genome variation in the *P. syringae***
979 **species complex.** Matrices show pairwise comparison between 193 isolates ordered according to
980 the phylogenetic tree rooted with the isolate UB246 and presented in Fig. 1. (a) Core genome
981 similarity is based upon the number of shared alleles at 1903 loci from the *P. syringae* reference
982 pan-genome detected in all isolates. Truncated sequences were not considered as alleles. (b)
983 Accessory genome similarity based upon gene presence or absence at 5581 non-core loci of the
984 pan-genome. Truncated genes are considered as present. Red labels symbolize crop pathogen
985 strains while blue labels symbolize isolates from other sources. Heat-map colors ranging from
986 white, through yellow, red to black represent values from the lowest to the highest number of
987 shared alleles and genes, in the first and second panels, respectively. The minimum number of
988 shared alleles in the core genome ranged from 0 to 1892 while the minimum number of shared
989 accessory genes ranged from 1144 to 3667.

990

991 **Figure S6. Congruence between evolutionary histories of 23 G1a crop pathogens based on *hopD1***
992 **and *hopQ1* and their core gene sequences.** (a) The maximum likelihood phylogenetic tree was
993 built with RAxML 8.2.6 (Stamatakis, 2014) from the concatenation of 4048 core gene sequences
994 using the General Time reversible model for nucleotide substitution and a gamma correction for
995 the distribution of rates among sites. (b) The evolutionary history was inferred in MEGA7 (Kumar
996 et al., 2016) from the concatenation of *hopD1* and *hopQ1* sequences by using the Maximum
997 Likelihood method based on the Tamura 3-parameter model using a gamma distribution of rates
998 among sites and allowing for some sites to be evolutionarily invariable. Trees are drawn to scale,
999 with branch lengths proportional to the number of substitutions per site. The analysis involved 23
1000 nucleotide sequences. All positions containing gaps and missing data were eliminated. The HGT-
1001 Detection program of the Trex platform (<http://www.trex.uqam.ca/>) supported identified
1002 discrepancies between trees identifying by iteration 13 events of horizontal gene transfer for the
1003 *hop* genes.

1004

1005 **SUPPLEMENTARY TABLE LEGENDS**

1006

1007 **Table S1.** *P. syringae* isolates and genomes used in this study sorted by phylogroup according to
1008 the nomenclature proposed by Berge et al. (2014)

1009

1010 **Table S2.** Genomes used in the rMLST tree representing the relationships between all genomes
1011 within the *Pseudomonas* genus. *P. syringae* genomes are listed in Table S1. All *P. syringae* isolates
1012 are in the HierBAPS cluster 4 (FigureS1) with the exception of the isolate UB246.

1013

1014 **Table S3.** List of genes associated with *P. syringae* 1a and 2d crop pathogens. A total of 75388
1015 and 5968 words (each 30pb long) were identified in *Pto* DC3000 and *Psy* B728a as associated to
1016 crop pathogens in the 1a and 2d phylogroups, respectively. These words were selected based on
1017 the probability threshold of 10^{-4} and were distributed in 571 and 222 genes, respectively. Genes
1018 are organized as a function of the number of words and their probability from the highest number
1019 of words and lower p-value to the lowest number of words and highest p-value. A brief description
1020 of the gene is also given.

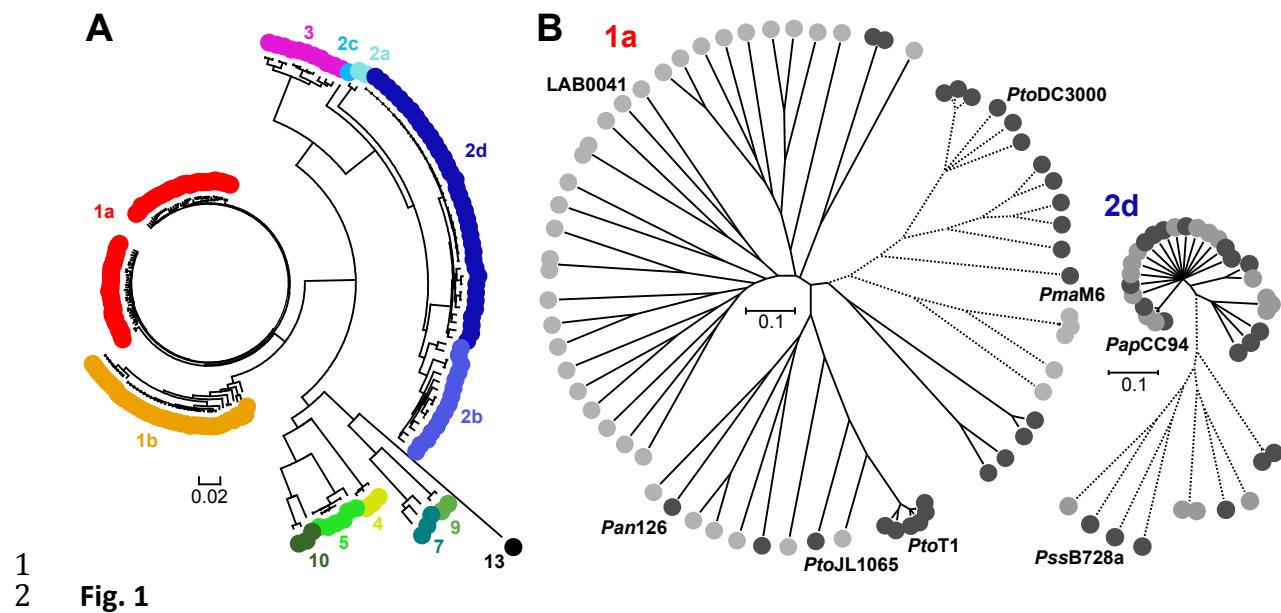
1021

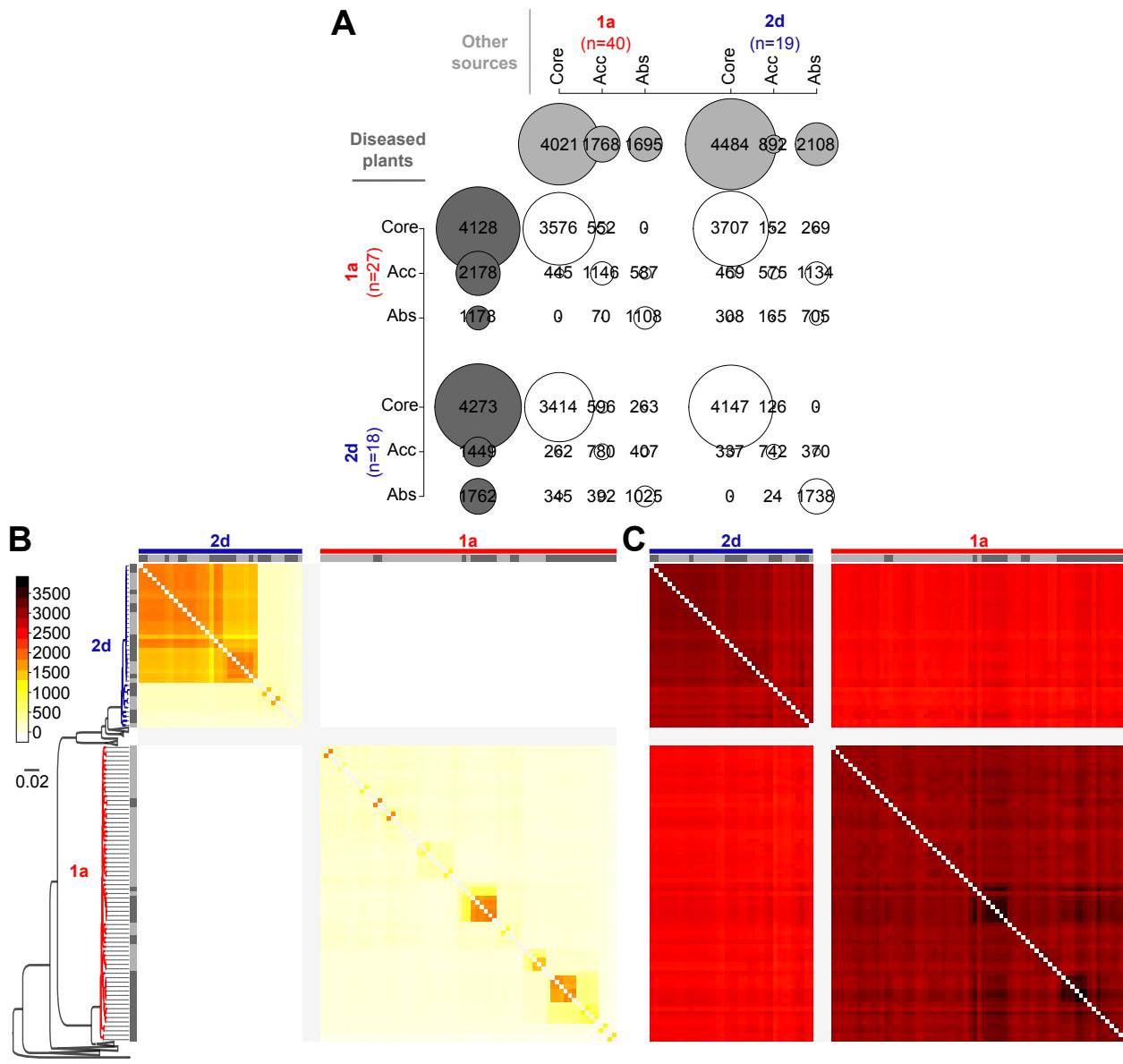
1022 **Table S4.** Repertoire of genes coding for Hop effector proteins blasted against genomes.
1023 Sequences were downloaded from the PPI website (<http://pseudomonas-syringae.org>, July 2014)

1024

1025 **Table S5.** List of *Pto* DC3000 and *Psy* B728a chromosomal genes showing intense homologous
1026 recombination. A total of 54 and 103 genes were associated with hot spots of recombination for
1027 phylogroups 1a and 2d, respectively. Genes were organized from the highest to the lowest value of
1028 recombination intensity (H_i) within each group. Information is provided if the gene was associated
1029 with the isolation source as well by GWAS. For phylogroup 1a, no recombining hot region matched
1030 genes associated with being a crop pathogen. For phylogroup 2d however, 12.6% of the genes in
1031 hot spots of recombination were associated with genes identified in the GWAS approach, but not
1032 with those that had the highest significance. The third spreadsheet lists orthologs genes associated
1033 to hot regions for recombination in both phylogroups.

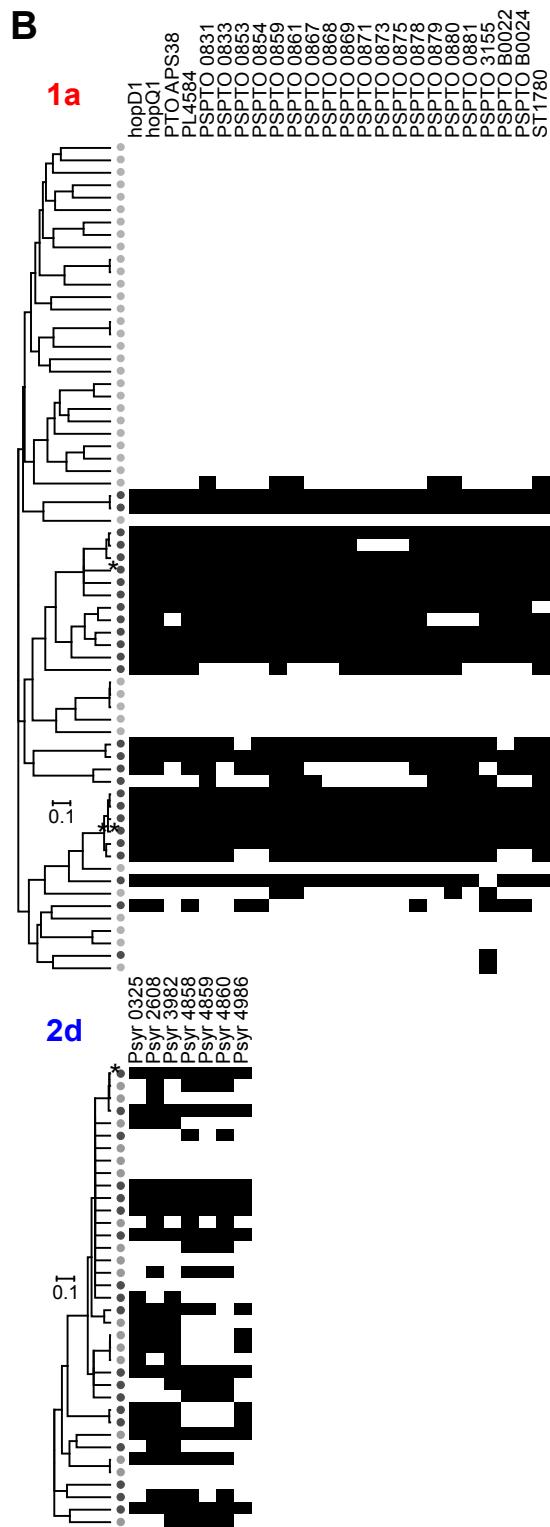
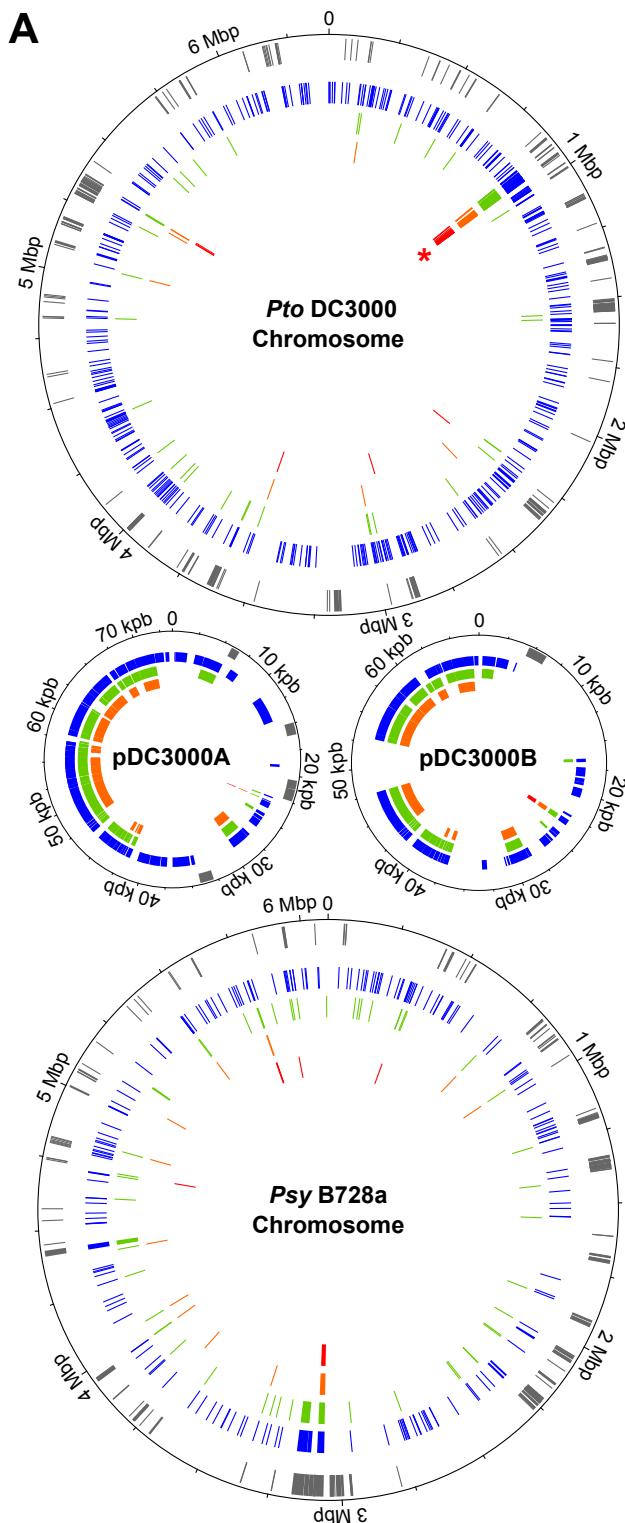
1034





3
4

Fig. 2



5
6

Fig. 3

7
8

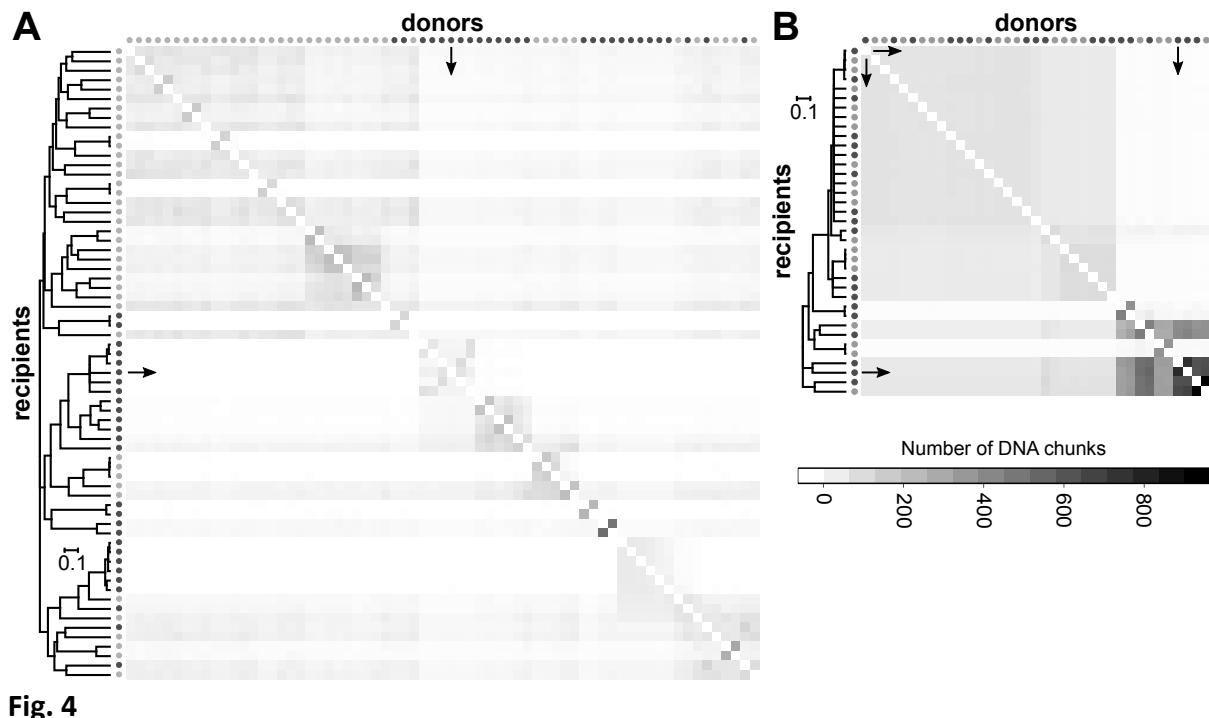
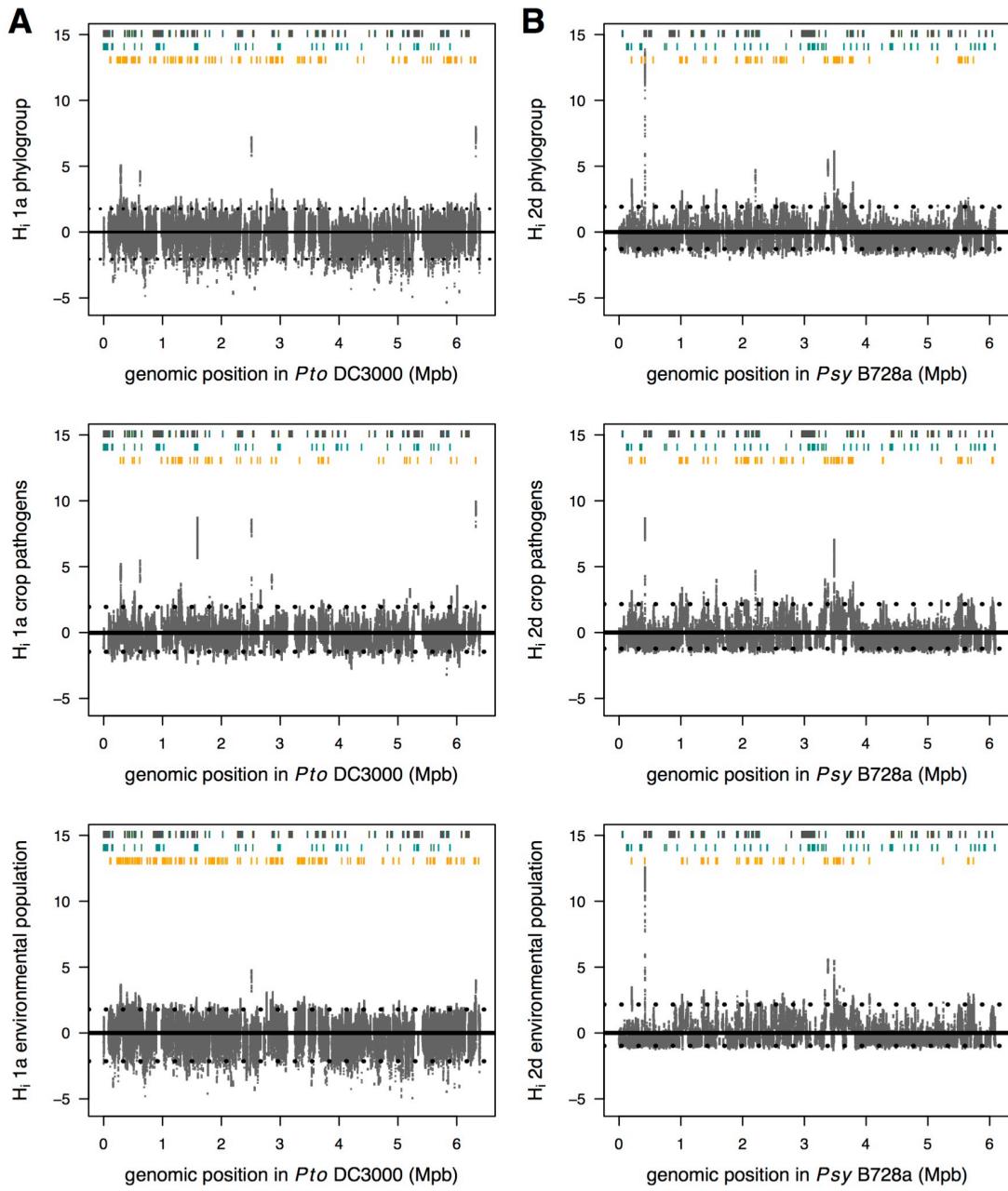
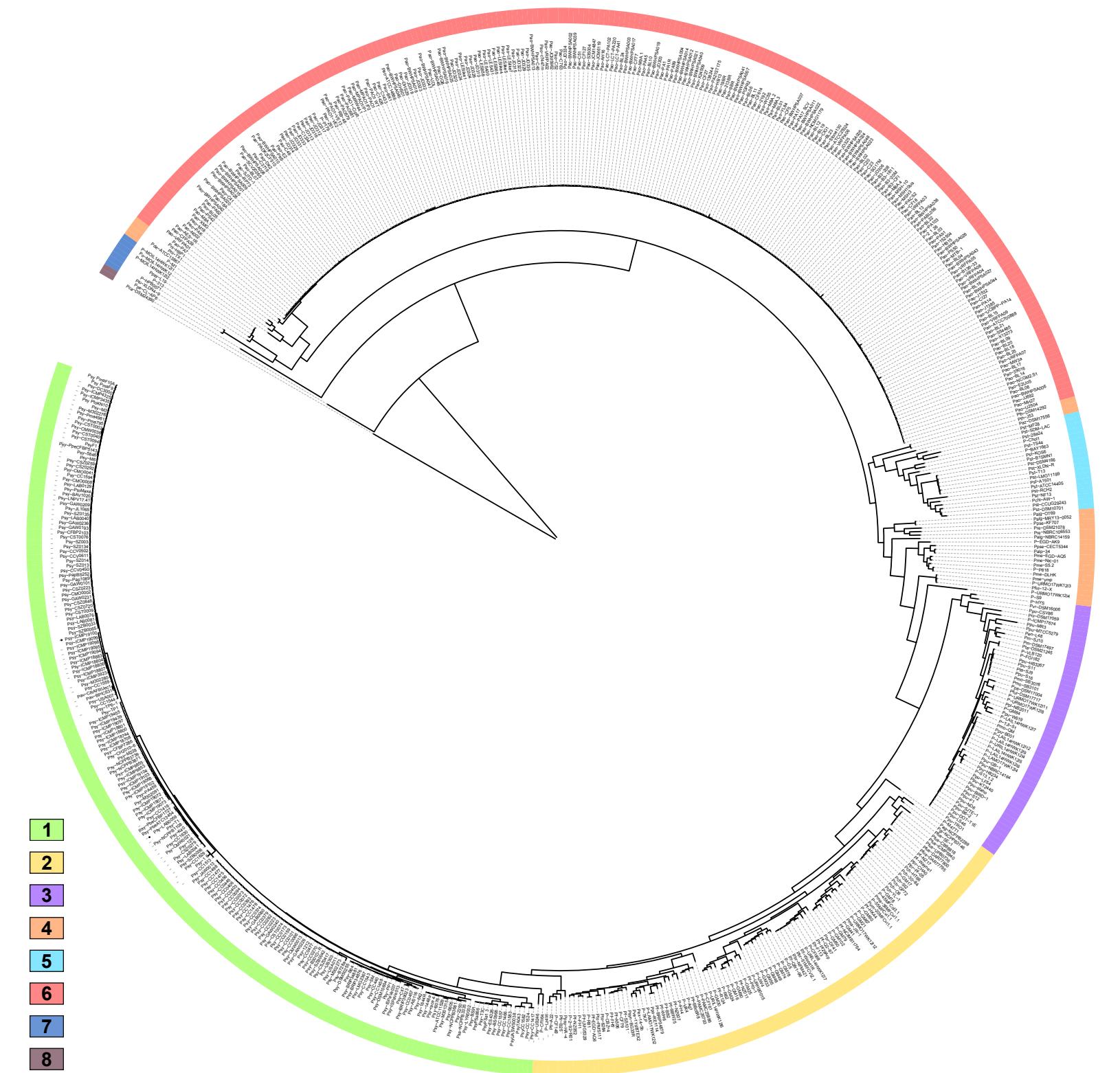


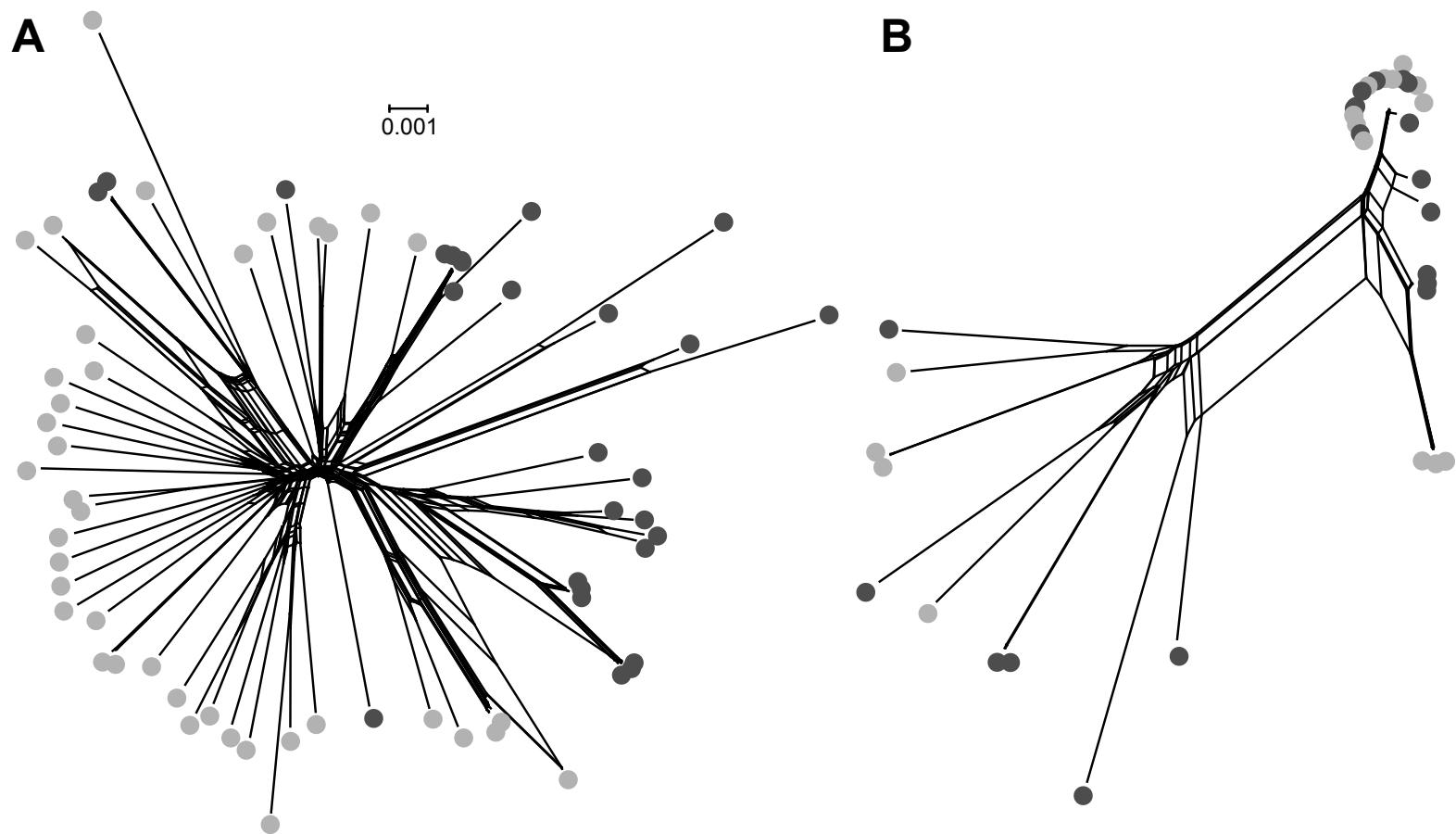
Fig. 4

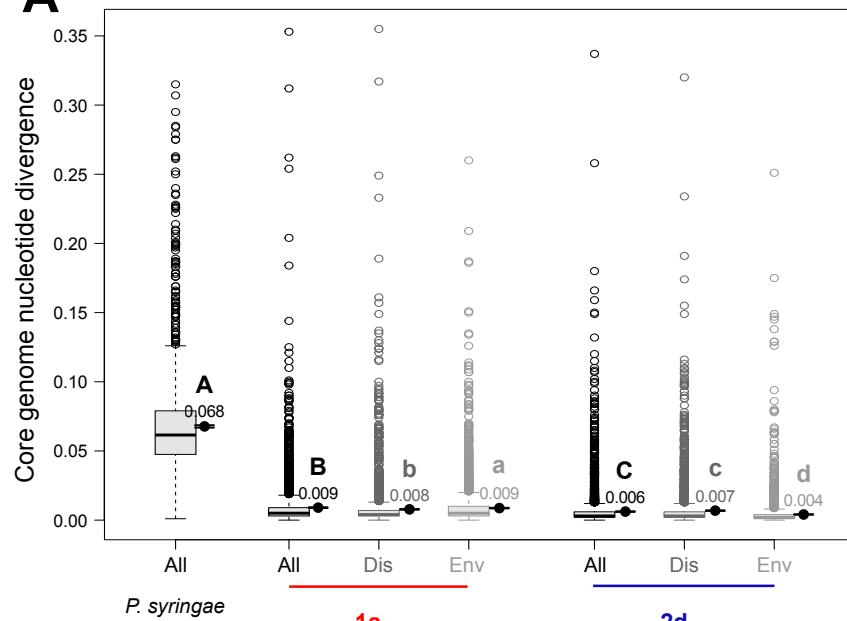
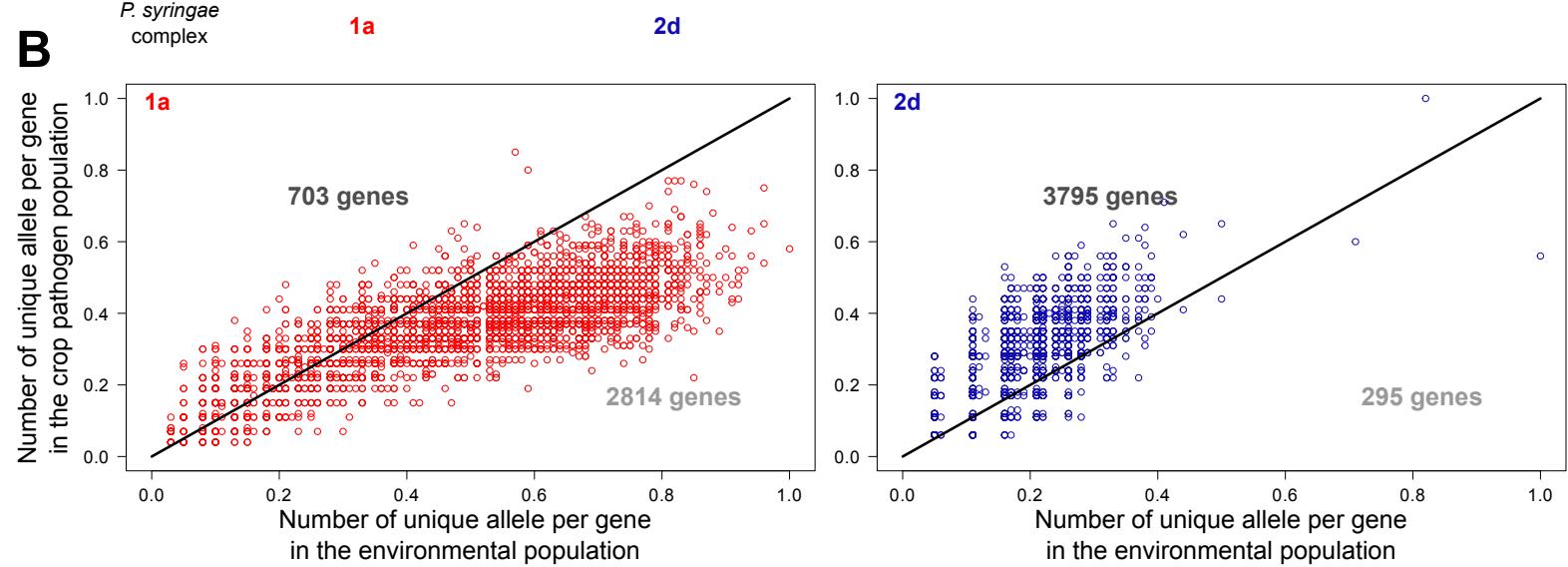


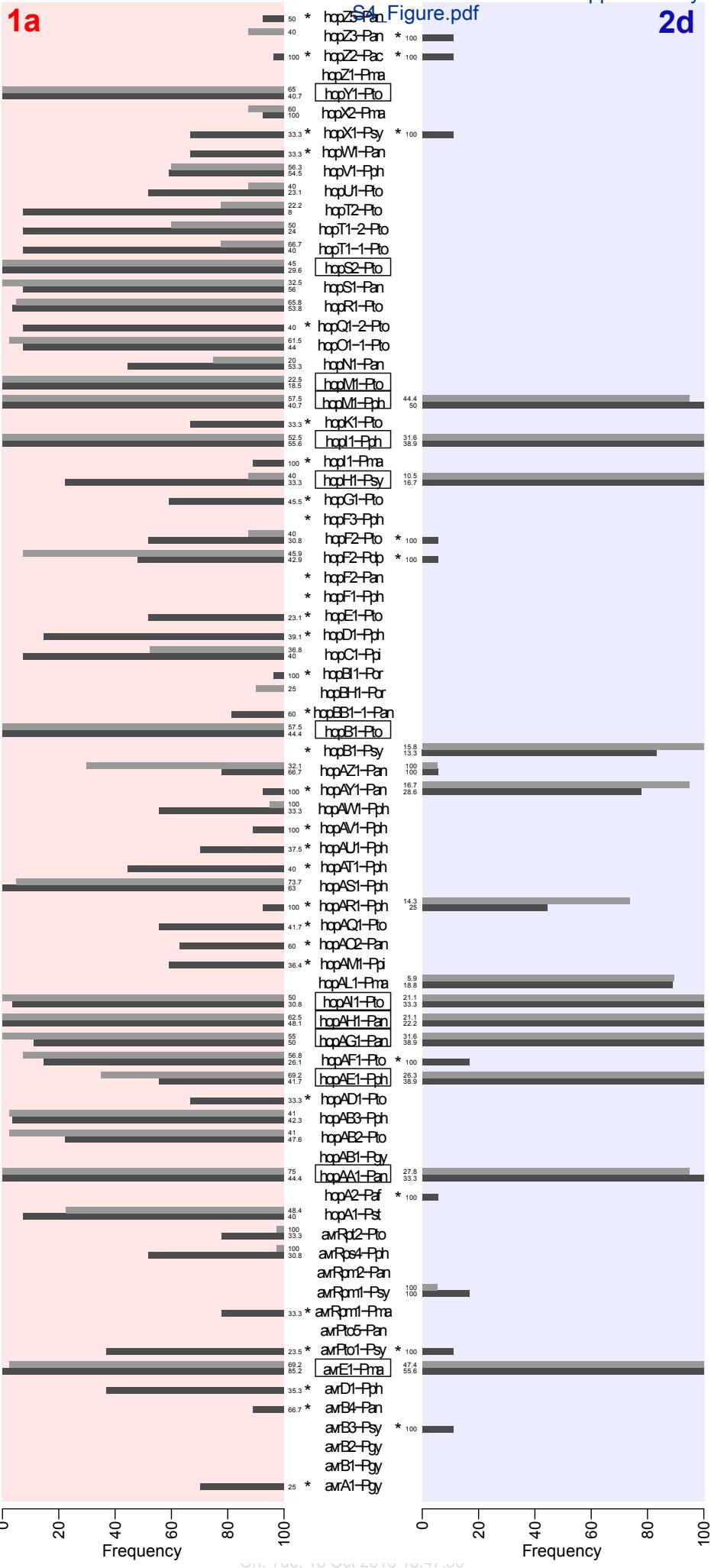
9
10
11
12

Fig. 5



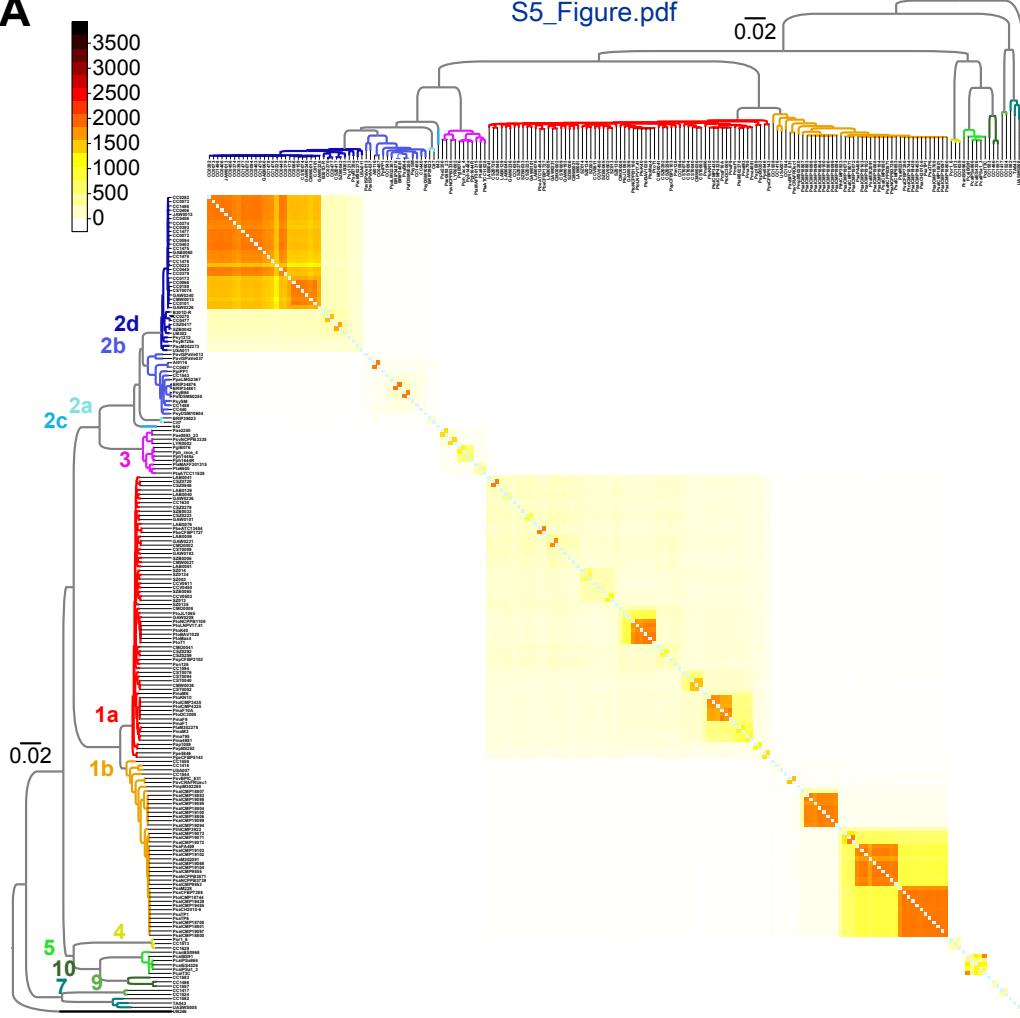
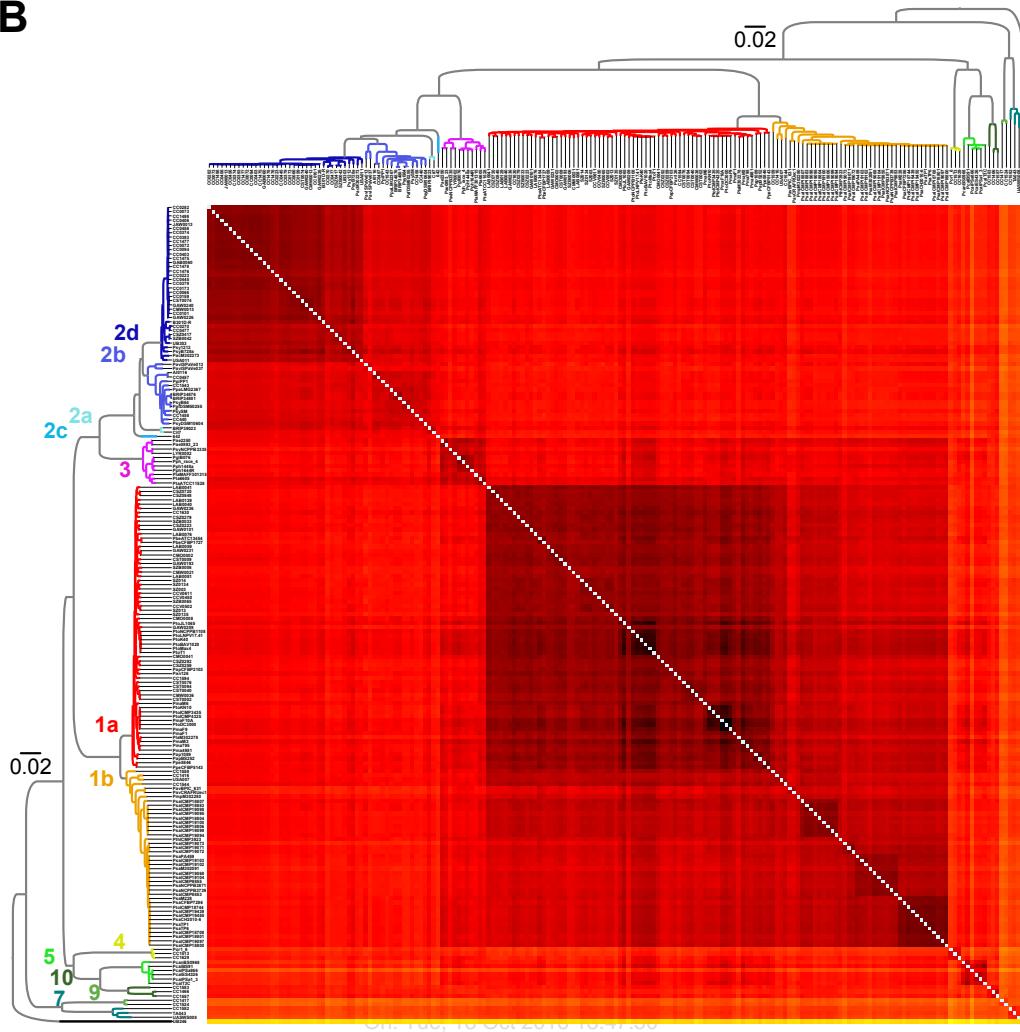


A**B**



Supplementary Figure 5

Click here to download Supplementary Material Files
S5_Figure.pdf

A**B**

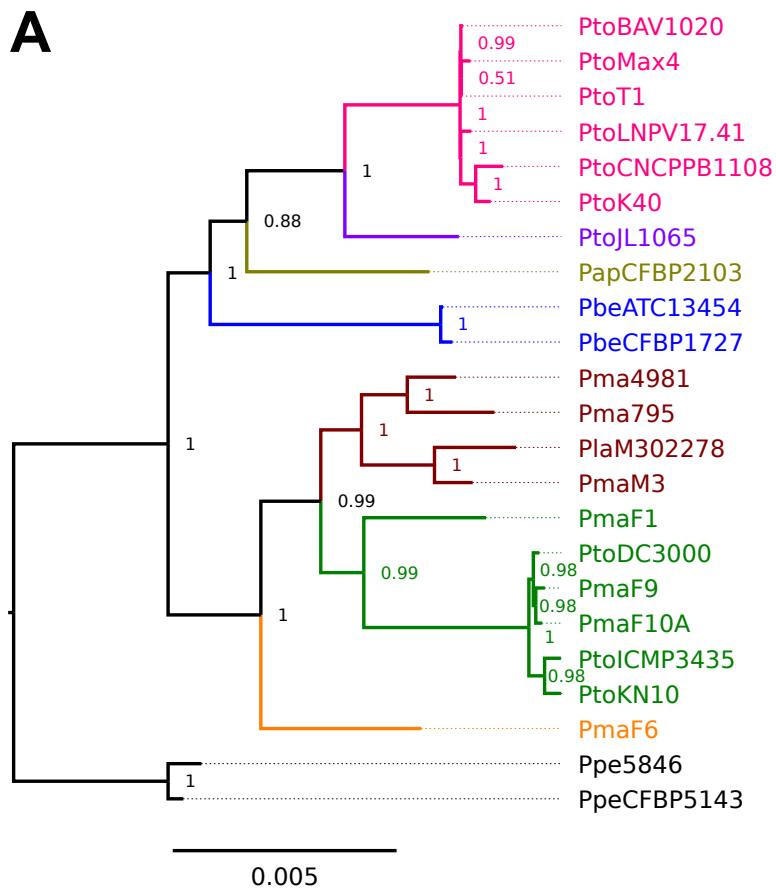
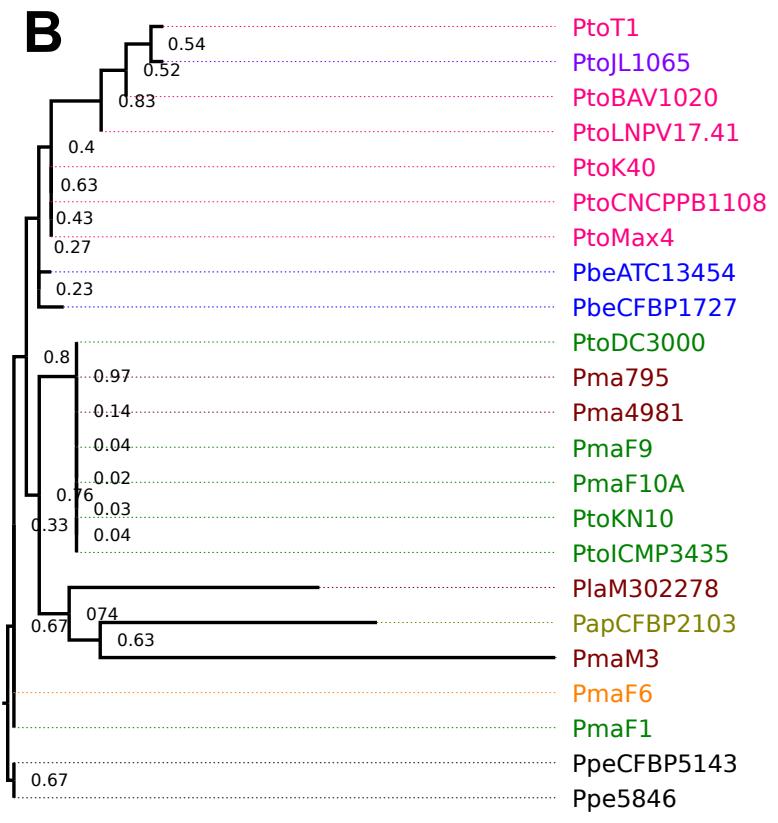
A**B**

Table S1. *P. syringae* isolates and genomes used in this study sorted by phylogroup according

Group	Isolate	Isolation	Isolation description
G2d	CC0066	France, <i>Cucumis melo</i>	Morris et al. (2000)
	CC0072	France, <i>Cucumis melo</i>	C.E. Morris ^c
	CC0073	France, <i>Cucumis melo</i>	Morris et al. (2000)
	CC0094	France, <i>Cucumis melo</i>	Morris et al. (2000)
	CC0101	France, <i>Cucumis melo</i>	Morris et al. (2000)
	CC0159	France, <i>Cucumis melo</i>	Morris et al. (2000)
	CC0173	France, <i>Cucumis melo</i>	C.E. Morris ^c
	CC0270	France, <i>Cucumis melo</i>	C.E. Morris ^c
	CC0282	France, <i>Cucumis melo</i>	C.E. Morris ^c
	CC0374	France, <i>Cucumis melo</i>	C.E. Morris ^c
	CC0379	France, <i>Cucumis melo</i>	C.E. Morris ^c
	CC0393	France, irrigation water	Morris et al. (2000)
	CC0403	France, irrigation water	Morris et al. (2000)
	CC0406	France, irrigation water	Morris et al. (2000)
	CC0445	France, <i>Cucumis melo</i>	C.E. Morris ^c
	CC0456	France, irrigation water	C.E. Morris ^c
	CC0477	France, <i>Cucumis melo</i>	C.E. Morris ^c
	CC1475	France, snowfall	Monteil et al. (2014b)
	CC1476	France, snowfall	Monteil et al. (2014b)
	CC1477	France, snowfall	Monteil et al. (2014b)
	CC1478	France, snowfall	Monteil et al. (2014b)
	CC1498	France, snowfall	Monteil et al. (2014b)
	CMW0013	France, river water	C.E. Morris ^b
	CST0074	France, rain	Monteil et al. (2014b)
	CSZ0417	France, stream water	Monteil et al. (2014a)
	GAB0060	France, epilithic biofilm	C.E. Morris ^b
	GAW0226	France, irrigation water	Berge et al. (2014)
	GAW0240	France, irrigation water	C.E. Morris ^b
	JAW0013	France, groundwater	C.E. Morris ^b
	SZB0042	France, epilithic biofilm	C.E. Morris ^b
	PsyB301D-R	UK, pear	Dudnik et al. (2014)

	UB303	France, stream water	Morris et al. (2010)
	USA011	USA, stream water	Morris et al. (2010)
	Psy1212	UK, pea	D.A. Baltrus
	PacM302273	Maple tree	Baltrus et al. (2011)
	PsyB728a	USA, common bean	Loper & Lindow (1987)
	CC0223	France, <i>Cucumis melo</i>	C.E. Morris ^c
G2b	AI0116	New Zealand, river water	Morris et al. (2010)
	CC0457	France, <i>Cucumis melo</i>	Morris et al. (2010)
	PsyDSM10604	UK, <i>Syringa vulgaris</i>	M. Rajan
	PafDSM50255	Wheat	D.A. Baltrus
	CC1458	USA, <i>Primula</i> sp.	Morris et al. (2008)
	CC1543	France, lake water	Morris et al. (2008)
	CC0440	France, <i>Cucumis melo</i>	Morris et al. (2008)
	Pav Ve037	Italy, hazelnut	O'Brien et al. (2012)
	PpiPP1	Japa, pea	Sarkar & Guttman (2004)
	BRIP34881	Australia, barley glumes	Gardiner et al. (2013)
	BRIP34876	Australia, barley glumes	Gardiner et al. (2013)
	PsyB64	USA, wheat	Dudnik et al. (2013)
	Pav Ve013	Italy, hazelnut	O'Brien et al. (2012)
	PsySM	USA, wheat	Dudnik et al. (2013)
G2a	BRIP39023	Australia, healthy barley	Gardiner et al. (2013)
	Cit7	USA, healthy orange tree	Baltrus et al. (2011)
	PpaLMG2367	Millet	Liu et al. (2012)
G2c	642	USA, weedy plant	Clarke et al. (2010)
G3	LYR0002	France, rain	Monteil et al. (2014b)
	Pta6605	Japan, Tobacco	D.J. Studholme
	Pph1644R	Mung bean	Baltrus et al. (2012)
	Pph_race_4	soybean	Baltrus et al. (2011)
	PglB076	USA, soybean	Qi et al. 2011
	PsvNCPPB_3335	France, Olive tree	P.P. Rodriguez-Palenzuela
	Pae2250	UK, Horse chesnut	Green et al. (2010)
	PaeNCPPB_3681	India, Horse chesnut	Green et al. (2010)
	PlaMAFF301315	Japan, cucumber	Baltrus et al. (2011)
	PtaATCC11528	Tobacco	D.J. Studholme

Pph1448a		Ethiopia, French bean	Joardar et al. (2005)
G1a	LAB0041	France, epilithic biofilm	Berge et al. (2014)
	CMO0008	France, rain	Monteil et al. (2014b)
	LAB0040	France, epilithic biofilm	Berge et al. (2014)
	CSZ0720	France, snowpack	Monteil et al. (2012)
	CSZ0279	France, snowpack	Monteil et al. (2012)
	CST0094	France, rain	Monteil et al. (2014b)
	CCV0611	France, snowpack	Monteil et al. (2012)
	GAW0231	France, irrigation water	Berge et al. (2014)
	CST0076	France, rain	Monteil et al. (2014b)
	CCV0502	France, leaf litter	Monteil et al. (2012)
	CMW0036	France, river water	Berge et al. (2014)
	SZB0033	France, epilithic biofilm	C.E. Morris ^b
	CST0009	France, rain	Monteil et al. (2014b)
	GAW0236	France, irrigation water	Berge et al. (2014)
	LAB0129	France, epilithic biofilm	Berge et al. (2014)
	CCV0450	France, snowpack	Monteil et al. (2012)
	CSZ0292	France, snowpack	Monteil et al. (2012)
	LAB0059	France, epilithic biofilm	C.E. Morris ^b
	Pap1089	USA, celery	Cai et al. (2011a)
	PapCFBP2103	USA, celery	Cai et al. (2011a)
	SZB0006	France, epilithic biofilm	Berge et al. (2014)
	GAW0193	France, irrigation water	C.E. Morris ^b
	CC1594	France, rain	Monteil et al. (2014b)
	SZ0134	France, stream water	Morris et al. (2010)
	LAB0076	France, epilithic biofilm	Berge et al. (2014)
	CST0040	France, rain	Monteil et al. (2014b)
	CST0002	France, rain	Monteil et al. (2014b)
	Ppe5846	France, peach	Cai et al. (2011a)
	SZ0135	France, stream water	Morris et al. (2010)
	GAW0209	France, epilithic biofilm	C.E. Morris ^b
	Pma795	New Zealand, cauliflower	Cai et al. (2011a)
	CSZ0259	France, stream water	Monteil et al. (2014a)
	CMW0021	France, irrigation water	C.E. Morris ^b
	LAB0081	France, epilithic biofilm	Berge et al. (2014)
	SZ003	France, stream water	Morris et al. (2010)
	PmaM6	UK, cauliflower	Debener et al. (1991)

	PmaM3	USA, cauliflower	Debener et al. (1991)
	PtoLNPV17.41	France, tomato	Zaccardelli et al. (2005)
	Pan126	UK, snapdragon	Yan et al. (2008)
	PtoJL1065	USA, tomato	Whalen et al. (1991)
	PmaF1	USA, spinach mustard	Zhao et al. (2000)
	PtoBAV1020	USA, tomato	Cai et al. (2011b)
	CMO0002	France, rain	Monteil et al. (2014b)
	CMO0041	France, rain	Monteil et al. (2014b)
	CSZ0223	France, stream water	Monteil et al. (2014a)
	CSZ0848	France, leaf litter	Monteil et al. (2012)
	GAW0101	France, irrigation water	C.E. Morris ^b
	PapBS252	USA, celery	Cai et al. (2011a)
	PbeATCC13454	USA, barberry	Cai et al. (2011a)
	PbeCFBP1727	New Zealand, barberry	Cai et al. (2011a)
	PmaICMP4981	Zimbabwe, cauliflower	Cai et al. (2011a)
	PpeCFBP5143	New Zealand, peach	Cai et al. (2011a)
	PtoMax4	Italy, tomato	Zaccardelli et al. (2005)
	SZ013	France, stream water	Morris et al. (2010)
	SZ014	France, stream water	Morris et al. (2010)
	SZB0065	France, epilithic biofilm	C.E. Morris ^b
	PmaF9	USA, spinach mustard	Zhao et al. (2000)
	PmaF10A	USA, turnip	Zhao et al. (2000)
	PtoKN10	Japan, tomato	Hwang et al. (2005)
	PtoICMP3435	New Zealand, woolly nightshade	Whalen et al. (1991)
	PtoICMP4325	Canada, tomato	Mitchell et al. (1983)
	PtoDC3000	UK, tomato	Cuppels and Ainsworth (1995)
	CC1630	USA, <i>Onobrychis viciifolia</i>	Berge et al. (2014)
	PtoNCPPB1108	UK, tomato	Yan et al. (2008)
	PtoK40	USA, tomato	Cai et al. (2011b)
	PtoT1	Canada, tomato	Ronald et al. (1992)
	PlaM302278	USA, cucumber	Hwang et al. (2005)
G1b	PavBPIC_631	Greece, hazelnut	Scortichini et al. (2013)
	PavCRAFRUec1	Italy, hazelnut	Scortichini et al. (2013)
	CC1416	USA, epilithic biofilm	Morris et al. (2007)
	CC1544	France, lake water	Morris et al. (2010)
	CC1559	France, snow	Morris et al. (2008)
	PsaICMP18806	New Zealand, kiwifruit	Butler et al. (2013)
	PsaM302091	Japan, kiwifruit	Baltrus et al. (2011)

	PsaCFBP7286	Italy, kiwifruit	Mazzaglia et al. (2012)
	PsaCH2010-6	China, kiwifruit	Mazzaglia et al. (2012)
	PsaICMP18708	New Zealand, kiwifruit	Butler et al. (2013)
	PtoICMP18744	New Zealand, kiwifruit	Butler et al. (2013)
	PsaICMP18800	New Zealand, kiwifruit	Butler et al. (2013)
	PsaICMP18801	New Zealand, kiwifruit	McCann et al. (2012)
	PsaICMP18804	New Zealand, kiwifruit	McCann et al. (2012)
	PsaICMP18807	New Zealand, kiwifruit	McCann et al. (2012)
	PsaICMP18883	New Zealand, kiwifruit	McCann et al. (2012)
	PsaICMP19068	Japan, kiwifruit	McCann et al. (2012)
	PsaICMP19071	Korea, kiwifruit	McCann et al. (2012)
	PsaICMP19072	Korea, kiwifruit	McCann et al. (2012)
	PsaICMP19073	Korea, kiwifruit	McCann et al. (2012)
	PsaICMP19094	New Zealand, kiwifruit	McCann et al. (2012)
	PsaICMP19095	New Zealand, kiwifruit	McCann et al. (2012)
	PsaICMP19097	New Zealand, kiwifruit	McCann et al. (2012)
	PsaICMP19098	New Zealand, kiwifruit	McCann et al. (2012)
	PsaICMP19099	New Zealand, kiwifruit	McCann et al. (2012)
	PsaICMP19100	New Zealand, kiwifruit	McCann et al. (2012)
	PsaICMP19102	Japan, kiwifruit	McCann et al. (2012)
	PsaICMP19103	Japan, kiwifruit	McCann et al. (2012)
	PsaICMP19104	Japan, kiwifruit	McCann et al. (2012)
	PsaICMP19439	Chile, kiwifruit	Butler et al. (2013)
	PsaICMP19455	Chile, kiwifruit	Butler et al. (2013)
	PsaICMP9853	Japan, kiwifruit	Butler et al. (2013)
	PsaICMP9855	Japan, kiwifruit	Mazzaglia et al. (2012)
	PsaPA459	Japan, kiwifruit	Mazzaglia et al. (2012)
	PsaM228	China, kiwifruit	Butler et al. (2013)
	PsaTP1	New Zealand, kiwifruit	Butler et al. (2013)
	PsaTP6-1	New Zealand, kiwifruit	Butler et al. (2013)
	PmpM302280	UK, <i>Prunus domestica</i>	Baltrus et al. (2011)
	PthICMP3923	Japan, <i>Camellia sinensis</i>	Mazzaglia et al. (2012)
	USA007	USA, creek water	Morris et al. (2010)
	PsaNCPPB3871	Italy, kiwifruit	Marceletti et al. (2011)
	PsaNCPPB3739	Japan, kiwifruit	McCann et al. (2012)
G4	CC1513	France, healthy <i>Hutchinsia alpina</i>	Morris et al. (2008)
	CC1629	USA, <i>Avena sativa</i>	Morris et al. (2010)
	Por1_6	Japan, <i>Oryzae sativa</i>	Takeuchi et al. (1992)
G5	PcalES4326	USA, radish	Hwang et al. (2005)

	PcalBS91	USA, broccoli	Sarris et al. (2013)
	PcalT3C	USA, turnip	Sarris et al. (2013)
	PcanBS0968	Hungary, <i>Cannabis sativa</i>	Sarris et al. (2013)
	PcalPSa866	Greece, arugula	Sarris et al. (2013)
	PcalPSa1_3	Greece, arugula	Sarris et al. (2013)
G10	CC1466	USA, healthy <i>Primula</i> sp.	Morris et al. (2008)
	CC1557	France, snow	Morris et al. (2008)
	CC1583	France, snow	Berge et al. (2014)
G9	CC1417	USA, epilithic biofilm	Morris et al. (2007)
	CC1524	France, stream water	Morris et al. (2010)
G7	CC1582	France, snow	Berge et al. (2014)
	TA043	France, stream water	Morris et al. (2010)
	UASWS0038	<i>Phytophthora</i> sp.-infected rhododendron	Lefort et al. (2013)
G13	UB246	France, river water	Morris et al. (2010)

^b isolation described in Berge et al. (2014)

^c isolation described in Morris et al. (2000)

^d genomes can be downloaded from <http://pacu.facom.ufms.br/Pcal/genomes>

ing to the nomenclature proposed by Berge et al. (2014)

Accession number NCBI	Number of Contigs	Total genome length	Min contig size	Max contig size	Mean contig size	Standard deviation contig size
This study	165	6033775	214	416361	36569	55779
This study	199	6104177	202	227761	30675	39790
This study	178	5980895	204	401007	33601	53016
This study	289	6087347	210	132807	21064	25884
This study	172	6027767	208	416166	35046	56050
This study	174	6013936	211	247742	34563	49812
This study	200	6111737	215	199299	30559	38087
This study	1175	5894131	203	49603	5017	5267
This study	238	6020869	230	123221	25298	27205
This study	475	6026378	204	113797	12688	15935
This study	353	6087176	201	117801	17245	18685
This study	498	6024357	201	95931	12098	15562
This study	231	6089668	202	274226	26363	41708
This study	215	6094469	204	370435	28347	45939
This study	203	6000915	203	265670	29562	42904
This study	222	6096181	203	373803	27461	45916
This study	175	5966800	201	186675	34096	37972
This study	239	6093610	202	233909	25497	39990
This study	719	6000868	204	67784	8347	9094
This study	267	6085100	201	213830	22791	35228
This study	575	6033288	206	85462	10493	13259
This study	221	6090515	213	353351	27559	44884
This study	566	6053034	213	86512	10695	11271
This study	229	6117272	201	145443	26713	29549
This study	685	5888353	201	76544	8597	9305
This study	530	5991265	212	94944	11305	12326
This study	347	6000752	204	87136	17294	18629
This study	193	6122101	214	235465	31721	41180
This study	235	6085838	213	215734	25898	41592
This study	658	5893272	206	57422	8957	9429
GCA_000585725.1	81	6036561	358	498139	74526	95360

GCA_000452565.2	478	5988839	199	230495	12529	19544
GCA_000452525.2	198	6429288	212	415631	32472	61151
GCA_000452465.2	338	6163906	200	306351	18237	36655
GCA_000145925.1	1179	6183769	100	50619	5245	7051
CP000075.1	1	6093698	6093698	6093698	6093698	NA
This study	463	5960432	503	81466	12874	12447
This study	233	5890279	212	171127	25281	26967
This study	152	5897069	203	296075	38797	54289
GCA_000597765.1	84	6072447	503	422665	72292	83755
GCA_000498595.1	669	5777269	199	120568	8636	14565
GCA_000452805.2	364	5801209	200	171448	15938	25126
GCA_000452725.2	441	5777476	199	209143	13101	20042
GCA_000452605.2	354	5725117	202	256307	16173	26170
GCA_000452445.2	152	5825140	145	234413	38324	45965
GCA_000452445.2	256	5949520	201	342107	23241	46859
GCA_000334055.1	157	6008189	212	357454	38269	63049
GCA_000334035.1	148	6005932	214	357425	40581	69850
GCA_000331385.1	42	5873298	315	766547	139841	190673
GCA_000302795.1	317	5975428	202	172456	18850	22556
CM001986.1	27	6140958	1504	1737860	227443	374799
CM001987.1						
GCA_000333995.1	71	5930576	203	448269	83530	110445
GCA_000145825.1	2655	6075147	100	34469	2289	3690
GCA_000282735.1	148	5993579	231	337491	40498	59108
GCA_000177515.1	296	5809390	208	117004	19627	24778
This study	380	5995356	207	136328	15778	23761
GCA_000275945.1	283	6159286	201	222375	21765	34611
GCA_000225805.1	1154	5870279	200	44097	5087	6335
GCA_000187065.2	108	5905212	833	445210	54678	97273
GCA_000187045.2	104	6236653	495	566593	59968	103440
GCA_000164015.2	397	5760120	467	114518	14510	16104
GCA_000163275.1	776	5926327	51	91451	7638	13651
GCA_000163255.1	841	5836346	51	93372	6940	9882
GCA_000146005.1	791	6759945	100	96173	8547	13170
GCA_000145945.1	1613	6158837	100	68422	3819	7790

CP000058.1	3	6112448	51711	5928787	2037483	3370208
CP000059.1						
CP000060.1						
This study	280	5731298	203	155604	20469	22494
This study	343	5927220	201	96843	17281	18607
This study	297	5695312	203	126589	19177	21247
This study	295	5762589	209	105609	19535	20773
This study	288	5757665	242	120666	19992	20942
This study	342	5875705	212	102739	17181	20055
This study	373	5722139	225	150679	15341	17202
This study	526	5805640	209	72435	11038	11401
This study	411	5814073	222	76532	14147	14225
This study	417	5922954	208	107477	14204	16317
This study	1125	5885656	202	31647	5232	5216
This study	513	5730128	206	83766	11170	11339
This study	541	5808699	256	75478	10737	10758
This study	1049	5870644	209	37625	5597	4909
This study	563	5610451	203	63903	9966	9328
This study	567	5826723	209	65339	10277	9794
This study	505	5722242	234	78708	11332	11549
This study	1054	5752984	202	39450	5459	5182
This study	511	6148426	223	88599	12033	12924
This study	652	5984216	203	56992	9179	9550
This study	1109	5627941	207	36325	5075	4483
This study	938	5887086	204	37168	6277	5367
This study	885	5680274	201	40168	6419	5995
This study	670	5977232	225	53061	8922	8370
This study	958	5627648	214	32821	5875	5183
This study	700	5913678	218	58437	8449	9473
This study	1242	5851041	201	24219	4711	4117
This study	570	6001251	205	158321	10529	14718
This study	742	5655963	202	47267	7623	7094
This study	1203	5673098	201	35533	4716	4048
This study	1042	6113474	201	51483	5868	6063
This study	819	5757301	220	53263	7030	6852
This study	1371	5739957	205	35287	4187	3610
This study	1438	5747568	218	27393	3997	3523
This study	134	5882459	223	236829	43899	55618
This study	453	5974223	209	120594	13189	15780

This study	599	6255438	202	96662	10444	12809
This study	367	6070520	205	162633	16541	24339
This study	228	6158168	201	185606	27010	33112
This study	414	6292112	219	128407	15199	19948
This study	455	6263009	242	85876	13765	15731
This study	309	6230260	210	249597	20163	33221
This study	220	5954325	203	306158	27066	37393
This study	128	5930905	210	318254	46336	61506
This study	188	5933561	201	286300	31562	49308
This study	148	5789446	220	303324	39118	54530
This study	201	5808985	209	237765	28901	34330
This study	105	6079836	202	297377	57904	66678
This study	362	6008746	203	174329	16599	26802
This study	473	6254908	202	123814	13224	17159
This study	398	6380930	201	225172	16033	27490
This study	448	6135905	208	162497	13697	25283
This study	258	6260812	202	196989	24267	41063
This study	162	6027666	226	258479	37208	49201
This study	179	5940307	210	208223	33187	45465
This study	121	5793050	230	334440	47877	64867
This study	499	6370103	504	105940	12766	15711
This study	507	6291646	504	94678	12410	14267
This study	362	6200395	502	105928	17129	20103
This study	496	6193349	513	105928	12487	14564
This study	609	6137266	512	84262	10078	11575
NC_004578.1	3	6538260	67473	6397126	2179420	3652642
NC_004633.1						
NC_004632.1						
GCA_000452625.2	283	6056064	201	182258	21400	28692
GCA_000177495.1	304	6082048	210	153621	20007	24979
GCA_000177455.1	582	6153658	203	104644	10574	14619
GCA_000172895.1	122	6145942	219	335366	50377	72837
GCA_000145885.1	798	5895184	102	57617	7388	9495
GCA_000444135.1	297	6625623	201	234505	22309	37967
GCA_000441975.1	547	5736089	1019	67487	10487	11032
GCA_000452845.2	336	5921932	199	136516	17625	21134
GCA_000452905.2	376	5883551	199	114546	15648	18325
GCA_000452685.2	365	5858350	200	108667	16051	17392
GCA_000344415.1	284	6257465	200	134790	22034	26309
GCA_000145865.1	941	5849032	100	48124	6216	8250

GCA_000245415.1 462	6124586	203	97519	13257	17118
GCA_000245475.1 563	6166361	201	84456	10953	13776
GCA_000344355.1 445	6229906	200	139630	14000	19732
GCA_000342185.1 442	6235607	200	143214	14108	21223
GCA_000344375.1 463	6228594	200	88251	13453	17240
GCA_000416945.1 610	6121169	201	100404	10035	12445
GCA_000416905.1 263	6253677	200	224101	23779	30079
GCA_000344535.1 148	6125821	249	274756	41391	47814
GCA_000416785.1 524	6282306	204	100484	11990	14295
GCA_000416705.1 637	6020802	201	100881	9452	13890
GCA_000416485.1 502	6003331	201	83918	11959	12930
GCA_000416885.1 290	6013211	201	169668	20736	24885
GCA_000416505.1 659	5927068	201	86792	8995	11132
GCA_000416745.1 620	6253438	203	93014	10087	11645
GCA_000416765.1 921	6165730	205	52120	6695	7417
GCA_000416725.1 648	6125309	202	80234	9453	11889
GCA_000416545.1 428	6285813	210	130718	14687	18397
GCA_000416805.1 461	6202406	203	115477	13455	15791
GCA_000416825.1 561	6181694	202	113072	11020	13180
GCA_000416585.1 523	5987154	203	113869	11448	14129
GCA_000416845.1 544	5989995	203	84624	11012	13125
GCA_000416525.1 517	5941175	201	133960	11492	14365
GCA_000416865.1 477	6216001	200	142020	13032	19503
GCA_000344555.1 415	6227178	211	127854	15006	21630
GCA_000344515.1 515	6002432	200	121081	11656	16576
GCA_000416665.1 375	5931093	201	133407	15817	20777
GCA_000245455.1 730	6415594	201	174501	8789	12683
GCA_000344475.2 419	6347454	200	237204	15150	22884
GCA_000344435.1 472	6220309	200	138717	13179	18237
GCA_000344455.1 483	6224105	201	143214	12887	18386
GCA_000145745.1 969	6039297	101	69145	6233	8986
GCA_000416465.1 378	6255742	205	198679	16550	24256
GCA_000452545.2 369	5970940	201	119183	16182	19036
GCA_000233795.2 462	5938086	200	122209	12853	16861
GCA_000233835.2 815	5927897	200	67329	7274	9029
GCA_000452765.2 164	5725032	208	223068	34909	46187
GCA_000452645.2 261	5932928	200	305884	22732	45201
GCA_000156995.1 2855	6704257	70	63953	2349	5173
GCA_000145845.1 878	6221751	100	104693	7087	10959

IMG ^d	401	6015117	201	155236	15001	19684
IMG ^d	464	5821618	100	140678	12547	18505
IMG ^d	808	5742917	203	64737	7108	9151
IMG ^d	367	5718447	211	649129	15582	67988
IMG ^d	489	6022892	154	684944	12317	62739
GCA_000452785.2	294	5591749	199	222713	19020	32510
CP007014.1	2	5811653	53629	5758024	2905827	4033617
CP007015.1						
GCA_000452665.2	308	5507197	199	224888	17881	25135
GCA_000452825.2	210	5648464	199	253590	26898	34126
GCA_000452745.2	264	5828366	199	204924	22078	29015
GCA_000452505.1	553	5985404	12	126122	10824	16947
GCA_000452485.1	545	5946509	29	145835	10912	18348
GCA_000307715.1	201	5910810	230	168057	29408	30476
GCA_000452865.1	1327	6176522	25	100273	4655	7211

N50	N90	N95
100916	21860	14123
66727	19855	13212
81331	20638	13326
46079	13746	9334
113307	22534	14263
86441	22716	13778
64468	20029	11983
8694	2401	1649
51344	16557	10865
25532	7882	4305
31835	10164	6807
25118	7824	4289
76555	20425	12098
82172	19817	12969
75385	25817	13496
81786	20552	12181
64716	20626	14277
73854	20958	12028
15534	4605	3111
67741	18015	9111
22139	5930	3522
70474	18021	11748
17243	6040	3867
54315	17336	11333
15040	4503	2863
19746	6598	4192
34439	9813	6571
75140	17921	13367
81948	20428	11879
15484	4654	3090
169605	39223	28384

31570	8073	4597
91862	21120	13421
61944	16830	10552
12409	3090	1621
6093698	6093698	6093698
22575	6740	4501

43727	14673	10249
103116	23686	12946
142141	48088	30871
27242	6774	2936
46904	12131	5556
32477	9050	4908
44744	10302	6599
86726	22154	13923
81886	20935	11147
132100	32463	22498
153727	39906	24282
329354	87413	61608
37856	10552	7293
671136	144930	110933

200657	59541	34604
6862	1381	634
101322	29231	17221

46825	12315	7624
46663	9970	4502
64341	18398	12970
11316	2692	1577
225100	43552	15451
202511	39061	18452
26591	7345	4732
29194	5946	2433
17923	3945	2507
22550	5518	3419
16098	3818	1572

5928787 5928787 5928787

38498	11930	7849
35241	10377	6293
38368	11729	6601
39480	10879	6817
37004	11670	8299
36128	9838	6190
25871	8508	5786
19188	6066	3871
25284	7667	5415
26848	8195	5239
8965	2649	1683
19850	6040	4065
18868	6616	4344
8500	3037	2167
15987	5479	3672
17556	5756	3781
20343	5986	3860
8423	2796	1869
22001	7009	3901
16575	4800	2903
7402	2615	1883
9511	3455	2314
10122	3347	2214
15433	4934	3139
8770	3223	2229
15153	4323	2689
7385	2529	1747
25957	6019	2831
12399	4003	2704
7047	2508	1663
10021	3030	1985
11960	3749	2480
6337	2277	1518
5909	2076	1439
109342	39638	21753
26312	7210	4634

19827	5781	3319
40596	11694	5775
54043	14939	9077
34668	9328	5150
28691	7631	4281
64056	19921	7244
67827	20197	12895
96405	34022	21839
96170	27691	20641
100047	33387	23004
59338	18007	10812
144160	48718	30891
53941	10676	5000
30251	8448	4263
44538	9907	4999
51194	8563	3472
82875	24668	8357
99098	28799	16035
83038	25487	15299
113640	44275	22825
27133	6849	3953
24806	6712	3779
34027	9853	5730
24365	6608	4100
18332	5375	3211
6397126	6397126	6397126

52218	10722	7532
47802	13750	8092
26013	6981	3486
150139	36607	24013
15738	4805	2910
79231	24793	11991
16957	5311	3155
32560	9246	6871
29886	9144	6110
30845	9220	5419
41975	13843	9001
14086	3678	1960

31420	9112	4730
24332	6788	3869
32941	9601	5066
35904	9600	4868
31235	9399	4781
21491	5383	3172
52055	13950	9196
90248	25142	16141
25493	6404	4193
25695	5299	3248
23149	6683	4328
43039	12805	7752
16986	4732	2936
18193	5561	3882
12305	3488	2191
19576	5070	3031
30557	8466	5346
27427	7400	5112
22859	6092	4038
23155	5891	3559
21642	5795	3408
23554	6082	3593
36822	9273	4367
39400	9919	5377
27018	7337	4259
38862	10090	5148
17643	4793	3092
38960	9969	5279
31162	8800	4964
31517	8924	4629
15161	3986	2165
42347	12229	6902
35399	8517	5546
27730	7815	4386
14838	3771	2381
85568	23119	17542
75208	18885	7568
10037	1680	607
17222	3864	2153

34219	8305	5119
29743	8237	5187
16366	3352	1852
248118	38125	2244
249178	46398	5460
60479	14624	7405
5758024	5758024	5758024
35492	8069	5870
50166	12957	9366
44224	10532	7307
31043	8731	5895
34387	9440	5893
48957	15216	10013
11025	3201	2158

Table S2. Genomes used in the rMLST tree representing the relationships between all genomes in the HierBAPS cluster 4 (FigureS1) with the exception of the isolate UB246.

HierBAPS cluster	Species name	Isolate	Accession number NCBI	Number of Contigs
2	<i>P. aeruginosa</i>	138244	GCA_000215775.4	24
2	<i>P. aeruginosa</i>	152504	GCA_000215795.3	389
2	<i>P. aeruginosa</i>	18A	GCA_000341565.1	179
2	<i>P. aeruginosa</i>	19BR	GCA_000223945.2	1
2	<i>P. aeruginosa</i>	213BR	GCA_000223965.2	1
2	<i>P. aeruginosa</i>	2192	GCA_000152545.1	82
2	<i>P. aeruginosa</i>	39016	GCA_000148745.1	349
2	<i>P. aeruginosa</i>	9BR	GCA_000223925.2	2
2	<i>P. aeruginosa</i>	AES-1R	GCA_000220025.2	1783
2	<i>P. aeruginosa</i>	AH16	GCA_000287875.1	130
2	<i>P. aeruginosa</i>	ATCC14886	GCA_000297275.1	211
2	<i>P. aeruginosa</i>	ATCC25324	GCA_000297295.1	1023
2	<i>P. aeruginosa</i>	ATCC700888	GCA_000297315.1	600
2	<i>P. aeruginosa</i>	B136-33	GCA_000359505.1	1
2	<i>P. aeruginosa</i>	B3-1811	GCA_000455545.1	341
2	<i>P. aeruginosa</i>	B3-208	GCA_000455405.1	418
2	<i>P. aeruginosa</i>	B3-20M	GCA_000455505.1	338
2	<i>P. aeruginosa</i>	B3-CFI	GCA_000455425.1	314
2	<i>P. aeruginosa</i>	C3719	GCA_000152525.1	1
2	<i>P. aeruginosa</i>	c7447m	GCA_000468935.1	1
2	<i>P. aeruginosa</i>	CI27	GCA_000297335.1	160
2	<i>P. aeruginosa</i>	CIG1	GCA_000295475.1	544
2	<i>P. aeruginosa</i>	DK2	GCA_000271365.1	1
2	<i>P. aeruginosa</i>	DQ8	GCA_000283055.1	376
2	<i>P. aeruginosa</i>	E2	GCA_000297355.1	196
2	<i>P. aeruginosa</i>	LCT-PA102	GCA_000258285.1	180
2	<i>P. aeruginosa</i>	LCT-PA220	GCA_000439855.1	190
2	<i>P. aeruginosa</i>	LCT-PA41	GCA_000439875.1	189
2	<i>P. aeruginosa</i>	LESB58	GCA_000026645.1	1
2	<i>P. aeruginosa</i>	M18	GCA_000226155.1	1
2	<i>P. aeruginosa</i>	MPAO1/P1	GCA_000247435.2	140
2	<i>P. aeruginosa</i>	MPAO1/P2	GCA_000247455.2	301
2	<i>P. aeruginosa</i>	MRW44.1	GCA_000282915.1	54
2	<i>P. aeruginosa</i>	MSH-10	GCA_000407905.1	19

2	<i>P. aeruginosa</i>	N002	GCA_000287815.2	235
2	<i>P. aeruginosa</i>	NCGM2.S1	GCA_000284555.1	1
2	<i>P. aeruginosa</i>	NCMG1179	GCA_000291745.1	92
2	<i>P. aeruginosa</i>	PA14	GCA_000404265.1	26
2	<i>P. aeruginosa</i>	PA21ST175	GCA_000342145.1	90
2	<i>P. aeruginosa</i>	PA45	GCA_000359565.1	124
2	<i>P. aeruginosa</i>	PA7	GCA_000017205.1	1
2	<i>P. aeruginosa</i>	PAb1	GCA_000172395.1	658
2	<i>P. aeruginosa</i>	PABL056	GCA_000290555.1	401
2	<i>P. aeruginosa</i>	PACS2	GCA_000168335.1	1
2	<i>P. aeruginosa</i>	PADK2CF510	GCA_000259025.1	71
2	<i>P. aeruginosa</i>	PAK	GCA_000408865.1	26
2	<i>P. aeruginosa</i>	PAO1	GCA_000006765.1	1
2	<i>P. aeruginosa</i>	PAO1-CipR	GCA_000414255.1	1677
2	<i>P. aeruginosa</i>	PAO579	GCA_000296325.1	16
2	<i>P. aeruginosa</i>	PAO581	GCA_000468555.1	1
2	<i>P. aeruginosa</i>	PGPR2	GCA_000412735.1	198
2	<i>P. aeruginosa</i>	RP73	GCA_000414035.1	1
2	<i>P. aeruginosa</i>	SJTD-1	GCA_000271985.1	104
2	<i>P. aeruginosa</i>	C1334	GCA_000412295.1	669
2	<i>P. aeruginosa</i>	C1426	GCA_000412555.1	421
2	<i>P. aeruginosa</i>	C1433	GCA_000412535.1	2020
2	<i>P. aeruginosa</i>	C763	GCA_000412355.1	844
2	<i>P. aeruginosa</i>	E2UoS	GCA_000412395.1	341
2	<i>P. aeruginosa</i>	J1385	GCA_000412515.1	2516
2	<i>P. aeruginosa</i>	J1532	GCA_000412475.1	2544
2	<i>P. aeruginosa</i>	MSH10bis	GCA_000412435.1	18
2	<i>P. aeruginosa</i>	MSH3	GCA_000412375.1	1240
2	<i>P. aeruginosa</i>	PA17	GCA_000412455.1	1245
2	<i>P. aeruginosa</i>	PA17_SCV	GCA_000412495.1	1853
2	<i>P. aeruginosa</i>	PA62	GCA_000412415.1	648
2	<i>P. aeruginosa</i>	UCBPPA14	GCA_000014625.1	1
2	<i>P. aeruginosa</i>	VRFPA01	GCA_000335395.1	617
2	<i>P. aeruginosa</i>	VRFPA02	GCA_000399805.1	186
2	<i>P. aeruginosa</i>	VRFPA03	GCA_000467675.1	150
2	<i>P. aeruginosa</i>	WC55	GCA_000455705.1	342
2	<i>P. aeruginosa</i>	XMG	GCA_000265035.1	225
5	<i>P. agari</i>	Pag-NCPPB2289	GCA_000280785.1	119
3	<i>P. alcaligenes</i>	Palg-NBRC14159	GCA_000467105.1	122
3	<i>P. alcaligenes</i>	Palg-OT69	GCA_000455385.1	223
3	<i>P. alcaliphila</i>	Pal34	GCA_000319815.1	18

3	<i>P. azotifigens</i>	Paz-DSM17556	GCA_000425625.1	97
5	<i>P. brassicacearum</i>	Pbr-NFM421	GCA_000194805.1	1
7	<i>P. caeni</i>	Pca-DSM24390	GCA_000421765.1	24
5	<i>P. chlororaphis</i>	Pch-O6	GCA_000264555.1	1
5	<i>P. chlororaphis</i>	Pch-30-84	GCA_000281915.1	1
5	<i>P. chlororaphis</i>	Pch-GP72	GCA_000237045.2	347
10	<i>P. cremoricolorata</i>	Pcr-DSM17059	GCA_000425745.1	27
3	<i>P. denitrificans</i>	Pde-ATCC13867	GCA_000349845.1	1
10	<i>P. entomophila</i>	Pen-L48	GCA_000026105.1	1
5	<i>P. extremaustralis</i>	Pex-14-3b	GCA_000242115.2	135
5	<i>P. fluorescens</i>	A506	GCA_000262325.2	2
5	<i>P. fluorescens</i>	BBc6R8	GCA_000297195.1	163
5	<i>P. fluorescens</i>	BRIP34879	GCA_000334015.1	110
5	<i>P. fluorescens</i>	BS2	GCA_000308175.1	122
5	<i>P. fluorescens</i>	EGD-AQ6	GCA_000465595.1	58
5	<i>P. fluorescens</i>	F113	GCA_000237065.1	1
5	<i>P. fluorescens</i>	HK44	GCA_000217955.2	32
5	<i>P. fluorescens</i>	LMG5329	GCA_000411675.1	253
5	<i>P. fluorescens</i>	NCIMB11764	GCA_000293885.2	1
5	<i>P. fluorescens</i>	NZ007	GCA_000280805.1	141
5	<i>P. fluorescens</i>	NZ011	GCA_000276585.1	973
5	<i>P. fluorescens</i>	NZ052	GCA_000275925.1	440
5	<i>P. fluorescens</i>	NZ17	GCA_000275905.1	1032
5	<i>P. fluorescens</i>	Pf0-1	GCA_000012445.1	1
5	<i>P. fluorescens</i>	Pf29Arp	GCA_000346775.1	206
5	<i>P. fluorescens</i>	Q2-87	GCA_000281895.1	1
5	<i>P. fluorescens</i>	Q8r1-96	GCA_000263695.2	1
5	<i>P. fluorescens</i>	R124	GCA_000292795.1	78
5	<i>P. fluorescens</i>	SBW25	GCA_000009225.1	2
5	<i>P. fluorescens</i>	SS101	GCA_000263675.2	1
5	<i>P. fluorescens</i>	Wayne1	GCA_000285355.1	337
5	<i>P. fluorescens</i>	WH6	GCA_000166515.1	1
5	<i>P. fluorescens</i>	Wood1R	GCA_000285615.1	1437
5	<i>P. fragi</i>	A22	GCA_000250595.1	48
5	<i>P. fragi</i>	B25	GCA_000250615.1	11
3	<i>P. fulva</i>	12-X	GCA_000213805.1	1
5	<i>P. fuscovaginae</i>	CB98818	GCA_000280575.1	261
5	<i>P. fuscovaginae</i>	DAR77795	GCA_000467005.1	482
5	<i>P. fuscovaginae</i>	DAR77800	GCA_000467025.1	791
5	<i>P. fuscovaginae</i>	ICMP5940	GCA_000467065.1	459
5	<i>P. fuscovaginae</i>	SE-1	GCA_000364705.1	407

5	<i>P. fuscovaginae</i>	UPB0736	GCA_000251185.1	102
5	<i>P. gingeri</i>	NCPPB3146	GCA_000280765.1	75
6	<i>P. luteola</i>	Plu-XLDN4-9	GCA_000282775.1	231
5	<i>P. mandelii</i>	36MFCvi1.1	GCA_000381285.1	35
5	<i>P. mandelii</i>	JR-1	GCA_000257545.2	34
3	<i>P. mendocrina</i>	DLHK	GCA_000287395.1	33
3	<i>P. mendocrina</i>	EGD-AQ5	GCA_000465575.1	19
3	<i>P. mendocrina</i>	NK-01	GCA_000204295.1	1
3	<i>P. mendocrina</i>	S5.2	GCA_000295795.1	46
3	<i>P. mendocrina</i>	ymp	GCA_000016565.1	1
10	<i>P. monteilii</i>	Pmo-QM	GCA_000262005.1	278
3	<i>P. nitroreducens</i>	TX1	GCA_000313755.1	138
10	<i>P. parafulva</i>	DSM17004	GCA_000425765.1	40
9	<i>P. pelagia</i>	CL-AP6	GCA_000410875.1	81
10	<i>P. plecoglossicida</i>	NB2011	GCA_000412715.1	87
5	<i>P. poae</i>	RE1114	GCA_000336465.1	1
5	<i>P. protegens</i>	CHA0	GCA_000397205.1	1
5	<i>P. protegens</i>	Pr-5	GCA_000012265.1	1
3	<i>P. pseudoalcaligenes</i>	CECT	GCA_000297075.1	24
3	<i>P. pseudoalcaligenes</i>	KF707	GCA_000262065.3	233
5	<i>P. psychrophila</i>	HA-4	GCA_000282975.1	145
1	<i>P. psychrotolerans</i>	L19	GCA_000236825.2	53
10	<i>P. putida</i>	B001	GCA_000285395.1	262
10	<i>P. putida</i>	B6-2	GCA_000226035.2	27
10	<i>P. putida</i>	BIRD-1	GCA_000183645.1	1
10	<i>P. putida</i>	CSV86	GCA_000319305.1	209
10	<i>P. putida</i>	DOT-T1E	GCA_000281215.1	2
10	<i>P. putida</i>	F1	GCA_000016865.1	1
10	<i>P. putida</i>	GB-1	GCA_000019125.1	1
10	<i>P. putida</i>	H8234	GCA_000410575.1	1
10	<i>P. putida</i>	HB3267	GCA_000325725.1	2
10	<i>P. putida</i>	KT2440	GCA_000007565.1	1
10	<i>P. putida</i>	LF54	GCA_000390005.1	180
10	<i>P. putida</i>	LS46	GCA_000294445.2	32
10	<i>P. putida</i>	MR3	GCA_000390445.1	1268
10	<i>P. putida</i>	MTCC5279	GCA_000411615.1	171
10	<i>P. putida</i>	NBRC14164	GCA_000412675.1	1
10	<i>P. putida</i>	ND6	GCA_000264665.1	3
10	<i>P. putida</i>	S11	GCA_000292775.1	196
10	<i>P. putida</i>	S12	GCA_000287915.1	258
10	<i>P. putida</i>	S16	GCA_000219705.1	1

10	<i>P. putida</i>	SJTE-1	GCA_000271965.1	207
10	<i>P. putida</i>	Idaho	GCA_000226475.2	839
10	<i>P. putida</i>	TRO1	GCA_000367825.1	222
10	<i>P. putida</i>	W619	GCA_000019445.1	1
3	<i>P. resinovorans</i>	DSM21078	GCA_000423545.1	59
3	<i>P. resinovorans</i>	NBRC106553	GCA_000412695.1	2
2	<i>Pseudomonas sp.</i>	2_1_26	GCA_000233495.1	51
1	<i>Pseudomonas sp.</i>	313	GCA_000316965.1	110
5	<i>Pseudomonas sp.</i>	35MFCvi1.1	GCA_000378525.1	26
5	<i>Pseudomonas sp.</i>	45MFCol3.1	GCA_000382025.1	23
5	<i>Pseudomonas sp.</i>	Ag1	GCA_000278565.1	113
5	<i>Pseudomonas sp.</i>	CBZ-4	GCA_000346755.1	186
5	<i>Pseudomonas sp.</i>	CF149	GCA_000416155.1	50
5	<i>Pseudomonas sp.</i>	CF161	GCA_000416215.1	132
5	<i>Pseudomonas sp.</i>	CFII64	GCA_000416235.1	54
5	<i>Pseudomonas sp.</i>	CFII68	GCA_000416195.1	128
5	<i>Pseudomonas sp.</i>	CFT9	GCA_000416255.1	67
8	<i>Pseudomonas sp.</i>	Chol1	GCA_000306015.1	199
3	<i>Pseudomonas sp.</i>	EGD-AK9	GCA_000465935.1	1016
5	<i>Pseudomonas sp.</i>	G5	GCA_000408945.1	223
5	<i>Pseudomonas sp.</i>	GM102	GCA_000282555.1	159
5	<i>Pseudomonas sp.</i>	GM16	GCA_000282155.1	128
5	<i>Pseudomonas sp.</i>	GM17	GCA_000282175.1	280
5	<i>Pseudomonas sp.</i>	GM18	GCA_000282195.1	140
5	<i>Pseudomonas sp.</i>	GM21	GCA_000282215.1	210
5	<i>Pseudomonas sp.</i>	GM24	GCA_000282235.1	399
5	<i>Pseudomonas sp.</i>	GM25	GCA_000282255.1	91
5	<i>Pseudomonas sp.</i>	GM30	GCA_000282275.1	180
5	<i>Pseudomonas sp.</i>	GM33	GCA_000282295.1	205
5	<i>Pseudomonas sp.</i>	GM41	GCA_000282315.1	164
5	<i>Pseudomonas sp.</i>	GM48	GCA_000282335.1	200
5	<i>Pseudomonas sp.</i>	GM49	GCA_000282355.1	345
5	<i>Pseudomonas sp.</i>	GM50	GCA_000282375.1	155
5	<i>Pseudomonas sp.</i>	GM55	GCA_000282395.1	163
5	<i>Pseudomonas sp.</i>	GM60	GCA_000282415.1	181
5	<i>Pseudomonas sp.</i>	GM67	GCA_000282435.1	183
5	<i>Pseudomonas sp.</i>	GM74	GCA_000282455.1	180
5	<i>Pseudomonas sp.</i>	GM78	GCA_000282475.1	235
5	<i>Pseudomonas sp.</i>	GM79	GCA_000282495.1	126
5	<i>Pseudomonas sp.</i>	GM80	GCA_000282515.1	282
10	<i>Pseudomonas sp.</i>	GM84	GCA_000282535.1	384

10	<i>Pseudomonas</i> sp.	HYS	GCA_000259195.1	64
10	<i>Pseudomonas</i> sp.	ICMP17674	GCA_000467045.1	359
5	<i>Pseudomonas</i> sp.	Lz4W	GCA_000346225.1	42
3	<i>Pseudomonas</i> sp.	M1	GCA_000317185.1	9
5	<i>Pseudomonas</i> sp.	M47T1	GCA_000263855.1	88
3	<i>Pseudomonas</i> sp.	P818	GCA_000418555.1	33
5	<i>Pseudomonas</i> sp.	PAMC25886	GCA_000242655.2	95
5	<i>Pseudomonas</i> sp.	PAMC26793	GCA_000313235.1	58
5	<i>Pseudomonas</i> sp.	R62	GCA_000257605.1	192
5	<i>Pseudomonas</i> sp.	R81	GCA_000257625.1	8
10	<i>Pseudomonas</i> sp.	S13.1.2	GCA_000292285.1	337
3	<i>Pseudomonas</i> sp.	S9	GCA_000222125.2	25
10	<i>Pseudomonas</i> sp.	TJI-51	GCA_000190455.2	1069
5	<i>Pseudomonas</i> sp.	UK4	GCA_000174915.1	2971
5	<i>Pseudomonas</i> sp.	UW4	GCA_000316175.1	1
8	<i>P. stutzeri</i>	A1501	GCA_000013785.1	1
8	<i>P. stutzeri</i>	ATCC14405	GCA_000237885.2	130
8	<i>P. stutzeri</i>	LMG11199	GCA_000219605.1	1
8	<i>P. stutzeri</i>	B1SMN1	GCA_000416345.1	78
8	<i>P. stutzeri</i>	CCUG29243	GCA_000267545.1	1
8	<i>P. stutzeri</i>	DSM10701	GCA_000279165.1	2
8	<i>P. stutzeri</i>	DSM4166	GCA_000195105.1	1
8	<i>P. stutzeri</i>	KOS6	GCA_000307775.1	79
8	<i>P. stutzeri</i>	MF28	GCA_000455665.1	91
8	<i>P. stutzeri</i>	NF13	GCA_000341615.1	82
8	<i>P. stutzeri</i>	RCH2	GCA_000327065.1	4
8	<i>P. stutzeri</i>	SDM-LAC	GCA_000235745.2	199
8	<i>P. stutzeri</i>	T13	GCA_000282955.1	71
8	<i>P. stutzeri</i>	TS44	GCA_000263395.1	78
8	<i>P. stutzeri</i>	XLDN-R	GCA_000280555.1	167
5	<i>P. synxantha</i>	BG33R	GCA_000263715.2	1
10	<i>P. taiwanensis</i>	DSM21245	GCA_000425785.1	73
3	<i>P. thermotolerans</i>	DSM14292	GCA_000364625.1	75
5	<i>P. tolaasii</i>	6264	GCA_000316215.1	688
5	<i>P. tolaasii</i>	PMS117	GCA_000276565.1	357
5	<i>P. umsongensis</i>	Pum-20MFCvi1.1	GCA_000377725.1	27
5	<i>P. veronii</i>	1YdBTEX2	GCA_000350565.1	63
10	<i>P. vranovensis</i>	DSM16006	GCA_000425805.1	42
2	<i>P. aeruginosa</i>	PAO1_WT	GCA_000414275.1	843
10	<i>P. monteilii</i>	SB3078	GCA_000510285.1	1
10	<i>P. monteilii</i>	SB3101	GCA_000510325.1	1

5	<i>P. brassicacearum</i>	DF41	GCA_000585995.1	1
5	<i>P. brassicacearum</i>	51MFCVI2.1	GCA_000510785.1	64
3	<i>P. alcaligenes</i>	MRY13-0052	GCA_000474255.1	237
5	<i>P. chlororaphis</i>	HT66	GCA_000597925.1	50
5	<i>P. chlororaphis</i>	YL-1	GCA_000512485.1	86
5	<i>P. chlororaphis</i>	St2	GCA_000506385.1	23
10	<i>P. taiwanensis</i>	SJ9	GCA_000500605.1	736
3	<i>P. thermotolerans</i>	J53	GCA_000513835.1	64
10	<i>P. mosselii</i>	DSM17497	GCA_000621225.1	59
10	<i>P. mosselii</i>	SJ10	GCA_000498975.1	1784
8	<i>P. chloritidismutans</i>	AW-1	GCA_000495915.1	77
3	<i>P. nitroreducens</i>	HBP1	GCA_000518065.1	212
10	<i>P. fulva</i>	DSM17717	GCA_000621265.1	54
1	<i>P. oleovorans</i>	MOIL14HWK12	GCA_000510765.1	19
5	<i>Pseudomonas sp.</i>	URHB0015	GCA_000620245.1	29
10	<i>Pseudomonas sp.</i>	URIL14HWK12:I	GCA_000483105.1	31
5	<i>Pseudomonas sp.</i>	URIL14HWK12:I	GCA_000514195.1	63
5	<i>Pseudomonas sp.</i>	URIL14HWK12:I	GCA_000514275.1	89
10	<i>Pseudomonas sp.</i>	URMO17WK12:I	GCA_000514235.1	38
5	<i>Pseudomonas sp.</i>	URMO17WK12:I	GCA_000514395.1	33
3	<i>Pseudomonas sp.</i>	URMO17WK12:I	GCA_000514255.1	20
3	<i>Pseudomonas sp.</i>	URMO17WK12:I	GCA_000514215.1	32
10	<i>Pseudomonas sp.</i>	URMO17WK12:I	GCA_000620365.1	54
1	<i>Pseudomonas sp.</i>	MOIL14HWK12:I	GCA_000510705.1	16
1	<i>Pseudomonas sp.</i>	MOIL14HWK12:I	GCA_000510745.1	22
2	<i>Pseudomonas sp.</i>	P179	GCA_000478485.2	51
10	<i>Pseudomonas sp.</i>	LAIL14HWK12:I	GCA_000514335.1	28
10	<i>Pseudomonas sp.</i>	LAIL14HWK12:I	GCA_000527055.1	54
10	<i>Pseudomonas sp.</i>	LAIL14HWK12:I	GCA_000514295.1	27
10	<i>Pseudomonas sp.</i>	LAIL14HWK12:I	GCA_000483465.1	33
10	<i>Pseudomonas sp.</i>	LAIL14HWK12:I	GCA_000514315.1	27
5	<i>Pseudomonas sp.</i>	LAMO17WK12:I	GCA_000514375.1	37
10	<i>Pseudomonas sp.</i>	LAMO17WK12:I	GCA_000514355.1	31
6	<i>Pseudomonas sp.</i>	HPB0071	GCA_000478505.2	34
5	<i>Pseudomonas sp.</i>	FH1	GCA_000510895.1	147
5	<i>Pseudomonas sp.</i>	FH4	GCA_000510915.1	64
8	<i>Pseudomonas sp.</i>	BAY1663	GCA_000582595.1	221
5	<i>Pseudomonas sp.</i>	2-92_R01	GCA_000503215.1	21
10	<i>Pseudomonas sp.</i>	VLB120	GCA_000494915.1	2
5	<i>Pseudomonas sp.</i>	TKP	GCA_000508205.1	1
10	<i>Pseudomonas sp.</i>	FGI182	GCA_000511325.1	1

8	<i>P. stutzeri</i>	28a24	GCA_000590475.1	1
5	<i>P. fluorescens</i>	FH5	GCA_000511155.1	88
2	<i>P. aeruginosa</i>	LES400	GCA_000583935.1	4
2	<i>P. aeruginosa</i>	LES431	GCA_000508765.1	1
2	<i>P. aeruginosa</i>	LESB65	GCA_000583955.1	3
2	<i>P. aeruginosa</i>	LESlike1	GCA_000583975.1	5
2	<i>P. aeruginosa</i>	LESlike4	GCA_000583995.1	3
2	<i>P. aeruginosa</i>	LESlike5	GCA_000583895.1	5
2	<i>P. aeruginosa</i>	LESlike7	GCA_000583915.1	4
2	<i>P. aeruginosa</i>	MTB-1	GCA_000504045.1	1
2	<i>P. aeruginosa</i>	PA1	GCA_000496605.1	1
2	<i>P. aeruginosa</i>	PA1R	GCA_000496645.1	1
2	<i>P. aeruginosa</i>	PAO1-VE2	GCA_000484495.1	4
2	<i>P. aeruginosa</i>	PAO1-VE13	GCA_000484545.1	4
2	<i>P. aeruginosa</i>	SCV20265	GCA_000510305.1	1
2	<i>P. aeruginosa</i>	YL84	GCA_000524595.1	1
2	<i>P. aeruginosa</i>	BK1	GCA_000572265.1	163
2	<i>P. aeruginosa</i>	BL01	GCA_000481125.1	20
2	<i>P. aeruginosa</i>	BL02	GCA_000481105.1	12
2	<i>P. aeruginosa</i>	BL03	GCA_000481085.1	22
2	<i>P. aeruginosa</i>	BL04	GCA_000481065.1	24
2	<i>P. aeruginosa</i>	BL05	GCA_000481045.1	15
2	<i>P. aeruginosa</i>	BL06	GCA_000481025.1	30
2	<i>P. aeruginosa</i>	BL07	GCA_000481005.1	29
2	<i>P. aeruginosa</i>	BL08	GCA_000480985.1	31
2	<i>P. aeruginosa</i>	BL09	GCA_000480965.1	29
2	<i>P. aeruginosa</i>	BL10	GCA_000480945.1	14
2	<i>P. aeruginosa</i>	BL11	GCA_000480925.1	30
2	<i>P. aeruginosa</i>	BL12	GCA_000480905.1	42
2	<i>P. aeruginosa</i>	BL13	GCA_000480885.1	32
2	<i>P. aeruginosa</i>	BL14	GCA_000480865.1	24
2	<i>P. aeruginosa</i>	BL15	GCA_000480845.1	22
2	<i>P. aeruginosa</i>	BL16	GCA_000480825.1	23
2	<i>P. aeruginosa</i>	BL17	GCA_000480805.1	18
2	<i>P. aeruginosa</i>	BL18	GCA_000480785.1	18
2	<i>P. aeruginosa</i>	BL19	GCA_000480765.1	20
2	<i>P. aeruginosa</i>	BL20	GCA_000480745.1	31
2	<i>P. aeruginosa</i>	BL21	GCA_000480725.1	18
2	<i>P. aeruginosa</i>	BL22	GCA_000480705.1	32
2	<i>P. aeruginosa</i>	BL23	GCA_000480685.1	43
2	<i>P. aeruginosa</i>	BL24	GCA_000480665.1	27

2	<i>P. aeruginosa</i>	BL25	GCA_000480645.1	22
2	<i>P. aeruginosa</i>	BWHPSA001	GCA_000481685.1	34
2	<i>P. aeruginosa</i>	BWHPSA002	GCA_000481665.1	21
2	<i>P. aeruginosa</i>	BWHPSA003	GCA_000481645.1	35
2	<i>P. aeruginosa</i>	BWHPSA004	GCA_000481625.1	16
2	<i>P. aeruginosa</i>	BWHPSA005	GCA_000481605.1	19
2	<i>P. aeruginosa</i>	BWHPSA006	GCA_000481585.1	20
2	<i>P. aeruginosa</i>	BWHPSA007	GCA_000481565.1	31
2	<i>P. aeruginosa</i>	BWHPSA008	GCA_000481545.1	21
2	<i>P. aeruginosa</i>	BWHPSA009	GCA_000481525.1	44
2	<i>P. aeruginosa</i>	BWHPSA010	GCA_000481505.1	30
2	<i>P. aeruginosa</i>	BWHPSA011	GCA_000481485.1	27
2	<i>P. aeruginosa</i>	BWHPSA012	GCA_000481465.1	10
2	<i>P. aeruginosa</i>	BWHPSA013	GCA_000481445.1	18
2	<i>P. aeruginosa</i>	BWHPSA014	GCA_000481425.1	13
2	<i>P. aeruginosa</i>	BWHPSA015	GCA_000481405.1	19
2	<i>P. aeruginosa</i>	BWHPSA016	GCA_000481385.1	22
2	<i>P. aeruginosa</i>	BWHPSA017	GCA_000481365.1	19
2	<i>P. aeruginosa</i>	BWHPSA018	GCA_000481345.1	25
2	<i>P. aeruginosa</i>	BWHPSA019	GCA_000481325.1	16
2	<i>P. aeruginosa</i>	BWHPSA020	GCA_000481305.1	32
2	<i>P. aeruginosa</i>	BWHPSA021	GCA_000481285.1	13
2	<i>P. aeruginosa</i>	BWHPSA022	GCA_000481265.1	24
2	<i>P. aeruginosa</i>	BWHPSA023	GCA_000481245.1	35
2	<i>P. aeruginosa</i>	BWHPSA024	GCA_000481225.1	14
2	<i>P. aeruginosa</i>	BWHPSA025	GCA_000481205.1	14
2	<i>P. aeruginosa</i>	BWHPSA026	GCA_000481185.1	23
2	<i>P. aeruginosa</i>	BWHPSA027	GCA_000481165.1	24
2	<i>P. aeruginosa</i>	BWHPSA028	GCA_000481145.1	42
2	<i>P. aeruginosa</i>	BWHPSA037	GCA_000520455.1	51
2	<i>P. aeruginosa</i>	BWHPSA038	GCA_000520435.1	37
2	<i>P. aeruginosa</i>	BWHPSA039	GCA_000520415.1	31
2	<i>P. aeruginosa</i>	BWHPSA040	GCA_000520395.1	20
2	<i>P. aeruginosa</i>	BWHPSA041	GCA_000520375.1	28
2	<i>P. aeruginosa</i>	BWHPSA042	GCA_000520355.1	11
2	<i>P. aeruginosa</i>	BWHPSA043	GCA_000520335.1	39
2	<i>P. aeruginosa</i>	BWHPSA044	GCA_000520315.1	12
2	<i>P. aeruginosa</i>	BWHPSA045	GCA_000520295.1	30
2	<i>P. aeruginosa</i>	BWHPSA046	GCA_000520275.1	22
2	<i>P. aeruginosa</i>	BWHPSA047	GCA_000520255.1	21
2	<i>P. aeruginosa</i>	BWHPSA048	GCA_000520235.1	13

2	<i>P. aeruginosa</i>	C20	GCA_000480515.1	15
2	<i>P. aeruginosa</i>	C23	GCA_000480495.1	15
2	<i>P. aeruginosa</i>	C40	GCA_000480475.1	22
2	<i>P. aeruginosa</i>	C41	GCA_000480455.1	14
2	<i>P. aeruginosa</i>	C48	GCA_000480435.1	24
2	<i>P. aeruginosa</i>	C51	GCA_000480415.1	26
2	<i>P. aeruginosa</i>	C52	GCA_000480395.1	28
2	<i>P. aeruginosa</i>	CF127	GCA_000481945.1	20
2	<i>P. aeruginosa</i>	CF18	GCA_000481925.1	15
2	<i>P. aeruginosa</i>	CF27	GCA_000481905.1	23
2	<i>P. aeruginosa</i>	CF5	GCA_000481885.1	18
2	<i>P. aeruginosa</i>	CF614	GCA_000480355.1	23
2	<i>P. aeruginosa</i>	CF77	GCA_000480375.1	24
2	<i>P. aeruginosa</i>	CFPA39	GCA_000568235.1	98
2	<i>P. aeruginosa</i>	DHS01	GCA_000496455.1	106
2	<i>P. aeruginosa</i>	DHS29	GCA_000503175.1	267
2	<i>P. aeruginosa</i>	HB13	GCA_000215775.4	24
2	<i>P. aeruginosa</i>	HB15	GCA_000215795.4	20
2	<i>P. aeruginosa</i>	JCM14847	GCA_000615565.1	1377
2	<i>P. aeruginosa</i>	JCM5962	GCA_000615485.1	1198
2	<i>P. aeruginosa</i>	JCM6119	GCA_000615525.1	1377
2	<i>P. aeruginosa</i>	JD303	GCA_000505885.1	2002
2	<i>P. aeruginosa</i>	JD304	GCA_000505905.1	2058
2	<i>P. aeruginosa</i>	JD306	GCA_000506085.1	2797
2	<i>P. aeruginosa</i>	JD310	GCA_000505925.1	2062
2	<i>P. aeruginosa</i>	JD312	GCA_000505945.1	2017
2	<i>P. aeruginosa</i>	JD313	GCA_000506105.1	2296
2	<i>P. aeruginosa</i>	JD314	GCA_000505965.1	2323
2	<i>P. aeruginosa</i>	JD315	GCA_000505985.1	1913
2	<i>P. aeruginosa</i>	JD316	GCA_000506125.1	1882
2	<i>P. aeruginosa</i>	JD317	GCA_000506145.1	2043
2	<i>P. aeruginosa</i>	JD318	GCA_000506005.1	2028
2	<i>P. aeruginosa</i>	JD320	GCA_000506165.1	2038
2	<i>P. aeruginosa</i>	JD322	GCA_000506185.1	1868
2	<i>P. aeruginosa</i>	JD323	GCA_000506025.1	1965
2	<i>P. aeruginosa</i>	JD324	GCA_000506205.1	2181
2	<i>P. aeruginosa</i>	JD325	GCA_000506225.1	2141
2	<i>P. aeruginosa</i>	JD326	GCA_000506045.1	1982
2	<i>P. aeruginosa</i>	JD328	GCA_000506065.1	1851
2	<i>P. aeruginosa</i>	JD329	GCA_000506265.1	2090
2	<i>P. aeruginosa</i>	JD331	GCA_000506305.1	2377

2	<i>P. aeruginosa</i>	JD332	GCA_000506325.1	2187
2	<i>P. aeruginosa</i>	JD333	GCA_000506345.1	1996
2	<i>P. aeruginosa</i>	JD334	GCA_000506285.1	2128
2	<i>P. aeruginosa</i>	JD335	GCA_000506365.1	1820
2	<i>P. aeruginosa</i>	JJ692	GCA_000481805.1	17
2	<i>P. aeruginosa</i>	M8A.1	GCA_000480615.1	18
2	<i>P. aeruginosa</i>	M8A.2	GCA_000480595.1	23
2	<i>P. aeruginosa</i>	M8A.3	GCA_000480575.1	19
2	<i>P. aeruginosa</i>	M8A.4	GCA_000480555.1	20
2	<i>P. aeruginosa</i>	M9A.1	GCA_000480535.1	27
2	<i>P. aeruginosa</i>	MH27	GCA_000513235.1	31
2	<i>P. aeruginosa</i>	MW3a	GCA_000590905.1	240
2	<i>P. aeruginosa</i>	PA103	GCA_000611975.1	270
2	<i>P. aeruginosa</i>	PA99	GCA_000611995.1	142
2	<i>P. aeruginosa</i>	PK6	GCA_000505825.1	187
2	<i>P. aeruginosa</i>	PS42	GCA_000520195.1	65
2	<i>P. aeruginosa</i>	PS50	GCA_000520175.1	15
2	<i>P. aeruginosa</i>	RB-48	GCA_000568115.1	183
2	<i>P. aeruginosa</i>	S35004	GCA_000481725.1	32
2	<i>P. aeruginosa</i>	S54485	GCA_000481825.1	19
2	<i>P. aeruginosa</i>	SG17M	GCA_000568215.1	17
2	<i>P. aeruginosa</i>	U2504	GCA_000481785.1	28
2	<i>P. aeruginosa</i>	UDL	GCA_000481845.1	18
2	<i>P. aeruginosa</i>	VRFPA04	GCA_000473745.1	86
2	<i>P. aeruginosa</i>	VRFPA05	GCA_000496325.1	170
2	<i>P. aeruginosa</i>	VRFPA06	GCA_000567865.1	277
2	<i>P. aeruginosa</i>	VRFPA07	GCA_000506805.1	140
2	<i>P. aeruginosa</i>	VRFPA08	GCA_000506885.1	197
2	<i>P. aeruginosa</i>	VRFPA09	GCA_000558345.1	80
2	<i>P. aeruginosa</i>	X13273	GCA_000481705.1	23
2	<i>P. aeruginosa</i>	X24509	GCA_000481865.1	23
2	<i>P. aeruginosa</i>	Z61	GCA_000520215.1	43
2	<i>P. aeruginosa</i>	Stone130	GCA_000478465.2	47
NA	<i>Francisella sp.</i>	TX077308	GCA_000219045.1	1

es within the *Pseudomonas* genus. *P. syringae* genomes are listed in Table S1. All *P. syringae*

Total genome length	Min contig size	Max contig size	Mean contig size	Standard deviation contig size	N50	N90
6534891	1200	1412346	272288	376020	696624	158939
6813259	509	165983	17515	21522	32927	9070
6093587	540	366561	34043	46042	63156	17831
6742964	6742964	6742964	6742964	na	6742964	6742964
6719211	6719211	6719211	6719211	na	6719211	6719211
6826253	2087	398738	83247	88682	174666	42503
6667333	709	214365	19105	25147	35210	8901
6849861	48358	6801503	3424931	4775195	6801503	6801503
6267620	200	27884	3516	3524	5915	1773
6765326	517	537227	52041	69675	94805	34118
6378034	226	295722	30228	38330	70107	20473
7103922	200	118634	6945	12764	24518	4000
6794802	202	137930	11325	16253	30042	7356
6421010	6421010	6421010	6421010	na	6421010	6421010
6679829	304	246717	19589	37975	82268	16781
6729352	301	216631	16099	31391	66611	14054
6727688	300	283883	19905	37980	77595	16558
6751908	300	263611	21503	41381	87472	17666
6222097	6222097	6222097	6222097	na	6222097	6222097
6262305	6262305	6262305	6262305	na	6262305	6262305
6772720	230	295257	42330	57972	96634	31246
6536241	201	109002	12016	16609	29475	7982
6402658	6402658	6402658	6402658	na	6402658	6402658
6776640	502	279897	18023	28450	43073	10120
6357331	229	324383	32436	45447	77529	20882
6752431	200	430527	37514	80787	220004	47720
6753508	200	430567	35545	77684	206082	46015
6753803	203	430463	35735	75021	181654	44752
6601757	6601757	6601757	6601757	na	6601757	6601757
6327754	6327754	6327754	6327754	na	6327754	6327754
6260813	500	408475	44721	71308	143328	34240
6276335	504	217745	20852	35362	55690	15997
6266909	526	656496	116054	173582	393582	64313
6466492	556	1753317	340342	487378	875324	153256

6050470	227	214048	25747	35200	63467	18888
6764661	6764661	6764661	6764661	na	6764661	6764661
7059701	103	6910294	76736	720282	6910294	6910294
6513056	513	1193617	250503	372555	952559	241501
6888495	205	777004	76539	141142	278105	72757
6615955	524	454497	53355	94953	220690	42536
6588339	6588339	6588339	6588339	na	6588339	6588339
6078600	200	103575	9238	13140	22431	5567
7282206	202	212723	18161	25935	39050	10291
6492423	6492423	6492423	6492423	na	6492423	6492423
6683163	2168	772924	94130	158199	259452	83771
6383688	674	1366580	245527	362114	742168	188730
6264404	6264404	6264404	6264404	na	6264404	6264404
5798124	216	72570	3458	4935	6740	1475
6015846	208	2560249	375991	787682	1975737	995987
6043974	6043974	6043974	6043974	na	6043974	6043974
6772433	234	6398761	34205	454610	6398761	6398761
6342034	6342034	6342034	6342034	na	6342034	6342034
6074058	219	485778	58405	96506	142678	48561
6302567	200	133228	9421	18349	39478	10073
6397762	200	244287	15197	30475	68372	15271
6201615	200	94373	3071	5421	9707	1708
6977933	200	122079	8268	16352	34710	8542
6634585	202	153920	19457	29095	52486	17691
6432261	200	59136	2557	4477	8235	1041
6270436	200	111649	2465	5534	7564	936
6462805	756	1285412	359045	421421	776334	192425
6320506	200	79386	5098	10249	22718	4536
6222894	200	87928	4999	10231	21489	4293
6178008	200	71174	3335	6348	11963	1848
6393253	201	123107	9867	18582	38142	8555
6537648	6537648	6537648	6537648	na	6537648	6537648
5937575	2020	159572	9624	11475	13434	4106
6474922	1224	389290	34812	41050	53240	17133
7037729	3635	212586	46919	42145	77051	19936
6844176	202	270921	20013	43492	92012	21937
6453808	203	305829	28684	44274	63052	18857
5511386	201	382485	46315	78256	158081	65485
4823390	344	283949	39536	60248	133047	24264
7029683	201	272539	31524	48077	99373	22969
5428961	1259	1769255	301609	451349	583285	370816

4997454	414	248425	51521	53710	90258	24708
6843248	6843248	6843248	6843248	na	6843248	6843248
3022325	1256	612766	125931	167944	312453	92381
6980251	6980251	6980251	6980251	na	6980251	6980251
6666321	6666321	6666321	6666321	na	6666321	6666321
6629881	209	137817	19107	23602	40509	11688
4653777	1093	561759	172363	177759	416153	109098
5696307	5696307	5696307	5696307	na	5696307	5696307
5888780	5888780	5888780	5888780	na	5888780	5888780
6586244	200	452701	48787	84027	135677	47095
6019547	56977	5962570	3009774	4175885	5962570	5962570
6943928	201	400614	42601	57658	97948	27792
5525290	203	358023	50230	73303	150526	44570
6123259	233	511304	50191	86985	155534	50360
6086255	219	440295	104936	133107	273104	104399
6845832	6845832	6845832	6845832	na	6845832	6845832
6195154	1125	1617628	193599	305774	456500	106729
6870021	215	164530	27155	30886	55424	14591
7015457	7015457	7015457	7015457	na	7015457	7015457
6542292	208	491244	46400	76389	127825	33441
6814431	200	178975	7004	17966	44326	9421
6827647	200	415705	15518	41908	99562	26252
6814088	200	78643	6603	9999	18017	4729
6438405	6438405	6438405	6438405	na	6438405	6438405
6118103	201	194977	29700	39274	70526	18287
6368168	6368168	6368168	6368168	na	6368168	6368168
6602611	6602611	6602611	6602611	na	6602611	6602611
6265484	602	651533	80327	113804	206694	47532
7147633	425094	6722539	3573817	4452967	6722539	6722539
6179543	6179543	6179543	6179543	na	6179543	6179543
6815803	300	163724	20225	30230	62450	16363
6270867	6270867	6270867	6270867	na	6270867	6270867
6681319	304	87034	4650	5728	7012	2340
5070893	507	898281	105644	178505	363383	55439
5020393	500	2851520	456400	845988	2851520	524646
4920769	4920769	4920769	4920769	na	4920769	4920769
6541075	226	202449	25062	28365	44486	14557
6252301	201	142282	12972	21859	39590	9205
5971953	194	96109	7550	10202	17776	4540
6378970	209	189576	13898	25245	47416	12106
6525804	302	127533	16034	20752	35927	9541

6708432	220	605802	65769	98383	205318	40504
7666642	200	699925	102222	149543	253573	58411
4626799	200	265571	20030	38421	76371	13621
6551619	995	1073718	187190	274346	441598	209829
7179265	503	1848410	211155	485292	1457030	366040
5072093	485	423315	153700	129731	271276	91720
5225569	1080	2105037	275030	485734	716166	172132
5434353	5434353	5434353	5434353	na	5434353	5434353
5343081	559	655445	116154	147567	285505	99744
5072807	5072807	5072807	5072807	na	5072807	5072807
5763138	201	170827	20731	28315	50586	14538
6694413	201	357048	48511	63976	111179	34664
4950980	1049	405206	123775	126949	251120	80242
4642307	517	492049	57313	93378	186161	42395
5413273	579	351897	62222	81602	143104	39275
5512241	5512241	5512241	5512241	na	5512241	5512241
6867980	6867980	6867980	6867980	na	6867980	6867980
7074893	7074893	7074893	7074893	na	7074893	7074893
4656168	7444	618070	194007	195138	414024	137008
6923443	205	597931	29715	68209	122872	17897
5235696	302	344370	36109	53255	103741	33818
5100236	220	728296	96231	172517	312035	66763
5740939	512	138829	21912	24221	41469	13434
6239598	333	903653	231097	265401	482035	152752
5731541	5731541	5731541	5731541	na	5731541	5731541
6469780	216	173327	30956	37350	66637	20337
6394153	133451	6260702	3197077	4332621	6260702	6260702
5959964	5959964	5959964	5959964	na	5959964	5959964
6078430	6078430	6078430	6078430	na	6078430	6078430
6870827	6870827	6870827	6870827	na	6870827	6870827
5956110	80360	5875750	2978055	4097960	5875750	5875750
6181863	6181863	6181863	6181863	na	6181863	6181863
5625016	202	392913	31251	65093	135867	41113
5874759	949	820599	183587	221690	407829	212968
2994716	556	9810	2362	1117	2464	1507
5215788	203	152482	30502	34226	60264	18396
6156701	6156701	6156701	6156701	na	6156701	6156701
6304310	101858	6085449	2101437	3450265	6085449	6085449
5970749	218	254923	30464	38842	63164	14799
6284656	200	185413	24360	31745	57275	16543
5984790	5984790	5984790	5984790	na	5984790	5984790

5551505	203	368746	26819	44097	79526	24264
6363067	506	67154	7585	7470	12056	3858
6315212	532	200608	28447	34602	62812	16990
5774330	5774330	5774330	5774330	na	5774330	5774330
6023887	1366	564385	102100	114182	182474	67501
6484828	198965	6285863	3242414	4304087	6285863	6285863
6447478	516	1998389	126422	335285	709633	272687
5203017	201	513188	47301	88562	172598	42474
6457554	1325	811557	248368	234432	433195	122095
6559141	1289	1763466	285181	396026	606427	154660
7252358	371	401317	64181	83798	159797	34379
6497331	339	282564	34932	42246	60442	19122
5154320	262	615098	103087	114225	187846	83622
6908137	189	336113	52335	75738	152761	31228
6415949	258	538515	118814	131820	204989	85296
5919597	201	307571	46247	55893	93612	26548
6205326	234	373852	92617	93878	201322	55812
4869033	205	299121	24468	45446	79285	17440
5107430	500	97583	5027	7934	11837	2246
7204561	525	367439	32308	49609	71126	16506
6657346	517	192171	41871	43892	88165	23666
6550518	447	278410	51176	62140	122673	25712
6786964	317	230552	24240	28504	44590	12757
6297618	431	240258	44983	48741	106048	24011
6609272	455	185369	31473	32798	57752	17913
6517063	406	107081	16334	19138	32656	7519
6350607	535	365551	69787	77515	137130	38434
6140967	527	184453	34117	34235	59627	17363
6727223	518	207002	32816	35147	61913	19035
6615479	538	308020	40339	47527	75073	23378
6443787	500	171623	32219	32883	59542	16830
6589890	530	143192	19102	19013	31212	10941
6692143	535	378902	43176	47944	68220	24243
6489331	572	299293	39812	48429	77637	20955
6424244	618	226607	35494	39284	62804	17333
6502113	545	293542	35531	41620	68050	16583
6104277	505	294898	33913	41975	75201	18847
7287561	611	204409	31011	33984	57174	15346
6707010	623	274135	53231	53949	96213	30048
6786162	540	188591	24065	28244	39805	12873
5818772	560	107668	15154	14797	24795	7542

5646028	525	503710	88220	103190	188174	47186
5064759	196	122687	14108	22936	45906	9370
4982906	5257	804687	118641	158639	236678	56758
6956853	1830	2281743	772984	647308	996899	528640
6310633	204	421490	71712	101350	171047	61701
5089000	904	1045301	154213	227026	331887	89546
7021839	208	652747	73915	96039	143372	32250
6771790	634	629961	116755	152864	262823	58415
6583569	502	3877914	34290	330009	3877914	2446977
6225770	670	6215135	778222	2196846	6215135	6215135
4955852	509	144579	14706	19458	31618	9818
4796529	1259	809251	191862	220411	445158	206905
5716742	209	59513	5348	6401	9921	2812
5759221	201	19243	1939	1967	3069	938
6183388	6183388	6183388	6183388	na	6183388	6183388
4567418	4567418	4567418	4567418	na	4567418	4567418
4525589	525	306010	34813	55939	88830	21511
4547930	4547930	4547930	4547930	na	4547930	4547930
5318847	525	729516	68191	127493	247369	56118
4709064	4709064	4709064	4709064	na	4709064	4709064
8883182	4174118	4709064	4441591	378264	4709064	4174118
4689946	4689946	4689946	4689946	na	4689946	4689946
5014616	2534	267277	63477	63103	118233	29055
4943564	214	357020	54325	76515	128670	33098
4674667	577	286061	57009	64160	128824	32225
4600489	2804	4575057	1150123	2283294	4575057	4575057
4233213	201	209516	21273	27330	39917	11845
4648939	308	541288	65479	103800	164111	37262
4278818	552	458632	54857	85139	130852	35893
4695416	200	235018	28117	49565	105529	26279
6297495	6297495	6297495	6297495	na	6297495	6297495
5412748	1006	332583	74148	79440	152033	55438
3752096	1128	310662	50028	72786	159554	45053
6233813	203	77809	9061	9538	15117	4713
7007821	201	353726	19630	42610	85555	17240
6458660	1113	811375	239210	239933	432848	122495
6680036	323	614100	106033	123681	229690	52750
5696307	1191	769681	135627	176018	228099	97718
6058177	211	149418	7187	11125	15831	3466
6000087	6000087	6000087	6000087	na	6000087	6000087
5945120	5945120	5945120	5945120	na	5945120	5945120

6652396	6652396	6652396	6652396	na	6652396	6652396
6564125	1210	435022	102565	107713	187784	61801
6876944	105	297772	29017	40372	64175	18218
7298618	1991	726830	145973	191747	410036	94769
6800976	252	615281	79082	115799	251854	48020
6590922	236	2489659	286562	591672	1449739	399220
6253055	399	178072	8496	13941	16485	3748
3753103	1096	283077	58643	66427	109861	36566
6255184	667	615386	106021	128068	210265	60736
5800006	297	56144	3252	4756	5872	1310
5056349	885	221301	65667	51185	103970	38945
7408198	200	292824	34945	48250	88764	20258
4766174	1020	568280	88263	118447	178183	74621
5003883	1533	1025370	263363	270445	516210	172589
6207020	1630	733861	214036	202101	365893	148844
5816732	1389	811972	187637	215682	399722	153983
6336549	1002	796202	100581	156308	312381	67830
6880294	1017	750198	77307	116182	208224	57889
4689301	1136	628184	123403	139871	258358	84251
6568341	1053	581037	199041	177790	360238	182720
4951748	1156	782076	247588	226308	486738	120511
4907807	1118	1152873	153369	250980	417125	76713
5748894	1271	453566	106461	109966	226795	68462
5005065	172	1025364	312817	347545	693389	336740
5037709	1369	672335	228987	193536	393964	114899
6803956	506	1319621	133411	265576	768958	82886
5573872	1191	1015160	199067	233726	412121	191035
5589422	1008	445840	103508	126484	258375	69361
5597233	1010	957804	207305	247737	358402	191181
5949467	1202	868090	180287	229641	518811	142650
5680941	1004	967241	210406	256079	437289	201042
6927773	1751	636005	187238	169646	309142	116513
5592746	1689	957733	180412	222158	417041	148572
5659196	1224	801836	166447	203917	440144	131465
7068472	552	430342	48085	71199	129878	31280
6029517	559	413212	94212	132636	324427	88161
4865289	506	224468	22015	28585	44207	12363
6419592	2112	1349399	305695	350348	622148	161759
5966222	321653	5644569	2983111	3763870	5644569	5644569
7012672	7012672	7012672	7012672	na	7012672	7012672
5891312	5891312	5891312	5891312	na	5891312	5891312

4731359	4731359	4731359	4731359	na	4731359	4731359
6183861	592	623907	70272	96042	133959	44673
6590821	226725	4752040	1647706	2099311	4752040	533641
6550070	6550070	6550070	6550070	na	6550070	6550070
6526806	1078597	3751411	2175602	1399259	3751411	1078597
6508670	226690	2782841	1301734	965716	1516505	724686
6523853	226500	3524630	2174618	1728501	3524630	2772723
6542950	177824	3540435	1308590	1371833	3540435	1078617
6467615	177893	3726210	1616904	1508545	3726210	1078602
6580038	6580038	6580038	6580038	na	6580038	6580038
6528877	6528877	6528877	6528877	na	6528877	6528877
6309305	6309305	6309305	6309305	na	6309305	6309305
6265285	785675	3202222	1566322	1136940	3202222	785675
6265299	785683	3201137	1566325	1136388	3201137	785683
6725183	6725183	6725183	6725183	na	6725183	6725183
6433441	6433441	6433441	6433441	na	6433441	6433441
6451158	500	268018	39578	51391	100874	26378
6396889	824	2089239	319845	598998	1917094	214251
6905378	2073	1977771	575449	669932	1588422	260321
7005139	1128	1677109	318416	428771	708679	214709
7319902	35729	2235033	304996	434067	374087	162981
6359317	382	2724692	423955	735647	1318254	258299
6982977	533	1483514	232766	368028	834817	220587
6490977	761	2021158	223827	451779	911952	295717
6946790	541	1522019	224090	322236	615288	130815
6866917	685	2030904	236791	414765	804025	204714
6490069	845	1491435	463577	520699	1001212	349585
6999295	762	2218406	233310	470699	806863	281148
7420344	560	1690647	176675	361057	957800	131545
7169659	792	1605164	224052	351069	726992	129244
7101427	746	1510669	295893	438748	932770	244570
6304293	713	1956892	286559	459980	689612	129232
6783015	1077	1928796	294914	530986	1675485	193722
6865653	1952	2306285	381426	553182	509143	244844
6486858	680	2094785	360381	593986	1357314	264395
6362156	909	1372122	318108	401479	813252	263750
6705662	839	1824146	216312	469824	1698925	264247
6808954	783	1455517	378276	431474	1005415	239788
6951239	1007	1959907	217227	410162	666915	212581
7088630	269	1223280	164852	285718	674699	145004
6909604	902	1241682	255912	376746	702598	343701

6532903	1081	1390338	296951	425683	1009843	405517
6414610	578	1191720	188665	307607	684459	91758
6920040	646	2206934	329526	573962	1572778	225776
6830561	804	1240929	195159	298740	588037	143614
6273839	1473	1818578	392115	514813	887626	198808
6773255	1738	1920153	356488	488381	546964	212192
6930020	1330	2704895	346501	630527	980240	228394
6753715	600	1507203	217862	366891	848663	214479
6473493	908	2194192	308262	529956	923870	382499
6413827	840	1108340	145769	268774	509252	94971
6854990	2339	965185	228500	275149	645398	113696
7044395	645	1911029	260904	486726	1259064	192291
6476343	1017	1900310	647635	722102	1839765	724339
6681668	1160	1362378	371204	408999	811336	264879
6279266	5516	1147037	483021	391489	1008680	288721
6414708	1197	1214303	337617	357685	680396	289432
6446132	810	2133634	293006	500593	793167	345011
6904844	825	1243991	363413	416393	859179	350097
6604024	1452	1333160	264161	329553	562517	186531
6408102	508	1192896	400507	423206	995645	345746
6526146	431	664236	203943	231095	462117	125097
6482332	1313	1241483	498641	435691	839136	509431
6859577	1186	1245849	285816	343605	641716	220335
6897964	781	1471039	197085	352937	707387	261115
6744500	1314	1993461	481750	535051	776349	266258
6323768	439	3220744	451698	827152	3220744	351444
6794049	765	1461550	295394	381658	715928	116717
6868846	750	2027587	286202	465398	741007	143029
7428714	820	1280239	176875	300987	854408	149762
7197184	1282	817457	141122	207929	401385	73881
6605348	624	2516749	178523	454578	1136197	141751
6933614	1050	1283640	223665	340940	660092	136751
6388179	2145	1350096	319409	401380	923623	212588
6862763	789	1371378	245099	352731	468727	147946
5886599	1770	1581875	535146	530641	727578	350872
6865435	1014	873294	176037	239746	650022	88215
6896541	1630	2154868	574712	688544	1523291	422001
7015865	1700	911602	233863	269885	541200	188477
6899001	5484	1255252	313591	330869	714321	226073
6358806	1342	1206891	302801	384433	1008424	114817
5939118	30582	1490138	456856	466929	703301	258688

6878986	1314	2073654	458600	513761	757249	375456
6878551	1314	1718638	458571	500442	945836	302193
6602175	720	1950341	300099	494762	820648	121033
6777954	830	2261570	484140	696481	1456048	663719
6386057	1202	1263444	266086	291569	482679	192663
6905129	1133	1270170	265582	401519	1125383	300844
7075006	904	1313554	252679	365440	788806	190793
6998147	796	2232268	349908	570964	1074942	267305
6486304	3370	2026322	432421	550296	1008812	262712
6500574	1864	1209927	282634	415521	1013663	211720
6330289	5339	1385135	351683	429302	870930	148565
6765233	540	1386206	294141	408448	837307	213773
7011519	1874	1893314	292147	424755	805509	185925
6192821	205	482376	63193	80095	113134	40881
6953788	411	327587	65602	75293	159116	42375
7108323	211	214942	26623	35075	66667	20579
6534891	1200	1412346	272288	376020	696624	158939
6643276	1123	2030140	332164	560070	1386613	233119
6493851	501	82085	4716	6933	8751	1947
6073945	501	58143	5071	6842	9318	2193
6493851	501	82085	4716	6933	8751	1947
6337654	201	33352	3166	3325	5261	1529
6419289	200	28823	3120	3269	5156	1564
6613538	200	17887	2365	2295	3888	1149
6403170	200	26888	3106	3275	5416	1525
6283749	200	29506	3116	3170	5108	1563
6327911	200	21645	2757	2857	4572	1331
6210321	200	35403	2674	2747	4422	1310
6392162	200	26489	3342	3456	5728	1721
6194931	200	27392	3292	3413	5578	1660
6493711	202	46383	3179	3391	5466	1579
6211825	200	25340	3064	3229	5416	1495
6405073	200	42212	3143	3404	5333	1555
6449563	200	37280	3453	3568	5962	1769
5982541	201	23354	3045	3145	5315	1502
6413181	200	26090	2941	3032	5054	1416
6258340	200	26076	2924	3012	5101	1465
6447494	201	29855	3254	3430	5760	1648
6330257	200	23640	3420	3517	5857	1687
6410124	201	26002	3068	3164	5143	1512
6312644	200	35128	2656	2892	4429	1305

6166120	200	25742	2820	2861	4590	1414
6331120	200	24188	3172	3296	5517	1545
6713254	201	38478	3155	3436	5459	1526
6373845	200	35262	3503	3735	6269	1745
6735382	1104	1707424	396199	554192	1406148	245012
6348330	1327	1352798	352685	424565	1061115	352200
6351482	867	1936126	276152	477570	1013936	348154
6331855	979	1940218	333256	495671	753209	240096
6351656	795	2911760	317583	660631	996518	331489
6402107	719	1399876	237116	411267	1162388	113631
7155691	5223	904570	230829	252719	502923	160986
6665300	207	292428	27773	42424	79915	22309
6716795	208	214042	24878	33839	53587	13190
6343906	220	330372	44676	56138	92875	23612
6045648	207	262844	32330	45976	74692	18179
7455553	728	1276113	114701	210607	387164	87602
6879858	1121	1513867	458658	442431	733360	424979
6175056	201	501972	33744	59405	106392	26566
6974149	290	1189597	217943	357625	765176	265178
7002867	701	2244494	368572	555124	748575	219579
6876775	632	4791879	404517	1224227	4791879	1946468
7049946	718	1352640	251784	338564	684359	155836
6288577	714	1305092	349366	422577	1017436	352770
6998695	1346	934439	81381	124403	148547	34142
7050363	2377	267421	41473	42370	65494	19041
6975032	725	203506	25181	32280	49557	11350
7177216	1205	310875	51266	58023	84034	25092
7035331	2071	218731	35713	39847	70821	16765
6711239	577	6165520	83891	688576	6165520	6165520
7037692	1773	1568358	305987	397232	726489	161236
6421104	786	1354942	279179	413654	908542	520424
6453260	652	1146904	150076	231682	461579	124259
7433566	431	1339088	158161	284659	546632	114988
2035931	2035931	2035931	2035931	/	2035931	2035931

isolates are

N95

106577

5382

10578

6742964

6719211

26394

5828

6801503

1145

23177

13557

1610

4455

6421010

7111

5060

6859

8829

6222097

6262305

16400

4405

6402658

5254

12429

29362

28885

23865

6601757

6327754

25016

7638

36810

121224

12083
6764661
6910294
117624
48282
22794
6588339
3319
5535
6492423
47884
102155
6264404
964
280066
6043974
7231
6342034
34105
5058
8117
519
2557
11343
441
424
124395
936
937
559
4333
6537648
3127
11661
15153
11584
11148
30378
14884
12893
141916

17488
6843248
73718
6980251
6666321
7163
95743
5696307
5888780
26968
5962570
17078
33621
27556
88381
6845832
72476
9668
7015457
18329
4406
15235
2590
6438405
10091
6368168
6602611
26132
425094
6179543
7548
6270867
1554
41489
524646
4920769
9454
5020
2328
5426
5553

25749
41885
6898
62671
327692
67805
122450
5434353
61559
5072807
9616
20864
40300
24037
25217
5512241
6867980
7074893
66571
11161
18564
50169
7178
63108
5731541
11564
6260702
5959964
6078430
6870827
5875750
6181863
23418
151250
1336
12057
6156701
6085449
10690
10763
5984790

14713
2619
11045
5774330
42194
6285863
238682
16805
98628
117867
20913
11681
60016
17944
52357
16729
34019
10992
1073
11082
15783
13751
7549
15846
12793
5159
24177
11739
11139
14902
12001
7724
16300
12698
10918
10538
13083
10881
19831
8345
5087

33803
5578
33837
528640
48148
37619
24272
28094
2446977
6215135
4764
66310
1691
606
6183388
4567418
11586
4547930
35404
4709064
4174118
4689946
22090
26105
19613
4575057
8320
20317
19157
14220
6297495
45572
16312
3196
9705
86808
38855
43799
1949
6000087
5945120

6652396
44052
11983
35040
30797
172405
2443
21401
38742
908
27425
12043
47011
109768
136929
69592
52796
29045
69406
123136
112561
62981
44446
81298
80928
38553
79816
28400
148568
79625
96423
77400
70224
49865
22509
50093
7187
150262
321653
7012672
5891312

4731359
32500
533641
6550070
1078597
724686
2772723
226649
1078602
6580038
6528877
6309305
785675
785683
6725183
6433441
15204
161335
214119
132987
96966
227929
110446
253228
105851
78777
218192
104821
82761
80147
60970
111646
172311
236783
191661
121099
106179
174370
98464
104826
192519

182365
72730
111819
76385
150814
90535
189504
98797
200667
84029
78113
56140
357700
210168
213269
228856
100186
103944
98103
108499
73446
349472
123810
193708
205602
110655
99492
89765
100023
42704
79904
101702
64280
99064
272633
60804
350153
120543
101093
93429
150113

220169
262728
88782
267293
101331
199372
125455
106760
231120
203393
105400
118606
112467
24704
24800
12948
106577
88066
1303
1414
1303
1022
1028
769
1019
1020
888
877
1092
1074
1067
965
1006
1144
968
951
919
1079
1111
1000
849

926
1029
1006
1144
91044
100787
100537
237650
110180
85167
116829
13965
8028
15606
11826
45693
251127
19674
171838
103559
1946468
85171
122365
27205
13691
6899
17179
10327
9848
114493
94555
44895
65807

2035931

Table S3A- List of genes associated with *P. syringae* 1a crop pathogens. A total of 7538 pathogens in the 1a phylogroup. These words were selected based on the probability threshold function of the number of words and their probability from the highest number of words and a brief description of the gene is also given.

Genes	Starting position (pb)	Ending position (pb)	Chromosome	p-value	
				<5.10 ⁻⁷	<5.10 ⁻⁶
Pto_PSPTO_0873	942019	943119	Chromosome	1154	62
Pto_hopQ1-1	948455	949798	Chromosome	927	13
Pto_PSPTO_0871	939884	941068	Chromosome	914	247
Pto_hopD1	946222	948339	Chromosome	784	274
Pto_PSPTO_0874	943206	944459	Chromosome	780	
Pto_PSPTO_0875	944512	945756	Chromosome	453	133
Pto_PSPTO_0859	929071	931740	Chromosome	113	1446
Pto_PSPTO_0869	938532	938897	Chromosome	97	62
Pto_PL4584	5351817	5353451	Chromosome	85	13
Pto_PSPTO_0867	937908	938276	Chromosome	49	198
Pto_PSPTO_0861	932461	933957	Chromosome	38	207
Pto_PSPTO_0831	900988	902916	Chromosome	29	282
Pto_PSPTO_0868	938296	938445	Chromosome	25	25
Pto_PSPTO_0881	952643	953524	Chromosome	18	43
Pto_PSPTO_0854	924071	924292	Chromosome	16	22
Pto_PSPTO_B0024	6493751	6494371	pDC3000B	16	4
Pto_PSPTO_0879	950387	951844	Chromosome	15	64
Pto_PSPTO_2123	2296562	2297482	Chromosome	14	
Pto_PSPTO_0878	949832	950149	Chromosome	11	38
Pto_PSPTO_0880	951854	952603	Chromosome	10	59
Pto_ST1780	914941	916050	Chromosome	9	
Pto_hopD	5350102	5350647	Chromosome	4	9
Pto_PSPTO_0853	923620	924033	Chromosome	2	16
Pto_PSPTO_B0022	6492959	6493159	Chromosome	2	16
Pto_PSPTO_3155	3545226	3548708	Chromosome	2	4
Pto_PSPTO_0833	903831	904853	Chromosome	1	18
Pto_PDC3000APS38	6420260	6420361	pDC3000A	1	12
Pto_traU-1	6524341	6527397	pDC3000B		261
Pto_PSPTO_0863	934867	935535	Chromosome		247
Pto_traU-2	6455722	6458769	pDC3000A		246
Pto_traO-2	6451310	6452668	pDC3000A		227
Pto_traO-1	6520586	6521944	pDC3000A		227

Pto_traN-1	6519522	6520580	pDC3000A	192
Pto_traN-2	6450246	6451304	pDC3000A	192
Pto_tral-1	6518408	6518761	pDC3000A	89
Pto_trbC	6535405	6537693	pDC3000B	84
Pto_PSPTOA0068	6466495	6468783	pDC3000A	84
Pto_tral-2	6449132	6449485	pDC3000A	79
Pto_trax-2	6460276	6460878	pDC3000A	78
Pto_trax-1	6528612	6529214	pDC3000B	78
Pto_traw-2	6459072	6460286	pDC3000A	69
Pto_traw-1	6527408	6528622	pDC3000B	69
Pto_PSPTO_0862	933947	934870	Chromosome	69
Pto_trat	6455079	6455738	pDC3000A	59
Pto_tray-2	6460907	6463051	pDC3000A	54
Pto_tray-1	6529245	6531389	pDC3000B	54
Pto_PSPTO_0860	932028	932483	Chromosome	45
Pto_PDC3000BST59	6523755	6524366	pDC3000B	44
Pto_PSPTO_B0057	6513709	6518382	pDC3000B	38
Pto_PSPTOA0049	6444937	6449106	pDC3000A	38
Pto_pilT	6439881	6440372	pDC3000A	34
Pto_PSPTO_B0049	6508887	6509339	pDC3000B	34
Pto_PSPTO_0866	937527	937892	Chromosome	33
Pto_PSPTO_0131	155443	156687	Chromosome	30
Pto_sucA	2419474	2422305	Chromosome	30
Pto_traq-1	6522660	6523202	pDC3000A	29
Pto_traq-2	6453384	6453926	pDC3000A	29
Pto_trba	6464138	6465337	pDC3000A	24
Pto_PSPTO_B0072	6532474	6533676	pDC3000B	24
Pto_PSPTO_4480	5043826	5045409	Chromosome	21
Pto_PSPTO_0858	928775	929074	Chromosome	18
Pto_tram-2	6449482	6450189	pDC3000A	10
Pto_trar-1	6523272	6523673	pDC3000A	10
Pto_trar-2	6454002	6454397	pDC3000A	10
Pto_PSPTOA0029	6424865	6426817	pDC3000A	9
Pto_PSPTO_B0036	6499257	6501212	pDC3000B	9
Pto_PSPTO_5621	923360	923584	Chromosome	9
Pto_ST2073	899534	900991	Chromosome	8
Pto_PSPTO_0850	918744	919160	Chromosome	7
Pto_PSPTO_B0046	6507724	6508257	pDC3000B	7
Pto_PSPTOA0040	6438757	6439290	pDC3000A	7
Pto_PSPTO_2701	2995844	2996782	Chromosome	1
Pto_PSPTO_5629	5328266	5329000	Chromosome	1
Pto_PSPTO_B0051	6509846	6510646	pDC3000B	
Pto_tral	6440879	6441679	pDC3000A	

Pto_PSPTO_0848	916412	917785	Chromosome
Pto_PSPTO_B0052	6510643	6511824	pDC3000B
Pto_PSPTOA0045	6441676	6442857	pDC3000A
Pto_traM-1	6518788	6519465	pDC3000A
Pto_PSPTO_B0050	6509376	6509849	pDC3000B
Pto_PSPTOA0043	6440409	6440882	pDC3000A
Pto_PSPTO_3305	3738804	3739646	Chromosome
Pto_traP-1	6521941	6522651	pDC3000A
Pto_PSPTO_B0047	6508247	6508555	pDC3000B
Pto_PDC3000APL45	6439379	6439588	pDC3000A
Pto_traP-2	6452710	6453375	pDC3000A
Pto_PSPTOA0065	6463395	6464045	pDC3000A
Pto_PSPTO_B0071	6531734	6532381	pDC3000B
Pto_ST2208	911253	911690	Chromosome
Pto_PSPTO_2293	2538321	2539313	Chromosome
Pto_motB	5612121	5613146	Chromosome
Pto_trbB	6465423	6466508	pDC3000A
Pto_PSPTO_B0048	6508573	6508845	pDC3000B
Pto_PSPTOA0041	6439606	6439878	pDC3000A
Pto_PSPTO_0587	645345	647486	Chromosome
Pto_PSPTO_A0072	6400690	6402831	pDC3000A
Pto_PSPTO_0857	928074	928778	Chromosome
Pto_trak-2	6444498	6444785	pDC3000A
Pto_trak-1	6513379	6513666	pDC3000B
Pto_PDC3000AST35	6444828	6444947	pDC3000A
Pto_PSPTO_2065	2245666	2246355	Chromosome
Pto_PSPTOA0061	6458793	6459029	pDC3000A
Pto_engA	1575936	1577405	Chromosome
Pto_PSPTO_4919	5567722	5568864	Chromosome
Pto_rluA	4386649	4387284	Chromosome
Pto_argE-2	351346	352494	Chromosome
Pto_PSPTO_2705	3000944	3001777	Chromosome
Pto_PSPTO_3680	4145626	4147656	Chromosome
Pto_PSPTO_2066	2246481	2246774	Chromosome
Pto_PSPTOA0047	6443636	6444187	pDC3000A
Pto_nuc	6512559	6513068	pDC3000B
Pto_PDC3000APL53	6444224	6444469	pDC3000A
Pto_PSPTO_B0055	6513105	6513350	pDC3000B
Pto_PSPTOA0023	6421738	6422088	pDC3000A
Pto_PSPTO_B0028	6495703	6495978	pDC3000B
Pto_PSPTO_3528	3981164	3982141	Chromosome
Pto_PSPTO_3223	3624316	3625293	Chromosome
Pto_PSPTO_B0075	6534333	6535418	pDC3000B

Pto_PSPTO_0946	1022855	1024093	Chromosome
Pto_PSPTO_T14	899143	899219	Chromosome
Pto_PSPTO_5625	4827116	4828339	Chromosome
Pto_PSPTO_2672	2967912	2968397	Chromosome
Pto_PSPTO_0121	145643	146884	Chromosome
Pto_hopAO1	5347118	5348524	Chromosome
Pto_PSPTO_3505	3957995	3958534	Chromosome
Pto_nhaA-2	5890278	5891453	Chromosome
Pto_cumA	1600668	1602041	Chromosome
Pto_PSPTO_3587	4046458	4046967	Chromosome
Pto_PSPTO_0840	910795	911211	Chromosome
Pto_cfl	5282597	5284126	Chromosome
Pto_PSPTO_B0016	6487233	6487586	pDC3000B
Pto_secD	1554625	1556493	Chromosome
Pto_PDC3000AST14	6419730	6419879	pDC3000A
Pto_rarD-1	529381	530268	Chromosome
Pto_PSPTO_5010	5696522	5697277	Chromosome
Pto_PSPTO_B0001	6471137	6472510	pDC3000B
Pto_PSPTO_0847	916047	916415	Chromosome
Pto_PSPTOA0057	6454441	6454716	pDC3000A
Pto_hopT1-2	5187920	5189089	Chromosome
Pto_avrPto1	4514766	4515260	Chromosome
Pto_hopO1-3	5187067	5187573	Chromosome
Pto_hopT2	5186263	5186646	Chromosome
Pto_hopC1	648896	649705	Chromosome
Pto_hopO1-2	5189095	5189991	Chromosome
Pto_PSPTOA0036	6435330	6436523	pDC3000A
Pto_PSPTO_4595	5190048	5190437	Chromosome
Pto_ST4045	4515377	4516348	Chromosome
Pto_hopH1	647588	648244	Chromosome
Pto_stbB	6489245	6489664	pDC3000B
Pto_PSPTO_0586	644311	645321	Chromosome
Pto_PSPTOA0004	6399656	6400666	pDC3000A
Pto_PSPTO_B0029	6496265	6496519	pDC3000B
Pto_PSPTOA0071	6470005	6470433	pDC3000A
Pto_PSPTOA0021	6420417	6421082	pDC3000A
Pto_PSPTOA0035	6434560	6435261	pDC3000A
Pto_stbC	6489661	6489954	pDC3000B
Pto_PSPTO_B0030	6496566	6497402	pDC3000B
Pto_PSPTO_B0073	6533666	6534043	pDC3000B
Pto_PSPTOA0024	6422222	6422599	pDC3000A
Pto_PSPTOA0033	6431422	6431820	pDC3000A
Pto_PSPTOA0070	6469194	6469817	pDC3000A

Pto_PSPTO_0899	980143	980610	Chromosome
Pto_PSPTO_3040	3420579	3421493	Chromosome
Pto_ST3313	4513488	4513715	Chromosome
Pto_PSPTOA0069	6468811	6469233	pDC3000A
Pto_lsc-3	6430034	6431329	pDC3000A
Pto_PSPTOA0028	6424525	6424854	pDC3000A
Pto_repA	6397477	6398790	pDC3000A
Pto_PSPTO_B0035	6498917	6499246	pDC3000B
Pto_PSPTOA0025	6422646	6423482	pDC3000A
Pto_PSPTOA0037	6436564	6437910	pDC3000A
Pto_PSPTO_B0023	6493197	6493418	pDC3000B
Pto_PSPTO_0849	917806	918744	Chromosome
Pto_PSPTO_2755	3066116	3069223	Chromosome
Pto_PSPTO_5514	6280450	6280761	Chromosome
Pto_saxG	2863432	2866575	Chromosome
Pto_PSPTO_4574	5169133	5171616	Chromosome
Pto_pilD	1003193	1004065	Chromosome
Pto_PSPTO_1016	1110675	1112279	Chromosome
Pto_PSPTO_4081	4588320	4591742	Chromosome
Pto_PSPTO_0997	1087661	1089664	Chromosome
Pto_PSPTO_0900	980624	980905	Chromosome
Pto_PSPTO_3320	3754868	3755323	Chromosome
Pto_PSPTO_5300	6025145	6027532	Chromosome
Pto_irp3	2872651	2873733	Chromosome
Pto_PSPTO_1117	1230850	1231533	Chromosome
Pto_ST1406	1115021	1115488	Chromosome
Pto_PSPTO_4004	4517561	4518493	Chromosome
Pto_PSPTO_5316	6044578	6045543	Chromosome
Pto_PDC3000BST27	6501643	6501780	pDC3000B
Pto_phnH	2827760	2828356	Chromosome
Pto_pvsA	2307860	2320870	Chromosome
Pto_PSPTO_1561	1723431	1724489	Chromosome
Pto_PSPTO_4542	5130995	5131495	Chromosome
Pto_PSPTO_2330	2585606	2585998	Chromosome
Pto_PSPTO_1471	1616531	1617343	Chromosome
Pto_hisS	1572682	1573971	Chromosome
Pto_PSPTO_0165	184922	186427	Chromosome
Pto_PSPTO_1291	1420117	1421523	Chromosome
Pto_PSPTO_2602	2884461	2890634	Chromosome
Pto_PSPTO_2744	3048952	3049881	Chromosome
Pto_PSPTO_3754	4243461	4244354	Chromosome

Pto_PSPTO_4868	5514202	5516991	Chromosome
Pto_soxG-1	500142	500774	Chromosome
Pto_xylB	2996807	2998288	Chromosome
Pto_PSPTO_1220	1336284	1337144	Chromosome
Pto_PSPTO_0114	134647	136311	Chromosome
Pto_PSPTO_4006	4518980	4521406	Chromosome
Pto_PSPTO_3225	3626411	3627220	Chromosome
Pto_nuoL	3813897	3815750	Chromosome
Pto_ileS	870407	873238	Chromosome
Pto_PSPTO_0591	649880	650980	Chromosome
Pto_PSPTO_B0074	6534040	6534240	pDC3000B
Pto_PSPTO_3851	4360417	4363731	Chromosome
Pto_PDC3000BST6	6503692	6504216	pDC3000B
Pto_gcvP	1399420	1402284	Chromosome
Pto_PSPTO_5450	6201983	6202288	Chromosome
Pto_terD	1021393	1021968	Chromosome
Pto_PSPTO_0753	800138	801334	Chromosome
Pto_PSPTO_0918	1000297	1000404	Chromosome
Pto_PSPTO_2978	3348163	3348912	Chromosome
Pto_PSPTO_5513	6280007	6280228	Chromosome
Pto_copA	4428608	4430377	Chromosome
Pto_PSPTO_B0037	6501306	6501515	pDC3000B
Pto_mmsA-3	3948230	3949732	Chromosome
Pto_PSPTO_0750	796999	799197	Chromosome
Pto_PSPTO_2149	2341852	2348307	Chromosome
Pto_lpxK	4352908	4353903	Chromosome
Pto_PSPTO_0327	356256	358577	Chromosome
Pto_PSPTO_1075	1179310	1180674	Chromosome
Pto_PSPTO_3918	4432322	4432771	Chromosome
Pto_PSPTO_4243	4775685	4779365	Chromosome
Pto_cti	3034729	3037041	Chromosome
Pto_PSPTO_0919	1000418	1000942	Chromosome
Pto_rnpA	6396329	6396730	Chromosome
Pto_PSPTO_5515	6280855	6281721	Chromosome
Pto_nadA	4466782	4467840	Chromosome
Pto_PSPTO_1012	1106014	1106421	Chromosome
Pto_PSPTO_2748	3053803	3054474	Chromosome
Pto_PSPTO_1015	1109641	1110651	Chromosome
Pto_PSPTO_4884	5532595	5534901	Chromosome
Pto_shcS2	5185621	5186094	Chromosome
Pto_PSPTO_1260	1382572	1383801	Chromosome
Pto_PSPTO_1860	2034000	2034491	Chromosome
Pto_PSPTO_4191	4719830	4722967	Chromosome

Pto_ftsE	473957	474628	Chromosome
Pto_PSPTO_2658	2954411	2955004	Chromosome
Pto_PSPTO_3886	4398715	4400487	Chromosome
Pto_PSPTO_4003	4517028	4517564	Chromosome
Pto_PSPTO_4008	4521997	4522404	Chromosome
Pto_PSPTO_1965	2152245	2152601	Chromosome
Pto_PSPTO_4005	4518582	4519010	Chromosome
Pto_PSPTO_4084	4594116	4598948	Chromosome
Pto_cadA-2	6298291	6300216	Chromosome
Pto_catF	4857791	4858990	Chromosome
Pto_cheA-1	994416	996518	Chromosome
Pto_coaE	1002573	1003196	Chromosome
Pto_cobH	5523330	5523956	Chromosome
Pto_dxs	750560	752455	Chromosome
Pto_etfA-2	2407768	2408697	Chromosome
Pto_fmt	199041	199985	Chromosome
Pto_frr	1696300	1696857	Chromosome
Pto_gabP	6088042	6089433	Chromosome
Pto_gph-1	622769	623569	Chromosome
Pto_gyrB	4147	6564	Chromosome
Pto_hopAH1	986113	987381	Chromosome
Pto_hppD	4009715	4010806	Chromosome
Pto_hrcV	1539401	1541488	Chromosome
Pto_hrpR	1522289	1523233	Chromosome
Pto_hutH-3	5998330	5999877	Chromosome
Pto_infB	5056608	5059133	Chromosome
Pto_irp1	2873730	2883251	Chromosome
Pto_kdpD	2486379	2489042	Chromosome
Pto_lon-1	4204235	4206631	Chromosome
Pto_minD	4385492	4386304	Chromosome
Pto_mltR	3004371	3005276	Chromosome
Pto_murC	4973697	4975157	Chromosome
Pto_murD	4977421	4978767	Chromosome
Pto_pepA	1393501	1394991	Chromosome
Pto_phcA	5145818	5147497	Chromosome
Pto_phnD	2824697	2825707	Chromosome
Pto_phoD	1103683	1105254	Chromosome
Pto_phrB	1234091	1235539	Chromosome
Pto_ppnK	4295734	4296624	Chromosome
Pto_PSPTO_0130	154316	155446	Chromosome
Pto_PSPTO_0166	186464	187390	Chromosome

Pto_PSPTO_0479	524756	525322	Chromosome
Pto_PSPTO_0679	736029	736715	Chromosome
Pto_PSPTO_0755	802509	803432	Chromosome
Pto_PSPTO_0817	881733	882836	Chromosome
Pto_PSPTO_0947	1024154	1024888	Chromosome
Pto_PSPTO_0994	1084494	1085132	Chromosome
Pto_PSPTO_1011	1105251	1105814	Chromosome
Pto_PSPTO_1042	1141370	1142032	Chromosome
Pto_PSPTO_1152	1266241	1267428	Chromosome
Pto_PSPTO_1272	1395310	1396437	Chromosome
Pto_PSPTO_1273	1396424	1397485	Chromosome
Pto_PSPTO_1482	1627718	1631359	Chromosome
Pto_PSPTO_1530	1687418	1688266	Chromosome
Pto_PSPTO_1693	1862367	1863722	Chromosome
Pto_PSPTO_1762	1933979	1934461	Chromosome
Pto_PSPTO_1785	1955032	1955802	Chromosome
Pto_PSPTO_1786	1956613	1957335	Chromosome
Pto_PSPTO_1789	1959574	1961844	Chromosome
Pto_PSPTO_1799	1971103	1971924	Chromosome
Pto_PSPTO_1869	2043020	2043310	Chromosome
Pto_PSPTO_2159	2370963	2372933	Chromosome
Pto_PSPTO_2220	2443525	2444244	Chromosome
Pto_PSPTO_2233	2466929	2468260	Chromosome
Pto_PSPTO_2235	2469471	2470577	Chromosome
Pto_PSPTO_2236	2470632	2472905	Chromosome
Pto_PSPTO_2239	2476507	2481195	Chromosome
Pto_PSPTO_2262	2506204	2507247	Chromosome
Pto_PSPTO_2263	2507215	2507652	Chromosome
Pto_PSPTO_2463	2718060	2720171	Chromosome
Pto_PSPTO_2552	2817858	2819156	Chromosome
Pto_PSPTO_2554	2820042	2823548	Chromosome
Pto_PSPTO_2573	2842191	2843192	Chromosome
Pto_PSPTO_2590	2860393	2861493	Chromosome
Pto_PSPTO_2635	2928229	2929125	Chromosome
Pto_PSPTO_2695	2988478	2989539	Chromosome
Pto_PSPTO_2738	3041973	3043580	Chromosome
Pto_PSPTO_2739	3043594	3044937	Chromosome
Pto_PSPTO_2933	3297568	3298149	Chromosome
Pto_PSPTO_3159	3551698	3552942	Chromosome
Pto_PSPTO_3174	3566109	3566426	Chromosome
Pto_PSPTO_3547	4002957	4003892	Chromosome
Pto_PSPTO_3568	4027084	4027209	Chromosome
Pto_PSPTO_3599	4059714	4060538	Chromosome
Pto_PSPTO_3661	4127968	4130346	Chromosome

Pto_PSPTO_3818	4325920	4328763	Chromosome
Pto_PSPTO_3841	4347221	4350607	Chromosome
Pto_PSPTO_3868	4381115	4381972	Chromosome
Pto_PSPTO_3870	4382255	4383436	Chromosome
Pto_PSPTO_3911	4425254	4427311	Chromosome
Pto_PSPTO_4068	4574899	4575966	Chromosome
Pto_PSPTO_4091	4602462	4604294	Chromosome
Pto_PSPTO_4104	4623994	4626840	Chromosome
Pto_PSPTO_4234	4767071	4768585	Chromosome
Pto_PSPTO_4671	5274267	5275259	Chromosome
Pto_PSPTO_4845	5485965	5490914	Chromosome
Pto_PSPTO_5030	5727198	5733161	Chromosome
Pto_PSPTO_5119	5820666	5823896	Chromosome
Pto_PSPTO_5318	6046664	6048010	Chromosome
Pto_PSPTO_5390	6120522	6121052	Chromosome
Pto_PSPTO_5447	6198383	6198811	Chromosome
Pto_PSPTO_5525	6291286	6291936	Chromosome
Pto_recG	87268	89343	Chromosome
Pto_rhlB	1375533	1376726	Chromosome
Pto_rnt	4685766	4686443	Chromosome
Pto_rpoB	678489	682562	Chromosome
Pto_ST305	1447343	1449169	Chromosome
Pto_ST3209	2947054	2948271	Chromosome
Pto_thiG	478385	479179	Chromosome
Pto_tvrR	4033365	4033988	Chromosome
Pto_uvrC	3398488	3400311	Chromosome
Pto_PSPTO_5448	6198893	6201319	Chromosome
Pto_mgoA	6209553	6213002	Chromosome
Pto_PSPTO_2181	2401837	2402646	Chromosome
Pto_PSPTO_2986	3357472	3358248	Chromosome
Pto_PSPTO_3596	4056266	4057120	Chromosome
Pto_PSPTO_4543	5131860	5133683	Chromosome
Pto_PSPTO_5636	1228041	1228448	Chromosome
Pto_surE	1724477	1725226	Chromosome
Pto_PSPTO_0055	75949	76971	Chromosome
Pto_PSPTO_0168	187946	189514	Chromosome
Pto_PSPTO_0208	233476	234099	Chromosome
Pto_PSPTO_0245	266158	267015	Chromosome
Pto_ST307	3063954	3066119	Chromosome
Pto_fhuF	1845930	1846676	Chromosome
Pto_PSPTO_5349	6082956	6083183	Chromosome
Pto_cheB-1	989768	990844	Chromosome
Pto_PSPTO_0151	174587	174820	Chromosome

Pto_PSPTO_1445	1584963	1585787	Chromosome
Pto_PSPTO_1918	2097934	2101455	Chromosome
Pto_PSPTO_2150	2348309	2356936	Chromosome
Pto_PSPTO_2922	3285734	3286171	Chromosome
Pto_PSPTO_3480	3926984	3928585	Chromosome
Pto_PSPTO_3848	4355091	4357307	Chromosome
Pto_PSPTO_3910	4424645	4425121	Chromosome
Pto_PSPTO_5641	2978065	2978625	Chromosome
Pto_radA	5243215	5244582	Chromosome
Pto_topA	3963478	3966087	Chromosome
Pto_PSPTO_0020	25187	25537	Chromosome
Pto_PSPTO_2588	2858488	2859525	Chromosome
Pto_PSPTO_3756	4245936	4247096	Chromosome
Pto_PSPTO_5469	6230154	6232115	Chromosome
Pto_PSPTO_5542	6310816	6312435	Chromosome
Pto_cyaA	249596	252442	Chromosome
Pto_ndH	998903	1000201	Chromosome
Pto_PSPTO_0241	261343	261693	Chromosome
Pto_PSPTO_3805	4310302	4311342	Chromosome
Pto_PSPTO_4007	4521415	4521939	Chromosome
Pto_PSPTO_5388	6119310	6119549	Chromosome
Pto_PSPTO_5451	6202574	6204055	Chromosome
Pto_mtgA	476853	477563	Chromosome
Pto_prpB	2528345	2529238	Chromosome
Pto_PSPTO_0425	469321	470814	Chromosome
Pto_PSPTO_0789	856374	857201	Chromosome
Pto_PSPTO_3859	4371357	4372211	Chromosome
Pto_PSPTO_4027	4536695	4537342	Chromosome
Pto_PSPTO_4657	5261062	5261952	Chromosome
Pto_PSPTO_5534	6301280	6302317	Chromosome
Pto_psyl	4377652	4378386	Chromosome
Pto_PSPTO_1628	1784377	1784709	Chromosome
Pto_PSPTO_2093	2263954	2264487	Chromosome
Pto_PSPTO_4037	4544346	4544879	Chromosome
Pto_rpoH	475778	476632	Chromosome
Pto_avrE1	1514116	1519503	Chromosome
Pto_PSPTO_0773	822968	823888	Chromosome
Pto_PSPTO_4885	5535156	5536421	Chromosome

Pto_PSPTO_5317	6045856	6046494	Chromosome
Pto_PSPTO_1880	2052948	2053790	Chromosome
Pto_PSPTO_2674	2969889	2970737	Chromosome
Pto_PSPTO_3038	3417166	3418668	Chromosome
Pto_PSPTO_3312	3745742	3746374	Chromosome
Pto_PSPTO_5340	6068635	6070296	Chromosome
Pto_hopM1	1511461	1513599	Chromosome
Pto_PDC3000BST40	6488120	6489070	pDC3000B
Pto_PSPTO_0816	881303	881704	Chromosome
Pto_PSPTO_1672	1843614	1844357	Chromosome
Pto_PSPTO_2224	2447757	2449760	Chromosome
Pto_PSPTO_2657	2953462	2954430	Chromosome
Pto_PSPTO_2934	3298155	3299129	Chromosome
Pto_xpt	131808	132377	Chromosome
Pto_PSPTO_0435	479190	479924	Chromosome
Pto_PSPTO_0735	784409	784771	Chromosome
Pto_PSPTO_1411	1551198	1551773	Chromosome
Pto_PSPTO_1455	1598062	1600338	Chromosome
Pto_PSPTO_1524	1680211	1682739	Chromosome
Pto_PSPTO_2645	2938265	2939854	Chromosome
Pto_PSPTO_5505	6271800	6272438	Chromosome
Pto_PSPTO_B0020	6490291	6491238	pDC3000B
Pto_PSPTOA0010	6408350	6411319	pDC3000A
Pto_ST995	214448	214963	Chromosome
Pto_amt-1	240728	242065	Chromosome
Pto_astB	2005413	2006759	Chromosome
Pto_hopAS1	521328	522167	Chromosome
Pto_pepN	4290096	4292762	Chromosome
Pto_PSPTO_0171	192549	193526	Chromosome
Pto_PSPTO_0335	364946	366733	Chromosome
Pto_PSPTO_0373	408969	411536	Chromosome
Pto_PSPTO_1436	1574010	1574651	Chromosome
Pto_PSPTO_1795	1967637	1968725	Chromosome
Pto_PSPTO_2114	2289022	2289612	Chromosome
Pto_PSPTO_2148	2333960	2341837	Chromosome
Pto_PSPTO_2955	3322024	3322977	Chromosome
Pto_PSPTO_3854	4365545	4366333	Chromosome
Pto_PSPTO_3867	4380318	4381115	Chromosome
Pto_PSPTO_5125	5832123	5833748	Chromosome
Pto_PSPTO_5380	6112219	6113838	Chromosome
Pto_ST744	912154	912267	Chromosome
Pto_PSPTO_2344	2598221	2599087	Chromosome
Pto_PSPTO_5171	5887357	5888721	Chromosome

Pto_PDC3000APS66	6434433	6434519	pDC3000A
Pto_PSPTO_1303	1432654	1432893	Chromosome
Pto_PSPTO_1653	1816911	1818464	Chromosome
Pto_PSPTO_2226	2452514	2454061	Chromosome
Pto_rluE	5433842	5434474	Chromosome
Pto_catI	4859769	4860626	Chromosome
Pto_glcB-2	4016048	4018225	Chromosome
Pto_hopR1	954407	960280	Chromosome
Pto_nuoG	3808794	3811511	Chromosome
Pto_PSPTO_0400	441859	443655	Chromosome
Pto_PSPTO_0733	781888	783615	Chromosome
Pto_PSPTO_1009	1102902	1103528	Chromosome
Pto_PSPTO_1140	1251599	1252762	Chromosome
Pto_PSPTO_1796	1968984	1969841	Chromosome
Pto_PSPTO_3029	3404189	3406057	Chromosome
Pto_PSPTO_3449	3892045	3893454	Chromosome
Pto_PSPTO_3541	3997736	3998896	Chromosome
Pto_PSPTO_3820	4329972	4330532	Chromosome
Pto_PSPTO_4344	4906645	4909305	Chromosome
Pto_PSPTO_4837	5477023	5478678	Chromosome
Pto_PSPTO_5089	5791486	5791944	Chromosome
Pto_thiE	5439438	5440055	Chromosome
Pto_thrS	2627367	2629289	Chromosome
Pto_PSPTO_0523	571196	572326	Chromosome
Pto_PSPTO_0995	1085145	1086818	Chromosome
Pto_PSPTO_3390	3831622	3832665	Chromosome
Pto_PSPTO_3690	4157261	4158121	Chromosome
Pto_flhB	2158962	2160098	Chromosome
Pto_glgB	3079691	3081916	Chromosome
Pto_PSPTO_0303	328831	330219	Chromosome
Pto_PSPTO_1457	1602041	1602550	Chromosome
Pto_PSPTO_1921	2103671	2105014	Chromosome
Pto_PSPTO_3691	4158118	4159659	Chromosome
Pto_PSPTO_3785	4287862	4288419	Chromosome
Pto_araH	2934266	2935255	Chromosome
Pto_phnG	2827296	2827760	Chromosome
Pto_PSPTO_0041	56961	57347	Chromosome
Pto_PSPTO_0244	265284	265997	Chromosome
Pto_PSPTO_2479	2738331	2739158	Chromosome
Pto_PSPTO_2771	3092048	3092578	Chromosome
Pto_PSPTO_4213	4746888	4747847	Chromosome
Pto_PSPTO_4795	5434554	5436017	Chromosome

Pto_fliK	2152700	2154124	Chromosome
Pto_PSPTO_0009	14512	15741	Chromosome
Pto_PSPTO_0754	801383	802432	Chromosome
Pto_PSPTO_0988	1077641	1080460	Chromosome
Pto_PSPTO_1440	1579125	1580273	Chromosome
Pto_PSPTO_3847	4354331	4354975	Chromosome
Pto_PSPTO_5029	5726735	5727205	Chromosome
Pto_PSPTO_B0002	6472628	6473857	pDC3000B
Pto_PSPTO_B0027	6495266	6495682	pDC3000B
Pto_PSPTOA0034	6431989	6433938	pDC3000A
Pto_recQ	1800803	1802932	Chromosome
Pto_hopT1-1	6418385	6419521	Chromosome
Pto_PSPTO_0912	992732	994387	Chromosome
Pto_PSPTO_1348	1480410	1482701	Chromosome
Pto_PSPTO_3283	3710282	3710995	Chromosome
Pto_PSPTO_3539	3995310	3996725	Chromosome
Pto_PSPTO_4051	4558522	4559097	Chromosome
Pto_PSPTO_5230	5950789	5951196	Chromosome
Pto_argA	352570	353916	Chromosome
Pto_hom	1624779	1626083	Chromosome
Pto_PDC3000APS2	6397358	6397480	pDC3000A
Pto_PDC3000BST39	6474585	6474725	pDC3000B
Pto_PSPTO_0299	324439	325494	Chromosome
Pto_PSPTO_0671	726055	728361	Chromosome
Pto_PSPTO_1652	1815904	1816479	Chromosome
Pto_PSPTO_1817	1984523	1985932	Chromosome
Pto_PSPTO_1964	2150505	2152211	Chromosome
Pto_PSPTO_2254	2496461	2497936	Chromosome
Pto_PSPTO_2700	2994802	2995356	Chromosome
Pto_PSPTO_2923	3286291	3287199	Chromosome
Pto_PSPTO_3529	3982432	3983868	Chromosome
Pto_PSPTO_5217	5939068	5940018	Chromosome
Pto_PSPTO_0164	183857	184807	Chromosome
Pto_PSPTO_4734	5362019	5362246	Chromosome
Pto_PSPTO_5630	191660	192508	Chromosome
Pto_ST1660	913394	914899	Chromosome
Pto_ST2741	644170	644250	Chromosome
Pto_PSPTO_0167	187452	187874	Chromosome
Pto_PSPTO_2917	3281523	3282293	Chromosome
Pto_PSPTO_4519	5093815	5104113	Chromosome
Pto_PSPTO_5135	5844940	5845683	Chromosome

Pto_gcd-1	2834486	2836852	Chromosome
Pto_PDC3000AST43	6416075	6416368	pDC3000A
Pto_PSPTO_0019	24096	25118	Chromosome
Pto_PSPTO_0788	855610	856377	Chromosome
Pto_PSPTO_1023	1116554	1117102	Chromosome
Pto_PSPTO_1062	1158828	1159091	Chromosome
Pto_PSPTO_3520	3971652	3972113	Chromosome
Pto_PSPTO_3755	4245060	4245632	Chromosome
Pto_PSPTO_4115	4639663	4640127	Chromosome
Pto_PSPTO_4291	4833443	4835200	Chromosome
Pto_PSPTO_4967	5631682	5632344	Chromosome
Pto_PSPTO_5378	6110636	6111406	Chromosome
Pto_PSPTO_B0077	6537771	6538199	pDC3000B
Pto_recF	3039	4142	Chromosome
Pto_PL3899	4516333	4517031	Chromosome
Pto_prpC	2529613	2530740	Chromosome
Pto_PSPTO_0465	510737	511840	Chromosome
Pto_PSPTO_2225	2450057	2452378	Chromosome
Pto_PSPTO_2494	2758195	2759103	Chromosome
Pto_PSPTO_2659	2955177	2955761	Chromosome
Pto_PSPTO_2894	3254445	3256850	Chromosome
Pto_PSPTO_3720	4199188	4200798	Chromosome
Pto_PSPTO_3739	4223245	4224360	Chromosome
Pto_PSPTO_4197	4729025	4730221	Chromosome
Pto_sucD	2426551	2427432	Chromosome
Pto_PSPTO_0902	982088	983065	Chromosome
Pto_PSPTO_2013	2199398	2200060	Chromosome
Pto_PSPTO_2589	2859734	2860069	Chromosome
Pto_PSPTO_2991	3362402	3363760	Chromosome
Pto_PSPTO_3222	3623952	3624167	Chromosome
Pto_PSPTO_3236	3656657	3657052	Chromosome
Pto_ST1431	582174	583040	Chromosome
Pto_ST660	3355616	3356056	Chromosome

8 words 30pb long were identified in *Pto* DC3000 as associated to crop yield of 5.10^{-4} and were distributed in 571 genes. Genes are organized as a d lower p-value to the lowest number of words and the highest p-value. A

p-value	Description
$<5.10^{-5}$	$<5.10^{-4}$
170	31 amidinotransferase family protein
183	125 type III effector HopQ1-1
755	87 macrolide efflux protein
339	603 type III effector HopD1
135	nikkomycin biosynthesis domain protein
37	62 hypothetical protein
440	484 hypothetical protein
62	268 hypothetical protein
83	97 pseudogene HopD1
51	15 hypothetical protein
290	807 hypothetical protein
261	523 hypothetical protein
34	62 hypothetical protein
60	573 hypothetical protein
9	59 hypothetical protein
9	265 resolvase, putative
220	884 hypothetical protein
16	3 sensor histidine kinase
35	324 hypothetical protein
64	405 hypothetical protein
35	105 STY4579- putative membrane protein in <i>Salmonella enterica</i>
16	4 pseudogene HopD
19	108 dnaK suppressor protein
17	7 DnaK suppressor domain-containing protein
12	17 indolepyruvate ferredoxin oxidoreductase
1	127 ISP _y 4, transposase
25	65 predicted in CGM
776	782 traU protein
84	179 hypothetical protein
761	537 traU protein
283	236 traO protein
278	236 traO protein

271	169	traN protein
271	169	traN protein
49	101	traL protein
655	756	trbC protein
655	618	TraG/TraD family conjugal transfer protein
40	96	traL protein
28	316	traX protein
	37	traX protein
363	242	traW protein
275	220	traW protein
30	456	hypothetical protein
138	257	traT protein
152	881	traY protein
61	84	traY protein
22	406	hypothetical protein
114	183	traT protein
420	1053	DNA primase TraC
387	1263	DNA primase, putative
46	206	pilT protein
46	206	pilT protein, putative
26	186	hypothetical protein
18	25	hemY protein
		2-oxoglutarate dehydrogenase E1 component
45	234	traQ protein
45	234	traQ protein
299	315	trbA protein
299	190	hypothetical protein
1		Tat (twin-arginine translocation) pathway signal sequence domain-containing protein
41	82	hypothetical protein
140	143	traM protein
26	12	traR protein
25	12	traR protein
202	715	relaxase/mobilization nuclease MobA, putative
177	433	relaxase/mobilization nuclease MobA, putative
8	49	hypothetical protein
120	138	unnamed protein product
37	45	hypothetical protein
17	266	hypothetical protein
12	345	hypothetical protein
37		fructokinase
		insertion sequence
346	238	tral protein
346	238	relaxase

302	457	hypothetical protein
202	354	type IV pilus biogenesis protein
202	354	type IV pilus biogenesis protein
131	138	traM protein
77	185	traH protein, putative
77	185	traH protein, putative
76	62	hypothetical protein
75	129	traP protein
74	157	hypothetical protein
74	113	predicted in CGM
70	129	traP protein
67	87	exclusion-determining protein, putative
65		surface exclusion protein, putative
47	25	predicted in CGM
45	75	esterase/lipase/thioesterase family protein
45		flagellar motor protein MotB
39	293	trbB protein
39	126	putative lipoprotein
39	126	putative lipoprotein
38	551	site-specific recombinase, phage integrase family
38	391	phage integrase family site specific recombinase
37	164	hypothetical protein
37	102	traK protein
37	70	traK protein
33	14	predicted in CGM
32		enoyl-CoA hydratase
31	65	hypothetical protein
30	75	GTP-binding protein EngA
30	21	high affinity branched-chain amino acid ABC transporter periplasmic amino acid-binding protein
30	3	ribosomal large subunit pseudouridine synthase A
30		acetylornithine deacetylase
26	5	mannitol ABC transporter permease
25	19	methyl-accepting chemotaxis protein
23	30	hypothetical protein
22	266	endonuclease
22	209	endonuclease
21	141	hypothetical protein
21	125	hypothetical protein
20	181	hypothetical protein
20	164	hypothetical protein
19	822	IS52, transposase
19	769	IS52, transposase
19	262	trbB protein, putative

19	39	tellurium resistance protein
19	28	tRNA-Pro
17	6	binary cytotoxin component
15	75	NAD(P)H-flavin oxidoreductase
12	49	membrane protein
12		type III effector HopAO1
10	20	cytochrome b561
9	16	sodium-proton antiporter NhaA
9	14	multicopper oxidase
8	76	hypothetical protein
7	3	hypothetical protein
4	14	coronafacic acid synthetase, ligase component
4	11	hypothetical protein
4		protein-export membrane protein SecD
3	58	IS801, transposase, truncated
3	22	rarD protein
3		amino acid ABC transporter, periplasmic amino acid-binding protein
2	106	RepA protein, putative
2	42	hypothetical protein
1		hypothetical protein
1008		type III effector HopT1-2
884		type III effector protein AvrPto1
774		type III effector HopO1-3
678		type III effector HopT2
616		type III effector HopC1
540		type III effector HopO1-2
455		hydroxyglutarate oxidase
414		hypothetical protein
357		predicted in CGM
333		type III effector HopH1
333		plasmid stability protein StbB
302		site-specific recombinase, phage integrase family
267		phage integrase family site specific recombinase
222		hypothetical protein
215		GntR family transcriptional regulator 73307
212		resolvase, putative
206		GntR family transcriptional regulator
205		plasmid stability protein StbC
204		hypothetical protein
204		PilT domain-containing protein
201		hypothetical protein
199		hypothetical protein
193		hypothetical protein

- 189 hypothetical protein
183 LysR family transcriptional regulator
180 predicted in CGM
174 hypothetical protein
164 levansucrase
162 mobilization protein MobB, putative
156 replication protein RepA
154 mobilization protein MobB, putative
154 hypothetical protein
143 major facilitator family transporter
138 hypothetical protein
128 hypothetical protein
120 AcrB/AcrD/AcrF family protein
119 hypothetical protein
108 aliphatic isothiocyanate resistance protein SaxG;
AcrB/AcrD/AcrF family
107 hypothetical protein
106 type IV pilus prepilin peptidase PilD
105 hypothetical protein
101 Rhs family protein
94 acyltransferase family protein
83 hypothetical protein
83 membrane protein
81 penicillin amidase family protein
78 yersiniabactin synthetase, thiazolinyl reductase
component
77 hypothetical protein
76 predicted in CGM
75 hypothetical protein
73 sulfonate ABC transporter periplasmic sulfonate-
binding protein
72 predicted in CGM
72 PhnH protein
67 pyoverdine chromophore precursor synthetase
66 hypothetical protein
65 GAF domain-containing protein
62 response regulator
61 membrane protein
60 histidyl-tRNA synthetase
60 sulfatase family protein
60 sensor histidine kinase
60 yersiniabactin non-ribosomal peptide synthetase
60 LysR family transcriptional regulator
60 hypothetical protein

60 sensor histidine kinase/response regulator RetS
60 sarcosine oxidase, gamma subunit
60 xylulokinase
58 fusaric acid resistance protein
54 GGDEF domain/EAL domain protein
54 tail fiber protein H
53 ISPsy4, transposition helper protein
52 NADH:ubiquinone oxidoreductase subunit L
51 isoleucyl-tRNA synthetase
51 impB/mucB/samB family protein
51 hypothetical protein
50 hypothetical protein
49 unknown
48 glycine dehydrogenase
47 hypothetical protein
47 tellurium resistance protein TerD
46 Bcr/CflA family multidrug resistance transporter
46 hypothetical protein
46 chitin-binding protein
46 hypothetical protein
45 copper resistance protein A
45 hypothetical protein
44 methylmalonate-semialdehyde dehydrogenase
44 copper-translocating P-type ATPase
44 pyoverdine sidechain peptide synthetase III, L-Thr-L-Ser component
43 tetraacylydisaccharide 4'-kinase
43 S1 RNA binding domain-containing protein
42 O-antigen ABC transporter, ATP-binding protein
42 hypothetical protein
42 urea amidolyase-related protein
41 esterified fatty acid cis/trans isomerase
41 MOSC domain-containing protein
41 ribonuclease P protein component
40 RpiR family transcriptional regulator
39 quinolinate synthetase
39 TonB system transport protein
39 hypothetical protein
37 integrase/recombinase XerC
36 hypothetical protein
36 type III chaperone ShcS2
35 cyanate MFS transporter
35 hypothetical protein
35 hypothetical protein

34 cell division ATP-binding protein FtsE
34 hypothetical protein
33 EAL domain/GGDEF domain protein
33 hypothetical protein
33 hypothetical protein
32 Hpt domain-containing protein
32 hypothetical protein
32 mannuronan C-5-epimerase
30 cadmium-translocating P-type ATPase
30 3-oxoadipyl-CoA thiolase
30 chemotaxis sensor histidine kinase CheA
30 dephospho-CoA kinase
30 precorrin-8X methylmutase
30 deoxyxylulose-5-phosphate synthase
30 electron transfer flavoprotein subunit alpha
30 methionyl-tRNA formyltransferase
30 ribosome recycling factor
30 gamma-aminobutyrate permease
30 phosphoglycolate phosphatase
30 DNA gyrase subunit B
30 type III effector HopAH1
30 4-hydroxyphenylpyruvate dioxygenase
30 type III secretion protein HrcV
30 type III transcriptional regulator HrpR
30 histidine ammonia-lyase
30 translation initiation factor IF-2
30 yersiniabactin polyketide/non-ribosomal peptide synthetase
30 sensor protein KdpD
30 ATP-dependent protease La
30 septum site-determining protein MinD
30 transcriptional activator MltR
30 UDP-N-acetylmuramate--L-alanine ligase
30 UDP-N-acetylmuramoylalanine--D-glutamate ligase
30 cytosol aminopeptidase
30 PhcA
30 phosphonates ABC transporter periplasmic phosphonates-binding protein
30 alkaline phosphatase D
30 deoxyribodipyrimidine photolyase
30 inorganic polyphosphate/ATP-NAD kinase
30 uroporphyrin-III C-methyltransferase
30 glycine/betaine family, ABC transporter, substrate-binding protein

30 hypothetical protein
30 outer membrane protein OmpW
30 LysR family transcriptional regulator
30 oxidoreductase, FAD-binding protein
30 hypothetical protein
30 carbonic anhydrase
30 GNAT family acetyltransferase
30 transcriptional activator ChrR
30 HDIG domain protein
30 membrane protein
30 membrane protein
30 sensory box histidine kinase/response regulator
30 hypothetical protein
30 23S rRNA m(5)U1939 methyltransferase
30 hypothetical protein
30 hypothetical protein
30 hypothetical protein
30 dimethylsulfoxide reductase
30 Ser/Thr protein phosphatase family protein
30 hypothetical protein
30 macrolide ABC efflux protein
30 hypothetical protein
30 hypothetical protein
30 hypothetical protein
30 hypothetical protein
30 YD repeat protein
30 hypothetical protein
30 hypothetical protein
30 TonB-dependent siderophore receptor
30 hypothetical protein
30 hypothetical protein
30 peptide ABC transporter ATP-binding protein
30 luciferase family protein
30 cation ABC transporter substrate-binding protein
30 hypothetical protein
30 carboxyl transferase domain protein
30 acyl-CoA dehydrogenase
30 TetR family transcriptional regulator
30 mutlidrug resistance protein
30 hypothetical protein
30 LysR family transcriptional regulator
30 hypothetical protein
30 hypothetical protein
30 xanthine dehydrogenase, C-terminal subunit

30 hypothetical protein
30 ribonuclease, Rne/Rng family protein
30 hypothetical protein
30 patatin family protein
30 protease II
30 AraC family transcriptional regulator
30 ABC transporter ATP-binding protein/permease
30 ATP-dependent helicase HepA
30 membrane protein TctA
30 alkaline D-peptidase
30 lipoprotein
30 sensor histidine kinase/response regulator
30 hypothetical protein
30 outer membrane porin, OprD family
30 hypothetical protein
30 hypothetical protein
30 hypothetical protein
30 ATP-dependent DNA helicase RecG
30 ATP-dependent RNA helicase RhIB
30 ribonuclease T
30 DNA-directed RNA polymerase subunit beta
30 Fimbrial biogenesis outer membrane usher protein
30 malate:quinone-oxidoreductase
30 thiazole biosynthesis protein ThiG
30 TetR-like virulence regulator
30 excinuclease ABC subunit C
28 hypothetical protein
27 ornithine acetyl transferase inhibitor
27 IclR family transcriptional regulator
27 branched-chain amino acid ABC transporter ATP-binding protein
27 membrane protein
27 GAF domain/GGDEF domain/EAL domain protein
27 hypothetical protein
27 stationary-phase survival protein SurE
26 ISPsy4, transposase
26 sulfate transporter family protein
26 hypothetical protein
26 GNAT family acetyltransferase
26 predicted in CGM
25 ferric iron reductase protein FhuF
25 prevent-host-death family protein
24 protein-glutamate methyltransferase CheB
24 hypothetical protein

- 24 peptidase, M23/M37 family
- 24 glycosyl transferase family protein
- 24 pyoverdine sidechain peptide synthetase IV, D-Asp-L-Ser component
- 24 acetyl-CoA carboxylase, biotin carboxyl carrier protein

- 24 methyl-accepting chemotaxis protein
- 24 DNA internalization-related competence protein ComEC/Rec2
- 24 hypothetical protein
- 24 hypothetical protein
- 24 DNA repair protein RadA
- 24 DNA topoisomerase I
- 23 hypothetical protein
- 23 luciferase family protein
- 23 hypothetical protein
- 23 hypothetical protein
- 23 glucan biosynthesis protein D
- 22 adenylate cyclase
- 22 NADH dehydrogenase
- 22 hypothetical protein
- 22 AraC family transcriptional regulator
- 22 hypothetical protein
- 22 bifunctional antitoxin/transcriptional repressor RelB

- 22 hypothetical protein
- 21 monofunctional biosynthetic peptidoglycan transglycosylase
- 21 methylisocitrate lyase
- 21 hypothetical protein
- 21 phosphonate ABC transporter, permease protein
- 21 hypothetical protein
- 21 LuxR family DNA-binding response regulator
- 21 zinc metallopeptidase
- 21 hypothetical protein
- 21 autoinducer synthesis protein Psyl
- 20 hypothetical protein
- 20 lysozyme
- 20 lysozyme
- 20 RNA polymerase sigma-32 factor
- 19 type III effector protein AvrE1
- 19 LysR family transcriptional regulator
- 19 branched-chain amino acid ABC transporter substrate-binding protein

- 19 antioxidant, AhpC/Tsa family
- 18 3-hydroxyisobutyrate dehydrogenase
- 18 3-oxoadipate enol-lactone hydrolase family protein
- 18 hypothetical protein
- 18 general secretion pathway protein J
- 18 MFS permease-like protein
- 17 type III effector HopM1
- 17 transposase
- 17 type IV pilus biogenesis protein
- 17 DNA-binding response regulator
- 17 hypothetical protein
- 17 xanthine dehydrogenase accessory factor XdhC
- 17 urea amidolyase-related protein
- 17 xanthine phosphoribosyltransferase
- 16 tRNA (guanine-N(7)-)methyltransferase
- 16 LrgA family protein
- 16 hypothetical protein
- 16 hypothetical protein
- 16 glycosyl transferase family protein
- 16 hypothetical protein
- 16 transcriptional regulator
- 16 hypothetical protein
- 16 Tn3 family transposase
- 16 predicted in CGM
- 15 ammonium transporter
- 15 succinylarginine dihydrolase
- 15 type III effector HopAS1
- 15 aminopeptidase N
- 15 ISPssy, transposase
- 15 sensory box histidine kinase
- 15 Rhs family protein
- 15 hypothetical protein
- 15 alkanesulfonate monooxygenase
- 15 hypothetical protein
- 15 pyoverdine sidechain peptide synthetase II, D-Asp-L-Thr component
- 15 transcriptional activator FtrA
- 15 inner membrane transport permease
- 15 dihydrodipicolinate synthetase family protein
- 15 hypothetical protein
- 15 allophanate hydrolase
- 15 predicted in CGM
- 14 membrane protein
- 14 Atz/Trz family protein

13 predicted in CGM
13 hypothetical protein
13 hypothetical protein
13 hypothetical protein
13 ribosomal large subunit pseudouridine synthase E
12 3-oxoadipate:succinyl-CoA transferase subunit A
12 malate synthase G
12 type III effector HopR1
12 NADH dehydrogenase subunit G
12 chitinase
12 MORN repeat-containing protein
12 isochorismatase family protein
12 hypothetical protein
12 sulfonate ABC transporter permease
12 peptide ABC transporter periplasmic peptide-binding protein
12 capsular polysaccharide biosynthesis protein
12 C4-dicarboxylate transporter/malic acid transport protein
12 hypothetical protein
12 insecticidal toxin protein
12 response regulator
12 hypothetical protein
12 thiamine-phosphate pyrophosphorylase
12 threonyl-tRNA synthetase
11 Fic family protein
11 methyl-accepting chemotaxis protein
11 tail protein D
11 hypothetical protein
10 flagellar biosynthetic protein FlhB
10 1,4-alpha-glucan-branching protein
10 response regulator
10 cytidine/deoxycytidylate deaminase family protein
10 glutamine synthetase
10 ea59 protein
10 rhomboid family protein
9 L-arabinose transporter permease protein
9 phosphonate metabolism protein PhnG
9 hypothetical protein
9 hypothetical protein
9 hypothetical protein
9 GNAT family acetyltransferase
9 tRNA mo(5)U34 methyltransferase
9 AMP nucleosidase

8 flagellar hook-length control protein FliK
8 ISPsy3, transposase
8 alcohol dehydrogenase
8 hypothetical protein
8 class I and II aminotransferase
8 MotA/TolQ/ExbB proton channel family protein
8 hypothetical protein
8 ISPsy3, transposase
8 hypothetical protein
8 GGDEF domain/EAL domain-containing protein
8 ATP-dependent DNA helicase RecQ
7 type III effector HopT1-1
7 methyl-accepting chemotaxis protein
7 sensory box/GGDEF domain/EAL domain-containing protein
7 GntR family transcriptional regulator
7 membrane protein PslK
7 hypothetical protein
7 flagellar basal body-associated protein FliL-like protein
6 N-acetylglutamate synthase
6 homoserine dehydrogenase
6 hypothetical protein
6 predicted in CGM
6 monooxygenase
6 TonB-dependent receptor
6 hypothetical protein
6 aromatic amino acid permease
6 response regulator
6 methyl-accepting chemotaxis protein
6 hypothetical protein
6 nitrogen assimilation transcriptional regulator
6 capsular polysaccharide biosynthesis protein PslA
6 sigma-54-binding protein
5 LysR family transcriptional regulator
5 hypothetical protein
5 hypothetical protein
5 putative transposase
5 predicted in CGM
4 hypothetical protein
4 LamB/YcsF family protein
4 non-ribosomal peptide synthetase, terminal component
4 staphylococcal nuclease-like protein

3 glucose dehydrogenase
3 RulA
3 ISPsy4, transposase
3 phosphonate ABC transporter, permease protein
3 MotA/TolQ/ExbB proton channel family protein
3 hypothetical protein
3 universal stress protein family
3 membrane protein
3 lipoprotein SlyB
3 sensor histidine kinase
3 hypothetical protein
3 LamB/YcsF family protein
3 GntR family transcriptional regulator
3 DNA replication and repair protein RecF
2 predicted in CGM
2 2-methylcitrate synthase
2 AraC family transcriptional regulator
2 autotransporter
2 LysR family transcriptional regulator
2 NAD(P)H-flavin oxidoreductase
2 lectin repeat domain protein
2 ABC transporter ATP-binding protein
2 ABC transporter ATP-binding protein
2 hypothetical protein
2 succinyl-CoA synthetase subunit alpha
1 ISPssy, transposase
1 hypothetical protein
1 predicted in CGM
1 hypothetical protein

Table S3B- List of genes associated with *P. syringae* 2d crop pathogen isolates. A total of 100 genes were selected based on the probability of being associated with crop pathogens in the 2d phylogroup. These words were selected based on the probability organized as a function of the number of words and their probability from the highest number to the lowest p-value. A brief description of the gene is also given.

Genes	Starting position (pb)	Ending position (pb)	Chromosome	p-value	
				<5.10 ⁻⁷	<5.10 ⁻⁶
Psyr_2608	2727077	2727598	/	50	8
Psyr_4858	5684012	5684632	/	13	1
Psyr_0325	306535	307275	/	11	
Psyr_4859	5697437	5701126	/	10	
Psyr_3982	4619924	4620172	/	4	
Psyr_4860	5705997	5706944	/	3	1
Psyr_4986	5767398	5767913	/	3	
Psyr_3131	3491713	3493119	/		30
Psyr_4055	4799474	4800997	/		22
Psyr_3703	4262499	4262912	/		21
Psyr_1678	1550472	1551584	/		17
Psyr_3381	3932531	3935143	/		17
Psyr_0653	523191	524084	/		9
Psyr_2766	3246619	3247356	/		6
Psyr_3316	3876371	3877027	/		4
Psyr_0834	889493	890302	/		4
Psyr_4609	5331356	5332621	/		4
Psyr_4253	4926055	4927320	/		4
Psyr_0135	3847440	3847554	/		1
Psyr_MR79	6000888	6001334	/		
Psyr_3216	3557635	3560826	/		
Psyr_OR1	6004080	6005495	/		
Psyr_2653	2941602	2942912	/		
Psyr_2731	3224311	3224826	/		
Psyr_1860	1877991	1879526	/		
Psyr_2330	2638817	2640151	/		
Psyr_2704	3114892	3131301	/		
Psyr_5132	5988256	5989338	/		
Psyr_3722	4300383	4301270	/		
Psyr_0194	144401	145225	/		
Psyr_2530	2719821	2720396	/		

Psyr_2616	2800728	2801507	/
Psyr_3315	3852819	3853109	/
Psyr_1093	1185994	1186458	/
Psyr_1376	1445063	1446310	/
Psyr_0115	969825	969939	/
Psyr_1976	2058369	2067563	/
Psyr_4314	4998639	4999121	/
Psyr_4997	5834418	5834795	/
Psyr_0346	342094	343092	/
Psyr_1257	1226571	1227557	/
Psyr_1785	1551681	1552526	/
Psyr_4805	5577665	5577982	/
Psyr_2059	2294713	2296194	/
Psyr_3579	4039846	4041279	/
Psyr_4990	5767943	5768749	/
Psyr_4921	5747721	5750000	/
Psyr_1853	1806495	1807214	/
Psyr_2052	2293422	2293937	/
Psyr_1977	2130979	2132070	/
Psyr_3977	4618847	4619611	/
Psyr_4606	5297802	5298662	/
Psyr_0324	293509	294450	/
Psyr_3879	4515853	4516818	/
Psyr_0143	44004	45059	/
Psyr_3047	3489280	3491607	/
Psyr_3798	4321250	4322692	/
Psyr_1064	1157518	1158321	/
Psyr_0551	519207	520478	/
Psyr_0826	750397	751362	/
Psyr_0872	892574	894691	/
Psyr_4032	4721850	4722632	/
Psyr_0341	340076	341305	/
Psyr_2254	2405242	2406717	/
Psyr_4818	5604413	5605564	/
Psyr_3313	3775147	3776871	/
Psyr_1060	1147555	1147647	/
Psyr_4204	4862509	4862841	/
Psyr_2358	2713737	2714576	/
Psyr_1027	994272	995843	/

Psyr_2293	2587752	2588618	/
Psyr_2336	2664682	2665752	/
Psyr_3639	4173615	4177013	/
Psyr_4036	4735007	4735888	/
Psyr_2910	3352726	3353229	/
Psyr_4352	4999429	4999713	/
Psyr_0179	125598	126362	/
Psyr_0783	557685	561638	/
Psyr_0831	770009	771502	/
Psyr_2909	3314762	3315982	/
Psyr_4044	4792033	4793250	/
Psyr_4949	5764268	5766655	/
Psyr_2337	2665760	2666185	/
Psyr_3646	4193706	4193972	/
Psyr_4669	5461535	5461783	/
Psyr_1794	1558832	1561372	/
Psyr_2687	3062695	3091068	/
Psyr_1368	1413086	1414756	/
Psyr_1992	2136914	2137459	/
Psyr_0323	259665	260747	/
Psyr_4035	4724066	4728559	/
Psyr_4069	4810295	4815148	/
Psyr_5007	5866174	5866995	/
Psyr_0170	95194	96192	/
Psyr_0292	215460	217535	/
Psyr_0662	526901	528943	/
Psyr_0821	741222	741620	/
Psyr_1050	1146154	1147374	/
Psyr_1318	1273249	1274778	/
Psyr_1339	1359885	1360088	/
Psyr_2295	2619171	2620058	/
Psyr_3435	3964004	3964957	/
Psyr_3485	3965087	3966247	/
Psyr_4024	4691427	4693844	/
Psyr_4063	4801339	4804092	/
Psyr_4097	4828329	4829486	/
Psyr_4135	4837649	4838500	/
Psyr_4471	5070353	5071936	/

Psyr_4713	5466494	5467975	/
Psyr_5058	5921771	5923126	/
Psyr_5100	5939010	5940017	/
Psyr_5104	5940336	5941073	/
Psyr_0330	310413	311423	/
Psyr_1279	1235580	1236641	/
Psyr_1960	2035656	2037194	/
Psyr_3287	3645405	3645818	/
Psyr_0322	220906	221754	/
Psyr_1125	1200514	1201926	/
Psyr_3974	4598442	4599353	/
Psyr_0043	363735	363830	/
Psyr_4203	4852564	4853853	/
Psyr_4429	5022527	5023450	/
Psyr_2536	2726540	2726998	/
Psyr_2968	3418796	3419257	/
Psyr_2350	2681686	2683137	/
Psyr_1230	1225706	1225987	/
Psyr_2136	2312799	2313851	/
Psyr_3944	4535974	4538115	/
Psyr_4202	4847137	4848462	/
Psyr_4707	5462199	5464175	/
Psyr_0196	161034	161111	/
Psyr_4474	5112792	5113877	/
Psyr_0148	81563	82132	/
Psyr_1048	1116474	1117472	/
Psyr_2885	3297079	3298422	/
Psyr_4031	4721038	4721853	/
Psyr_4992	5782166	5783257	/
Psyr_1201	1206856	1207635	/
Psyr_2225	2337733	2338308	/
Psyr_3797	4316968	4318965	/
Psyr_4217	4885088	4886275	/
Psyr_5047	5918400	5921003	/
Psyr_0679	549707	552841	/
Psyr_0945	926782	929652	/
Psyr_1607	1495153	1497789	/
Psyr_2306	2625734	2626885	/

Psyr_2675	2950727	2951677	/
Psyr_5008	5911246	5915475	/
Psyr_0986	934924	936588	/
Psyr_1672	1517134	1517967	/
Psyr_3799	4329247	4330221	/
Psyr_2294	2614437	2615447	/
Psyr_4811	5584702	5585952	/
Psyr_0206	194094	194330	/
Psyr_0336	338837	340024	/
Psyr_0366	343725	345008	/
Psyr_0441	363882	364115	/
Psyr_0486	390409	391221	/
Psyr_1367	1388001	1389869	/
Psyr_2343	2666274	2668085	/
Psyr_4605	5205329	5207935	/
Psyr_4760	5534830	5537118	/
Psyr_3858	4514362	4515162	/
Psyr_OR6	6055558	6056601	/
Psyr_OR7	6086712	6088607	/
Psyr_0201	185020	185427	/
Psyr_0246	205070	205804	/
Psyr_0288	206567	209947	/
Psyr_0491	463498	463953	/
Psyr_1013	945989	946876	/
Psyr_2261	2495175	2495714	/
Psyr_3973	4557364	4558491	/
Psyr_4620	5333982	5335508	/
Psyr_0076	772096	773633	/
Psyr_0088	5951649	5953186	/
Psyr_3828	4411388	4431295	/
Psyr_ST1728	430658	430747	/
Psyr_1039	1080200	1081054	/
Psyr_3013	3461411	3462967	/
Psyr_3312	3754317	3755408	/
Psyr_3878	4515442	4515795	/
Psyr_0506	478811	479206	/
Psyr_0786	622501	624405	/
Psyr_1958	1887204	1887479	/
Psyr_2273	2586853	2587755	/

Psyr_3151	3527685	3531746	/
Psyr_4079	4815168	4819511	/
Psyr_0275	205820	206563	/
Psyr_0481	368978	372079	/
Psyr_2418	2714622	2715584	/
Psyr_2940	3405489	3405950	/
Psyr_OR16	6052034	6052981	/
Psyr_0440	355995	359093	/
Psyr_2248	2398036	2398314	/
Psyr_0195	155026	156522	/
Psyr_0424	349706	350386	/
Psyr_0503	477339	478118	/
Psyr_1012	941320	942753	/
Psyr_2226	2391814	2393892	/
Psyr_2351	2707924	2709108	/
Psyr_2715	3131727	3172340	/
Psyr_3502	3967320	3967673	/
Psyr_3517	3968246	3968761	/
Psyr_3636	4154856	4155668	/
Psyr_3815	4390342	4390947	/
Psyr_4287	4997122	4998558	/
Psyr_4604	5154317	5164012	/
Psyr_4726	5478853	5479857	/
Psyr_4871	5715688	5716245	/
Psyr_5061	5926344	5928011	/
Psyr_1085	1184228	1185088	/
Psyr_2048	2259273	2267903	/
Psyr_2615	2734832	2736034	/
Psyr_2845	3285422	3286450	/
Psyr_3238	3606839	3607426	/
Psyr_3616	4099140	4099925	/
Psyr_4842	5635726	5636064	/
Psyr_1084	1174500	1175477	/
Psyr_2014	2244792	2252792	/
Psyr_2831	3262145	3263539	/
Psyr_4680	5461905	5462147	/
Psyr_4795	5547340	5548176	/

5968 words (each 30pb long) were identified in *Psy* B728a as associated with a p-value threshold of 10^{-4} and were distributed in 222 genes. Genes are sorted by number of words and lower p-value to the lowest number of words and the highest p-value.

p-value	Description
$<5.10^{-5}$	$<5.10^{-4}$
23	amino acid adenylation
11	penicillin amidase
18	general substrate transporter:TonB box, N-terminal
3	ISPsy8, transposase OrfA
25	3-hydroxyacyl-CoA-ACP transferase
8	integrase catalytic subunit
31	YD repeat-containing protein secretion protein HlyD
8	hypothetical protein
9	flavodoxin/nitric oxide synthase
6	hypothetical protein
3	RND efflux system, outer membrane lipoprotein, NodT
14	regulatory protein, MerR
8	hypothetical protein
31	ISPsy8, transposase OrfA
19	glutamyl-Q tRNA(Asp) synthetase
2	anthranilate synthase component I
10	peptide chain release factor 3
9	hypothetical protein
156	C4 antisense RNA
78	hypothetical protein
78	16S ribosomal RNA
31	ISPsy8, transposase OrfA
30	3-hydroxyphenylpropionic transporter MhpT
30	ankyrin
30	FAD dependent oxidoreductase
30	UDP-glucose 4-epimerase
30	tRNA uridine 5-carboxymethylaminomethyl modification protein GidA
27	amino acid adenylation
27	short-chain dehydrogenase
26	major facilitator transporter

24	86	amino acid adenylation
19		hypothetical protein
18		permease YjgP/YjgQ
15		DNA mismatch repair protein MutS
12	11	hypothetical protein
11		ISPsy8, transposase OrfA
11		beta-ketoacyl synthase
10	11	YD repeat-containing protein
9		acriflavin resistance protein
8		2-isopropylmalate synthase
7	4	deoxyribodipyrimidine photolyase-like protein
6	6	YD repeat-containing protein
6	3	histidine kinase, HAMP region: chemotaxis sensory transducer
6	2	peptidyl-tRNA hydrolase domain-containing protein
4	5	YD repeat-containing protein
4		hypothetical protein
3	27	chorismate synthase
3	3	lipoprotein
3		glutamyl-tRNA synthetase
2	66	urea short-chain amide or branched-chain amino acid uptake ABC transporter periplasmic solute-binding protein
2		Outer membrane autotransporter barrel
1	10	AraC family transcriptional regulator
1	10	hypothetical protein
1	8	chemotaxis sensory transducer protein
1	4	helix-hairpin-helix DNA-binding motif-containing protein
	195	hypothetical protein
	128	hypothetical protein
	120	DNA topoisomerase IV subunit B
	108	glucose-6-phosphate isomerase
	105	alkaline phosphatase
	104	YD repeat-containing protein
	99	hypothetical protein
	90	phosphonate metabolism PhnJ
	77	chorismate mutase
	75	C4-dicarboxylate transporter/malic acid transport protein
	73	Sel1 repeat-containing protein
	73	hypothetical protein
	72	acetylornithine aminotransferase
	60	hypothetical protein

60 hypothetical protein
60 regulatory proteins, IclR
60 Beta-lactamase-like:RNA-metabolising metallo-beta-lactamase
60 virulence factor MVIN-like protein
54 regulatory protein, GntR
53 leucyl-tRNA synthetase
45 hypothetical protein
45 protein-glutamate O-methyltransferase
45 response regulator receiver:sigma-54 factor, interaction region
45 DEAD/DEAH box helicase
45 phospholipase D/transphosphatidylase
43 hypothetical protein
40 hypothetical protein
40 phosphoserine aminotransferase
40 insulinase-like:peptidase M16, C-terminal
39 amino acid adenylation
39 PepSY-associated TM helix family protein
38 carboxylesterase
36 zinc-containing alcohol dehydrogenase superfamily protein
35 2-octaprenyl-3-methyl-6-methoxy-1,4-benzoquinol hydroxylase
35 insecticidal toxin protein
33 sensor histidine kinase
31 aldo/keto reductase
30 RNA-binding S4
30 delta-aminolevulinic acid dehydratase
30 2OG-Fe(II) oxygenase
30 phosphoenolpyruvate-protein phosphotransferase
30 hypothetical protein
30 phosphoenolpyruvate carboxylase
30 hypothetical protein
30 ABC transporter, transmembrane region:ABC transporter
30 chemotaxis phosphatase, CheZ
30 protein-glutamate O-methyltransferase
30 hypothetical protein
30 hypothetical protein
30 cell division protein FtsZ
30 UDP-N-acetylglucosamine 1-carboxyvinyltransferase
30 hypothetical protein

30 sarcosine oxidase subunit beta
30 N-acetyltransferase GCN5
30 hypothetical protein
30 transport system permease protein
29 hypothetical protein
29 CBS:transporter-associated region
27 non-ribosomal peptide synthase:amino acid
adenylation
27 DNA topoisomerase I
26 2-octaprenyl-6-methoxyphenyl hydroxylase
26 GGDEF
25 ABC transporter
24 hypothetical protein
24 SsrA-binding protein
24 ABC transporter
23 regulatory protein LysR
23 hydrophobe/amphiphile efflux-1 HAE1
22 hypothetical protein
21 preprotein translocase subunit SecD
21 molybdenum cofactor biosynthesis protein
21 peptidase S16, ATP-dependent protease La
21 Sodium:neurotransmitter symporter
21 hypothetical protein
20 hypothetical protein
20 hypothetical protein
18 hypothetical protein
18 secretion protein HlyD
18 xylose transporter ATP-binding subunit
18 Fe-S type hydro-lyase tartrate/fumarate alpha
region:Fe-S type hydro-lyase tartrate/fumarate beta
region
18 YD repeat-containing protein
17 type III secretion protein HrpT
17 bifunctional 5,10-methylene-tetrahydrofolate
dehydrogenase/ 5,10-methylene-tetrahydrofolate
cyclohydrolase
16 NAD(P)H dehydrogenase (quinone)
16 peptidase S33, tricorn interacting factor 1
16 phosphoribosylaminoimidazole carboxylase ATPase
subunit
15 methylmalonate-semialdehyde dehydrogenase
15 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase
15 transcriptional regulator GntR
15 major facilitator transporter

15 transcriptional regulator GntR
15 transcriptional regulator GntR
14 rRNA (guanine-N(2)-)methyltransferase
14 threonine dehydratase
14 hypothetical protein
13 hypothetical protein
13 AraC family transcriptional regulator
12 hypothetical protein
12 Outer membrane autotransporter barrel
12 histidine utilization repressor
12 hypothetical protein
12 hypothetical protein
12 zinc-containing alcohol dehydrogenase superfamily protein
12 N-acetyltransferase GCN5
12 transcriptional regulator PrtN
12 hypothetical protein
11 nucleoside-specific channel-forming protein, Tsx
11 5S ribosomal RNA
11 5S ribosomal RNA
10 ATP-dependent DNA helicase RecG
9 glycine cleavage system aminomethyltransferase T
9 carbonate dehydratase
9 chemotaxis sensory transducer
9 hypothetical protein
9 peptidase M20:peptidase M20
9 binding-protein dependent transport system inner membrane protein
9 hypothetical protein
8 xanthine phosphoribosyltransferase
8 PAS:GGDEF
8 hypothetical protein
8 predicted in CGM
7 FecR protein
7 magnesium chelatase subunit ChID
7 regulatory protein LysR
7 ABC transporter
6 trifunctional transcriptional regulator/proline dehydrogenase/pyrroline-5-carboxylate dehydrogenase
6 CheW-like protein
6 non-ribosomal peptide synthase:amino acid adenylation
6 glutamate--ammonia ligase

6 type II secretion system protein E
6 hypothetical protein
5 periplasmic solute binding protein
5 dihydroorotate
5 shikimate 5-dehydrogenase
5 PAS
5 16S ribosomal RNA
4 cobyric acid a,c-diamide synthase
4 phosphonate-binding periplasmic protein
3 hypothetical protein
3 hypothetical protein
3 hypothetical protein
3 group 1 glycosyl transferase
3 formyltetrahydrofolate deformylase
3 tRNA synthetase
3 major facilitator transporter
3 hypothetical protein
3 hypothetical protein
3 polysaccharide biosynthesis protein CapD
3 hypothetical protein
3 tRNA (uracil-5)-methyltransferase
3 hypothetical protein
3 hypothetical protein
3 hypothetical protein
3 pyruvate carboxylase subunit A
2 SAM-dependent methyltransferase
2 potassium-transporting ATPase subunit B
2 amino acid adenylation
2 hypothetical protein
2 3-oxoacid CoA-transferase
2 heat shock protein HtpX
2 phosphoenolpyruvate-protein phosphotransferase
1 histone-like DNA-binding protein
1 hypothetical protein
1 hypothetical protein
1 integrase catalytic subunit
1 lysine exporter protein LysE/YggA

Table S4. T3E effectors genes used to blast against genomes. The Hop database at <http://pseudomonas-syringae.org> updated in July 2014 was used in this study. Each gene refers to an accession number in public databases. Several alleles were analyzed for some genes.

Effector	Pathovar	Accession	Source
avrA1	<i>Pgy</i>	M15194	NCBI
avrB1	<i>Pgy</i>	M21965	NCBI
avrB2	<i>Pgy</i>	M22219	NCBI
avrB3	<i>Psy</i>	AF232005	NCBI
avrB4	<i>Pan</i>	A250_04358	PPI
avrB4_1	<i>Pph</i>	AAX12108	NCBI
avrD1	<i>Pph</i>	AAV68721	NCBI
avrE1	<i>Pma</i>	AF458405	NCBI
avrPto1	<i>Psy</i>	AAY39946	NCBI
avrPto5	<i>Pan</i>	A250_19604	PPI
avrRpm1	<i>Psy</i>	AAY35802	NCBI
avrRpm1	<i>Pma</i>	X67808	NCBI
avrRpm2	<i>Pan</i>	A250_26288	PPI
avrRps4	<i>Pph</i>	AAZ38042	NCBI
avrRpt2	<i>Pto</i>	L11355	NCBI
hopA1	<i>Pst</i>	AAO58779	NCBI
hopA2	<i>Paf</i>	AAP23114	NCBI
hopAA1	<i>Pan</i>	A237_25960	NCBI
hopAB1	<i>Pgy</i>	AJ439728	NCBI
hopAB2	<i>Pto</i>	AY074795	NCBI
hopAB3	<i>Pph</i>	CP000058	NCBI
hopAD1	<i>Pto</i>	AF458398	NCBI
hopAE1	<i>Pph</i>	AAZ36433	NCBI
hopAF1	<i>Pto</i>	AAO55088	NCBI
hopAG1	<i>Pan</i>	A237_16469	NCBI
hopAH1	<i>Pan</i>	IYO_08085	PPI
hopAI1	<i>Pto</i>	AAO54440	NCBI
hopAL1	<i>Pma</i>	AF458049	NCBI
hopAM1	<i>Ppi</i>	X84843	NCBI
hopAO2	<i>Pan</i>	A250_04298	PPI
hopAQ1	<i>Pto</i>	AF458394	NCBI
hopAR1	<i>Pph</i>	M86401	NCBI

hopAS1	<i>Pph</i>	AAZ37064	NCBI
hopAT1	<i>Pph</i>	AY803994	NCBI
hopAU1	<i>Pph</i>	AY803995	NCBI
hopAV1	<i>Pph</i>	AY803996	NCBI
hopAW1	<i>Pph</i>	AY803997	NCBI
hopAY1	<i>Pan</i>	A250_04348	PPI
hopAZ1	<i>Pan</i>	A250_18013	PPI
hopB1	<i>Psy</i>	AAN85183	NCBI
hopB1	<i>Pto</i>	AF232004	NCBI
hopBB1_1	<i>Pan</i>	A250_04288	PPI
hopBH1	<i>Por</i>	EGI07563_1	NCBI
hopBI1	<i>Por</i>	EGI04913	NCBI
hopC1	<i>Ppi</i>	AJ277496	NCBI
hopD1	<i>Pph</i>	AJ277494	NCBI
hopE1	<i>Pto</i>	AY208297	NCBI
hopF1	<i>Pph</i>	AF231452	NCBI
hopF2	<i>Pto</i>	AAO54046	NCBI
hopF2	<i>Pdp</i>	AAP23118	NCBI
hopF2	<i>Pan</i>	A250_04273	PPI
hopF3	<i>Pph</i>	AAZ37227	NCBI
hopG1	<i>Pto</i>	AY208296	NCBI
hopH1	<i>Psy</i>	AAY36933	NCBI
hopI1	<i>Pph</i>	AAZ33342	NCBI
hopI1	<i>Pma</i>	AF458047	NCBI
hopK1	<i>Pto</i>	AAO53599	NCBI
hopM1	<i>Pto</i>	AAO54897	NCBI
hopM1	<i>Pph</i>	CP000058	NCBI
hopN1	<i>Pan</i>	A250_20479	PPI
hopO1_1	<i>Pto</i>	AF458392	NCBI
hopQ1_2	<i>Pto</i>	AAO58166	NCBI
hopR1	<i>Pto</i>	AF458397	NCBI
hopS1	<i>Pan</i>	A237_1150	PPI
hopS2	<i>Pto</i>	AAO58034	NCBI
hopT1_1	<i>Pto</i>	AF458399	NCBI
hopT1_2	<i>Pto</i>	AF458400	NCBI
hopT2	<i>Pto</i>	AAO58036	NCBI
hopU1	<i>Pto</i>	AAO54045	NCBI
hopV1	<i>Pph</i>	AAZ34233	NCBI
hopW1	<i>Pan</i>	A237_16224	PPI
hopX1	<i>Psy</i>	AAF71495	NCBI
hopX2	<i>Pma</i>	AF458041	NCBI

hopY1	<i>Pto</i>	AF458403	NCBI
hopZ1	<i>Pma</i>	AF458043	NCBI
hopZ2	<i>Pac</i>	DQ986428	NCBI
hopZ3	<i>Pan</i>	DQ986456	NCBI
hopZ5	<i>Pan</i>	IYO_23140	PPI

Table S5A- List of Pto DC3000 genes recombining within and between phylogroup 1a & same isolation source represented genes for which recombination hotness H_i value for all those in the top 1% interval that are shared by populations from the two isolation sources

Genes	Max Intensity of		
	Whole population	Environmental populations	Crop pathogens populations
Pto_apaH	-	2.08	-
Pto_aroh	-	2.00	-
Pto_aruG	-	1.87	-
Pto_astB	1.82	-	-
Pto_astD	2.44	1.87	2.12
Pto_betA	-	1.91	-
Pto_betB	-	2.34	-
Pto_braE	1.82	-	-
 Pto_cadA_1	-	1.80	-
Pto_cca	2.29	2.07	-
Pto_ccoP	1.99	-	-
Pto_cobQ	-	2.19	-
Pto_cyaA	1.94	1.80	-
Pto_cysA	2.30	1.91	-
Pto_cysE	-	1.98	-
Pto_cysT	2.24	-	2.44
Pto_dadX	1.95	2.08	-
Pto_dsbD	1.81	1.94	-
Pto_dxs	1.91	1.99	-
Pto_engA	2.36	2.50	-
Pto_fhuF	2.04	1.94	2.31
Pto_folC	1.97	1.84	-
Pto_folK_1	2.78	2.80	-
 Pto_gabD_2	1.99	2.21	-
Pto_gabT_2	2.13	2.16	-
Pto_gacS	-	1.97	-
Pto_ggt	-	2.10	-
Pto_glmS	-	1.98	-
Pto_glxK	2.03	-	2.42
Pto_glyA_1	1.92	-	-
Pto_gnl	1.83	-	-
Pto_hisS	2.18	2.39	-
Pto_hopAA1_2	-	-	2.19

Pto_hopAH1	-	-	1.96
Pto_hopB1	-	-	2.08
Pto_iciA	-	2.10	-
Pto_ileS	-	2.15	-
Pto_inaA	2.12	-	-
Pto_irp1	2.11	-	-
Pto_irp3	-	3.12	-
Pto_ispA	-	2.19	-
Pto_ispH	-	2.21	-
Pto_kptA	1.84	-	-
Pto_lsc_1	2.16	-	8.72
Pto_lsyA_1	2.04	2.13	-
Pto_maf_1	-	2.03	-
Pto_mdcD	1.80	-	-
Pto_metH	1.83	1.82	-
Pto_metQ_1	-	1.82	-
Pto_moeA	1.90	2.00	-
Pto_murl	2.65	2.28	1.85
Pto_nadB	-	2.03	2.47
Pto_nadE	1.85	-	-
Pto_nfrB	2.18	-	3.11
Pto_nhaA_1	2.67	-	2.07
Pto_nrdB	-	1.97	-
Pto_oprl	-	-	2.92
Pto_parE	-	1.94	-
Pto_pdxA	1.91	-	-
Pto_pgm	1.77	-	-
Pto_pgm	-	2.01	-
Pto_phab	-	2.01	-
Pto_phhA	-	-	2.05
Pto_phoQ	1.79	1.90	-
Pto_pilB	-	2.14	-
Pto_pilC	-	1.79	-
Pto_pncB	-	-	2.30
Pto_PSPTO_0092	-	1.90	-
Pto_PSPTO_0095	1.89	-	-
Pto_PSPTO_0255	2.48	2.33	-
Pto_PSPTO_0264	3.15	2.48	2.78
* Pto_PSPTO_0265	5.08	3.68	5.20
* Pto_PSPTO_0266	3.76	2.80	3.85
Pto_PSPTO_0270	-	2.01	-
Pto_PSPTO_0303	2.10	1.93	-
Pto_PSPTO_0310	2.32	-	-
Pto_PSPTO_0312	2.26	-	-

Pto_PSPTO_0313	2.62	-	2.42
Pto_PSPTO_0314	2.01	-	-
Pto_PSPTO_0319	-	-	2.07
Pto_PSPTO_0338	2.45	-	-
Pto_PSPTO_0339	2.14	2.03	-
Pto_PSPTO_0362	1.92	2.41	-
Pto_PSPTO_0364	1.83	2.63	-
Pto_PSPTO_0378	-	1.92	-
Pto_PSPTO_0379	-	2.10	-
Pto_PSPTO_0382	-	2.19	-
Pto_PSPTO_0408	-	2.13	-
Pto_PSPTO_0435	2.25	-	-
Pto_PSPTO_0452	1.81	-	2.34
Pto_PSPTO_0453	1.76	-	2.27
Pto_PSPTO_0475	-	-	2.66
Pto_PSPTO_0476	1.88	-	3.12
Pto_PSPTO_0487	-	2.02	-
Pto_PSPTO_0490	2.22	2.10	-
Pto_PSPTO_0491	2.23	2.03	-
Pto_PSPTO_0492	2.08	3.03	-
Pto_PSPTO_0493	1.97	2.53	-
Pto_PSPTO_0521	-	1.91	-
Pto_PSPTO_0547	-	1.94	-
Pto_PSPTO_0555	-	2.23	-
Pto_PSPTO_0556	2.00	2.63	-
Pto_PSPTO_0557	1.91	2.05	-
*	Pto_PSPTO_0561	4.61	2.50
	Pto_PSPTO_0746	1.97	2.18
	Pto_PSPTO_0790	-	2.06
	Pto_PSPTO_0795	1.81	1.84
	Pto_PSPTO_0820	1.99	2.05
	Pto_PSPTO_0937	-	2.01
	Pto_PSPTO_0954	1.81	1.87
Pto_PSPTO_0965	-	1.86	-
Pto_PSPTO_0990	-	1.81	-
Pto_PSPTO_0997	-	-	2.08
Pto_PSPTO_1001	1.91	-	-
Pto_PSPTO_1023	-	1.83	-
Pto_PSPTO_1047	1.74	2.15	-
Pto_PSPTO_1048	1.92	2.13	-
Pto_PSPTO_1054	1.93	2.20	-
Pto_PSPTO_1059	-	2.28	-

Pto_PSPTO_1070	-	-	2.11
Pto_PSPTO_1078	-	-	2.44
Pto_PSPTO_1115	-	1.95	-
Pto_PSPTO_1159	-	-	3.21
Pto_PSPTO_1160	-	-	3.10
Pto_PSPTO_1174	-	2.14	1.84
Pto_PSPTO_1175	-	2.04	-
Pto_PSPTO_1193	2.12	-	2.44
Pto_PSPTO_1195	2.68	1.84	3.71
Pto_PSPTO_1200	-	-	2.62
Pto_PSPTO_1202	-	-	2.02
Pto_PSPTO_1203	-	-	2.29
Pto_PSPTO_1204	1.91	-	2.55
Pto_PSPTO_1208	1.75	-	-
Pto_PSPTO_1214	1.75	2.12	-
Pto_PSPTO_1215	2.02	-	2.33
Pto_PSPTO_1217	1.96	-	-
Pto_PSPTO_1251	-	2.05	-
Pto_PSPTO_1303	1.97	-	-
Pto_PSPTO_1344	2.12	-	1.99
Pto_PSPTO_1349	2.26	2.12	-
Pto_PSPTO_1358	-	1.93	-
Pto_PSPTO_1359	-	1.90	-
Pto_PSPTO_1360	1.94	-	-
Pto_PSPTO_1440	2.48	2.14	-
Pto_PSPTO_1441	2.48	2.69	-
Pto_PSPTO_1577	1.89	1.89	1.81
Pto_PSPTO_1621	-	-	2.05
Pto_PSPTO_1639	2.09	1.89	1.81
Pto_PSPTO_1640	2.16	2.07	-
Pto_PSPTO_1642	2.52	2.30	2.70
Pto_PSPTO_1681	2.24	1.95	-
Pto_PSPTO_1730	1.82	-	-
Pto_PSPTO_1734	1.86	-	-
Pto_PSPTO_1770	-	2.11	-
Pto_PSPTO_1792	-	1.81	-
Pto_PSPTO_1861	-	1.91	-
Pto_PSPTO_1899	-	1.88	-
Pto_PSPTO_1919	-	1.90	-
Pto_PSPTO_2100	-	-	2.19
Pto_PSPTO_2105	-	1.80	-
Pto_PSPTO_2126	1.86	-	-
Pto_PSPTO_2138	-	1.81	-

Pto_PSPTO_2140	1.88	2.60	-	
Pto_PSPTO_2141	1.77	2.28	-	
Pto_PSPTO_2145	2.49	1.84	2.57	
Pto_PSPTO_2146	2.25	-	2.27	
Pto_PSPTO_2270	1.97	-	4.18	
*	Pto_PSPTO_2271	7.20	4.77	8.59
	Pto_PSPTO_2347	1.93	-	-
	Pto_PSPTO_2348	2.33	-	-
	Pto_PSPTO_2349	1.80	-	-
	Pto_PSPTO_2354	-	1.83	-
	Pto_PSPTO_2418	1.90	-	3.19
	Pto_PSPTO_2509	1.98	3.08	-
	Pto_PSPTO_2569	1.89	1.97	-
	Pto_PSPTO_2570	2.28	2.51	-
	Pto_PSPTO_2580	1.77	1.86	-
	Pto_PSPTO_2583	-	2.16	-
*	Pto_PSPTO_2587	3.26	2.53	4.40
	Pto_PSPTO_2603	-	1.95	-
	Pto_PSPTO_2637	2.10	2.13	-
	Pto_PSPTO_2639	-	-	2.06
	Pto_PSPTO_2647	1.94	1.98	-
	Pto_PSPTO_2648	2.27	1.97	-
	Pto_PSPTO_2649	2.52	2.04	-
	Pto_PSPTO_2671	1.77	-	-
	Pto_PSPTO_2673	-	2.05	-
	Pto_PSPTO_2721	-	2.15	-
	Pto_PSPTO_2738	1.77	-	-
	Pto_PSPTO_2739	2.39	2.05	-
	Pto_PSPTO_2936	2.37	2.38	-
	Pto_PSPTO_2937	2.30	1.93	-
	Pto_PSPTO_2951	-	2.33	-
	Pto_PSPTO_2952	-	1.93	-
	Pto_PSPTO_2959	-	1.82	-
	Pto_PSPTO_2962	1.98	-	1.96
	Pto_PSPTO_2984	-	2.06	-
	Pto_PSPTO_2988	-	2.09	-
	Pto_PSPTO_3029	-	2.12	-
	Pto_PSPTO_3030	2.05	2.17	-
	Pto_PSPTO_3031	-	2.13	-
	Pto_PSPTO_3032	2.49	2.75	-
	Pto_PSPTO_3036	2.14	2.14	-
	Pto_PSPTO_3037	1.78	2.12	-
	Pto_PSPTO_3111	-	2.16	-
	Pto_PSPTO_3128	1.77	-	-

Pto_PSPTO_3159	-	1.87	-
Pto_PSPTO_3165	-	2.19	-
Pto_PSPTO_3170	-	1.80	-
Pto_PSPTO_3199	-	2.03	-
Pto_PSPTO_3237	2.49	1.95	1.87
Pto_PSPTO_3244	1.85	-	-
Pto_PSPTO_3251	1.77	-	-
Pto_PSPTO_3256	2.30	1.86	-
Pto_PSPTO_3279	-	-	2.34
Pto_PSPTO_3281	-	-	2.27
Pto_PSPTO_3282	-	-	2.29
Pto_PSPTO_3283	-	-	1.97
Pto_PSPTO_3284	-	1.82	-
Pto_PSPTO_3302	-	-	2.14
Pto_PSPTO_3336	1.89	1.89	-
Pto_PSPTO_3339	1.85	-	-
Pto_PSPTO_3340	2.22	2.06	-
Pto_PSPTO_3341	1.91	-	-
Pto_PSPTO_3342	1.90	-	-
Pto_PSPTO_3343	-	2.12	-
Pto_PSPTO_3356	-	1.83	-
Pto_PSPTO_3379	-	-	2.03
Pto_PSPTO_3384	-	-	2.11
Pto_PSPTO_3593	-	1.86	-
Pto_PSPTO_3688	-	1.79	-
Pto_PSPTO_3720	-	1.82	-
Pto_PSPTO_3876	-	1.81	-
Pto_PSPTO_3911	2.56	2.46	-
Pto_PSPTO_3913	-	2.28	-
Pto_PSPTO_3917	-	1.86	-
Pto_PSPTO_4151	-	-	2.12
Pto_PSPTO_4213	-	1.91	-
Pto_PSPTO_4351	1.88	-	-
Pto_PSPTO_4352	1.89	-	-
Pto_PSPTO_4358	1.80	-	-
Pto_PSPTO_4369	-	1.83	-
Pto_PSPTO_4460	1.97	2.43	-
Pto_PSPTO_4531	1.91	-	-
Pto_PSPTO_4534	1.83	-	-
Pto_PSPTO_4559	2.41	-	2.54
Pto_PSPTO_4612	-	-	3.31
Pto_PSPTO_4665	-	1.91	-
Pto_PSPTO_4796	1.94	-	-

Pto_PSPTO_4798	1.76	-	-
Pto_PSPTO_4856	2.08	1.96	-
Pto_PSPTO_4863	1.75	2.04	-
Pto_PSPTO_4907	-	1.80	-
Pto_PSPTO_4915	2.10	1.94	-
Pto_PSPTO_4945	-	1.84	-
Pto_PSPTO_5133	2.26	-	-
Pto_PSPTO_5135	-	2.01	-
Pto_PSPTO_5159	2.09	-	-
Pto_PSPTO_5175	1.94	2.02	-
Pto_PSPTO_5184	1.93	-	-
Pto_PSPTO_5185	1.93	-	-
Pto_PSPTO_5190	2.19	1.91	-
Pto_PSPTO_5191	2.11	1.98	2.73
Pto_PSPTO_5192	2.20	-	2.31
Pto_PSPTO_5254	1.77	-	-
Pto_PSPTO_5271	-	2.02	-
Pto_PSPTO_5272	-	2.19	-
Pto_PSPTO_5273	-	2.22	-
Pto_PSPTO_5352	1.86	1.91	-
Pto_PSPTO_5395	-	1.90	-
Pto_PSPTO_5481	2.22	-	-
Pto_PSPTO_5535	2.10	2.25	-
Pto_PSPTO_5546	2.06	2.35	-
Pto_PSPTO_5547	2.74	2.84	-
Pto_PSPTO_5549	1.83	1.92	-
Pto_PSPTO_5551	2.91	2.63	2.36
* Pto_PSPTO_5552	7.99	4.01	9.94
Pto_PSPTO_5553	1.82	-	2.11
Pto_PSPTO_5554	-	-	2.10
Pto_PSPTO_5636	1.95	-	-
Pto_pta	1.94	-	-
Pto_rbsA_1	-	-	2.13
Pto_rfbA	-	-	2.52
Pto_rfbB_1	2.00	-	-
Pto_ribD	-	1.93	-
Pto_rpoD	-	1.89	-
Pto_sodA	1.92	2.26	-
Pto_soxA_1	2.09	2.33	-
Pto_sppA	-	2.25	-
Pto_terA	-	1.88	-
Pto_thyA	2.12	1.91	3.53

Pto_trpG	-	1.83	-
Pto_typA	2.55	3.07	-
Pto_tyrB	-	1.87	-
Pto_xylA	2.02	2.29	-
Pto_xylF	-	1.95	-
Pto_znuC	1.90	-	-

populations from the two isolation sources. Genes recombining within populations from the at least one SNP was in the top 2.5% interval. Stars associated to recombining genes denote

Description	Orthologs in <i>Psy</i> B728a genes
bis(5'-nucleosyl)-tetraphosphatase, symmetrical	Psyr_2373
chorismate mutase	Psyr_4818
arginine N-succinyltransferase subunit beta	Psyr_3563
succinylarginine dihydrolase	
succinylglutamic semialdehyde dehydrogenase	Psyr_3562
choline dehydrogenase	Psyr_4732
betaine aldehyde dehydrogenase BADH	Psyr_4733
high-affinity branched-chain amino acid ABC transporter, permease	
protein BraE	
cadmium-translocating P-type ATPase	Psyr_4837
tRNA nucleotidyltransferase	Psyr_4634
cytochrome c oxidase, cbb3-type subunit III	
cobyrinic acid synthase	Psyr_3676
adenylate cyclase	Psyr_0177
sulfate ABC transporter, ATP-binding protein CysA	Psyr_0081
serine O-acetyltransferase	Psyr_1235
sulfate ABC transporter permease CysT	Psyr_0083
alanine racemase	Psyr_0237
thiol:disulfide interchange protein DsbD	Psyr_4398
deoxyxylulose-5-phosphate synthase	Psyr_0604
GTP-binding protein EngA	Psyr_1252
ferric iron reductase protein FhuF	Psyr_3714
bifunctional folylpolyglutamate synthase/dihydrofolate synthase	Psyr_1665
	Psyr_4635
2-amino-4-hydroxy-6- hydroxymethylhydropteridine pyrophosphokinase	
succinate-semialdehyde dehydrogenase	Psyr_0091
4-aminobutyrate aminotransferase	Psyr_0090
sensor histidine kinase/response regulator GacS	Psyr_3698
gamma-glutamyltranspeptidase	Psyr_0691
glucosamine--fructose-6-phosphate aminotransferase	Psyr_5117
glycerate kinase	Psyr_3115
serine hydroxymethyltransferase	Psyr_4712
gluconolactonase	
histidyl-tRNA synthetase	Psyr_1249
type III effector HopAA1-1	

type III effector HopAH1	Psyr_0779
type III effector HopB1	
chromosome initiation inhibitor	Psyr_4057
isoleucyl-tRNA synthetase	Psyr_0710
inaA protein	
yersiniabactin polyketide/non-ribosomal peptide synthetase	
yersiniabactin synthetase, thiazolinyl reductase component	
geranyltransterase	Psyr_0605
hydroxymethylbutenyl pyrophosphate reductase	Psyr_0713
RNA 2'-phosphotransferase	
levansucrase	Psyr_0754
diaminopimelate decarboxylase	
maf protein	Psyr_1642
malonate decarboxylase subunit beta	
5-methyltetrahydrofolate--homocysteine methyltransferase	Psyr_2464
D-methionine-binding lipoprotein MetQ	Psyr_0351
molybdopterin biosynthesis protein MoeA	Psyr_2137
glutamate racemase	Psyr_0952
L-aspartate oxidase	Psyr_3959
NH(3)-dependent NAD(+) synthetase	
bacteriophage N4 adsorption protein B	Psyr_1030
sodium-proton antiporter NhaA	Psyr_4210
ribonucleotide-diphosphate reductase subunit beta	Psyr_3721
outer membrane lipoprotein OprI	Psyr_2071
DNA topoisomerase IV subunit B	Psyr_0551
pyridoxal phosphate biosynthetic protein PdxA	
phosphoglucomutase, alpha-D-glucose phosphate-specific	
phosphoglucomutase, alpha-D-glucose phosphate-specific	Psyr_2908
poly(3-hydroxyalkanoate) depolymerase	
phenylalanine-4-hydroxylase	Psyr_3575
sensor protein PhoQ	Psyr_3708
type IV pilus biogenesis protein PilB	Psyr_0798
type IV pilus biogenesis protein PilC	Psyr_0797
nicotinate phosphoribosyltransferase	Psyr_0595
aldehyde dehydrogenase family protein	Psyr_0227
phospholipase D	
oxidoreductase, FAD-binding protein	Psyr_0151
peptide ABC transporter ATP-binding protein	Psyr_0142
peptide ABC transporter permease	Psyr_0141
peptide ABC transporter permease	Psyr_0140
biopolymer transport protein ExbB	Psyr_0136
response regulator	Psyr_0089
sulfate ABC transporter permease CysW	
hypothetical protein	

cyclic nucleotide-binding protein	
iron ABC transporter periplasmic iron-binding protein	
type IV pilus biogenesis protein	Psyr_0249
N-acetylmuramoyl-L-alanine amidase family protein	
diguanylate cyclase	Psyr_0266
DeoR family transcriptional regulator	Psyr_4814
sugar ABC transporter periplasmic sugar-binding protein	Psyr_4812
DNA-binding heavy metal response regulator	Psyr_4800
heavy metal sensor histidine kinase	Psyr_4799
NAD-dependent DNA ligase LigB	Psyr_4796
hypothetical protein	Psyr_4768
tRNA (guanine-N(7)-)methyltransferase	
putative acyltransferase	Psyr_4724
hypothetical protein	Psyr_4722
hypothetical protein	
hypothetical protein	
ABC transporter permease	Psyr_4694
Ser/Thr protein phosphatase family protein	Psyr_4691
Tat (twin-arginine translocation) pathway signal sequence domain-containing protein	Psyr_4690
molybdate transport regulator ModE	Psyr_4689
competence protein ComF	Psyr_4688
DNA-binding protein	
hypothetical protein	Psyr_4631
hypothetical protein	Psyr_4623
nucleotidyltransferase family protein	Psyr_4622
DnaJ-like protein DjIA	Psyr_4621
hypothetical protein	
hypothetical protein	Psyr_0647
phosphonate ABC transporter, ATP-binding protein	Psyr_0694
membrane protein	Psyr_0698
AcrB/AcrD/AcrF family protein	
hypothetical protein	Psyr_0805
phosphoenolpyruvate-protein phosphotransferase, EI/HPr/EIIA components	Psyr_0821
sensor histidine kinase	Psyr_0832
hypothetical protein	Psyr_0855
acyltransferase family protein	
hypothetical protein	
MotA/TolQ/ExbB proton channel family protein	Psyr_0877
hypothetical protein	Psyr_0896
membrane protein	Psyr_0897
outer membrane porin, OprD family	Psyr_0903
methyl-accepting chemotaxis protein	Psyr_0905

lipopolysaccharide biosynthesis protein	
lipopolysaccharide biosynthesis protein	
membrane protein	Psyr_0955
bmp family protein	Psyr_1001
ABC transporter ATP-binding protein	Psyr_1002
glycoside hydrolase family protein	Psyr_1012
membrane protein	
hypothetical protein	Psyr_1029
hypothetical protein	Psyr_1031
hypothetical protein	Psyr_1034
hypothetical protein	Psyr_1036
RNA polymerase sigma-70 family protein	
regulatory protein	
regulatory protein	
membrane protein	Psyr_1042
D-isomer specific 2-hydroxyacid dehydrogenase family protein	Psyr_1043
outer membrane efflux protein	
hypothetical protein	
hypothetical protein	
hypothetical protein	
hypothetical protein	
ABC transporter permease	Psyr_1167
ABC transporter permease	Psyr_1168
bmp family protein	
class I and II aminotransferase	Psyr_1253
hydrolase	Psyr_1254
hypothetical protein	Psyr_3802
hypothetical protein	Psyr_3759
bifunctional tRNA (mnm(5)s(2)U34)-methyltransferase/FAD-dependent cmnm(5)s(2)U34 oxidoreductase	Psyr_3742
hypothetical protein	Psyr_3741
hypothetical protein	Psyr_3740
hypothetical protein	
NAD(P)H dehydrogenase, quinone family	
putative rhodanese-related sulfurtransferase	
LysR family transcriptional regulator	Psyr_3624
rhodanese-like domain-containing protein	Psyr_3602
short chain dehydrogenase	Psyr_3541
sulfatase family protein	Psyr_3506
hypothetical protein	
hypothetical protein	Psyr_1895
thiamine biosynthesis lipoprotein	Psyr_1900
yjbN family TIM-barrel protein	
ABC transporter periplasmic substrate-binding protein	Psyr_1948

cation ABC transporter ATP-binding protein	Psyr_1950
cation ABC transporter substrate-binding protein	Psyr_1951
iron-regulated membrane protein	Psyr_1955
pyoverdine biosynthesis regulatory gene	Psyr_1956
ErfK/YbiS/YcfS/YnhG family protein	Psyr_2070
hypothetical protein	
TetR family transcriptional regulator	
shikimate 5-dehydrogenase	
RNA pseudouridine synthase family protein	
glycosyl transferase family protein	Psyr_2138
branched-chain amino acid ABC transporter substrate-binding protein	Psyr_2205
penicillin-binding protein	Psyr_2313
amidase family protein	Psyr_2260
peptidase, M20/M25/M40 family	Psyr_2261
membrane protein	Psyr_2271
glutamine amidotransferase, class-II protein	Psyr_2274
hypothetical protein	
ABC transporter ATP-binding protein/permease	
senescence marker protein-30 family protein	Psyr_2370
L-arabinose transporter ATP-binding protein	Psyr_2372
hypothetical protein	
hypothetical protein	
hypothetical protein	Psyr_2381
GntR family transcriptional regulator	
aldehyde dehydrogenase family protein	Psyr_2407
membrane protein	Psyr_2454
carboxyl transferase domain protein	
acyl-CoA dehydrogenase	Psyr_2470
LamB/YcsF family protein	Psyr_2722
LysR family transcriptional regulator	Psyr_2723
sigma-54 dependent transcriptional regulator	
glycosidase	Psyr_2736
hypothetical protein	
hypothetical protein	Psyr_2747
xenobiotic compound monooxygenase, DszA family	Psyr_2869
branched-chain amino acid ABC transporter permease	
peptide ABC transporter periplasmic peptide-binding protein	Psyr_2902
peptide ABC transporter permease	Psyr_2903
peptide ABC transporter permease	Psyr_2904
peptide ABC transporter ATP-binding protein	Psyr_2905
hypothetical protein	Psyr_2909
transcriptional regulator	
sensory box histidine kinase	Psyr_2978
glycosyl hydrolase family protein	

multidrug resistance protein	Psyr_3025
lipoprotein	Psyr_3032
glycosyl transferase family protein	Psyr_3035
luciferase family protein	Psyr_3065
methyl-accepting chemotaxis protein	
AraC family transcriptional regulator	
dipeptide ABC transporter ATP-binding protein	
iron ABC transporter, periplasmic iron-binding protein	Psyr_3094
methyl-accepting chemotaxis protein	Psyr_3114
carbohydrate diacid regulator	Psyr_3116
lipoprotein	Psyr_3117
GntR family transcriptional regulator	Psyr_3118
major facilitator family transporter	Psyr_3119
AcrB/AcrD/AcrF family protein	Psyr_3132
citrate transporter	Psyr_3167
DsrF family protein	
DsrH family protein	
hypothetical protein	
hypothetical protein	
glutathione S-transferase domain protein	Psyr_3173
isocitrate dehydrogenase	Psyr_3186
methyl-accepting chemotaxis protein	
hypothetical protein	
tail tape measure protein	Psyr_3363
membrane protein	Psyr_1786
ABC transporter ATP-binding protein	Psyr_1755
aspartyl aminopeptidase	Psyr_1609
protease II	Psyr_1574
hypothetical protein	Psyr_1572
hypothetical protein	Psyr_1568
LuxR family DNA-binding response regulator	Psyr_3890
tRNA mo(5)U34 methyltransferase	Psyr_3947
hypothetical protein	
hydrolase	
oxidoreductase FAD/FMN-binding protein	
lipoprotein	Psyr_4065
hypothetical protein	Psyr_4153
methyl-accepting chemotaxis protein	
peptide ABC transporter substrate-binding protein	
dipeptide ABC transporter substrate-binding protein	Psyr_4236
LysR family transcriptional regulator	Psyr_1478
exonuclease	Psyr_4300
sensor histidine kinase	

bifunctional hydroxy-methylpyrimidine kinase/ hydroxy-phosphomethylpyrimidine kinase	
TPR domain-containing protein	Psyr_4396
hypothetical protein	
hypothetical protein	Psyr_4449
high affinity branched-chain amino acid ABC transporter ATP-binding protein	Psyr_0601
N-acetylmuramoyl-L-alanine amidase family protein	Psyr_0569
penicillin-binding protein	
staphylococcal nuclease-like protein	Psyr_0400
methyl-accepting chemotaxis protein	
hypothetical protein	Psyr_0363
acyl-CoA dehydrogenase family protein	
acyl-CoA dehydrogenase family protein	
hypothetical protein	Psyr_0348
AcrB/AcrD/AcrF family protein	Psyr_0346
RND family efflux transporter MFP subunit	Psyr_0345
radical SAM domain-containing protein	
histidine transporter, periplasmic histidine-binding protein	Psyr_4829
histidine ABC transporter, permease protein	Psyr_4830
histidine ABC transporter, ATP-binding protein	Psyr_4831
methyl-accepting chemotaxis protein	Psyr_4907
class III aminotransferase	Psyr_4934
hypothetical protein	
hypothetical protein	Psyr_5084
hypothetical protein	
acyltransferase family protein	Psyr_5087
sensor histidine kinase	Psyr_5089
ABC transporter permease	Psyr_5091
hypothetical protein	
methyl-accepting chemotaxis protein	Psyr_5092
methyl-accepting chemotaxis protein	Psyr_5093
hypothetical protein	
phosphate acetyltransferase	
ribose ABC transporter ATP-binding protein	Psyr_2152
glucose-1-phosphate thymidyltransferase	
dTDP-glucose 4,6-dehydratase	
riboflavin biosynthesis protein RibD	Psyr_4462
RNA polymerase sigma-70 factor	Psyr_4641
superoxide dismutase	Psyr_4152
sarcosine oxidase, alpha subunit	Psyr_4715
signal peptide peptidase SppA, 36K type	Psyr_1641
tellurium resistance protein TerA	Psyr_0809
thymidylate synthase	Psyr_4840

anthranilate synthase component II	Psyr_4581
GTP-binding protein TypA	Psyr_4815
aromatic-amino-acid aminotransferase	Psyr_4898
xylose isomerase	Psyr_2883
D-xylose ABC transporter, periplasmic-D xylose binding protein	Psyr_2884
zinc ABC transporter ATP-binding protein	

Table S5B- List of Psy B728a genes recombining within and between ph
within populations from the same isolation source represented genes fo
2.5% interval . Stars associated to recombining genes denote those in th
sources.

Psy B728a genes	Max Intensity of recombination H _i			
	Whole population	Environmental populations	Crop pathogens populations	
Psyr_0160	-	-	2.24	
*	Psyr_0195	4.02	3.48	3.15
	Psyr_0197	2.89	2.45	2.29
	Psyr_0336	-	-	2.17
	Psyr_0346	2.38	-	2.31
	Psyr_0390	-	2.55	-
*	Psyr_0392	13.87	12.60	8.68
*	Psyr_0393	8.78	9.64	4.30
	Psyr_0394	5.08	4.97	2.18
	Psyr_0506	2.16	-	-
	Psyr_0862	-	-	2.26
	Psyr_0868	2.34	-	2.91
	Psyr_0872	2.27	-	-
	Psyr_0873	2.26	-	-
	Psyr_0874	2.03	-	-
	Psyr_0875	1.97	-	-
	Psyr_0885	-	-	2.17
	Psyr_0896	-	-	2.16
*	Psyr_0903	3.11	2.92	3.39
	Psyr_0913	-	2.30	-
	Psyr_0952	1.96	-	2.93
	Psyr_0969	-	-	2.42
	Psyr_0975	2.22	-	2.36
	Psyr_0976	2.27	2.38	2.27
	Psyr_0984	-	-	2.18
	Psyr_1186	-	2.70	-
	Psyr_1213	-	2.38	-
*	Psyr_1214	2.75	3.11	3.05
	Psyr_1215	2.65	3.24	2.67
	Psyr_1218	-	2.52	-
	Psyr_1257	1.96	-	-
	Psyr_1277	-	-	2.22

	Psyr_1278	-	2.07	2.50
	Psyr_1376	2.12	-	-
	Psyr_1391	2.10	2.15	2.19
*	Psyr_1392	3.22	3.10	4.01
	Psyr_1393	2.19	2.11	2.74
	Psyr_1413	-	2.48	-
	Psyr_1691	2.28	-	2.34
	Psyr_1704	2.09	2.78	2.30
	Psyr_1720	-	2.24	-
	Psyr_1721	-	2.23	-
	Psyr_1743	-	-	2.62
	Psyr_1781	-	-	2.42
	Psyr_1794	2.19	-	2.46
	Psyr_1797	1.92	3.18	2.04
	Psyr_1798	2.01	2.36	2.25
	Psyr_1799	-	-	2.20
	Psyr_1800	-	-	2.26
	Psyr_1801	-	-	2.41
	Psyr_1818	1.96	-	-
	Psyr_1819	2.50	2.21	2.34
	Psyr_1936	4.73	2.22	4.69
	Psyr_1957	2.00	2.07	2.56
	Psyr_1971	-	2.37	-
	Psyr_1977	-	2.34	-
	Psyr_1981	-	2.29	-
	Psyr_1982	-	2.27	2.04
	Psyr_1991	1.94	2.18	2.28
	Psyr_1992	2.07	-	2.36
	Psyr_1995	-	-	2.57
	Psyr_2152	1.95	2.41	-
	Psyr_2153	2.05	2.48	2.33
	Psyr_2241	-	2.19	-
	Psyr_2249	2.00	-	-
	Psyr_2250	2.17	-	-
	Psyr_2251	1.95	-	-
	Psyr_2253	-	-	2.24
	Psyr_2260	-	-	2.27
	Psyr_2274	2.21	2.14	2.56
	Psyr_2276	-	2.92	2.23
	Psyr_2282	-	2.99	-

Psyr_2298	2.26	2.64	2.28
Psyr_2299	2.03	-	2.21
Psyr_2338	2.18	-	2.81
Psyr_2426	-	-	2.17
Psyr_2427	-	-	2.65
Psyr_2433	-	-	2.20
Psyr_2445	-	2.37	-
Psyr_2574	2.03	2.53	2.12
Psyr_2743	1.94	-	3.05
Psyr_2744	-	3.26	2.36
Psyr_2746	2.58	-	4.05
Psyr_2747	2.95	-	3.93
Psyr_2748	-	-	2.93
Psyr_2749	-	-	2.42
Psyr_2751	-	-	2.31
Psyr_2752	2.00	-	2.68
Psyr_2754	1.98	-	2.42
Psyr_2755	2.53	-	3.21
Psyr_2756	2.24	2.18	2.72
Psyr_2757	-	-	2.40
*	Psyr_2803	5.51	5.60
	Psyr_2865	-	-
*	Psyr_2902	3.23	3.47
*	Psyr_2903	5.88	4.95
*	Psyr_2904	6.13	5.48
*	Psyr_2905	4.00	4.04
	Psyr_2932	2.06	2.50
	Psyr_2937	1.98	2.58
*	Psyr_2938	2.87	3.24
*	Psyr_2939	2.96	3.22
	Psyr_2952	-	2.19
	Psyr_2953	2.14	2.29
	Psyr_2954	2.51	2.31
	Psyr_2955	2.40	2.33
	Psyr_2956	3.21	2.37
	Psyr_2957	2.25	-
	Psyr_2958	2.10	-
	Psyr_2977	2.34	-
	Psyr_2979	2.26	3.55
	Psyr_2989	-	2.31
	Psyr_3020	1.95	-
			2.61

Psyr_3021	2.18	-	2.65
Psyr_3037	-	3.36	-
Psyr_3102	-	-	2.50
Psyr_3119	2.14	-	3.21
Psyr_3130	2.22	-	-
Psyr_3131	2.91	-	2.63
Psyr_3132	2.24	-	2.05
Psyr_3134	2.01	-	2.63
Psyr_3151	-	-	2.17
Psyr_3156	1.96	-	-
Psyr_3157	2.17	2.35	2.40
Psyr_3158	2.99	2.61	2.38
* Psyr_3159	3.70	2.81	3.50
* Psyr_3160	3.77	2.79	3.53
Psyr_3161	3.00	2.23	3.81
Psyr_3167	-	2.58	-
Psyr_3399	-	2.52	-
Psyr_3400	2.38	2.94	-
Psyr_3591	-	-	2.16
Psyr_3593	-	-	2.19
Psyr_3600	1.94	-	2.36
Psyr_3601	-	-	2.23
Psyr_4314	2.01	-	-
Psyr_4365	-	-	2.42
Psyr_4387	-	2.33	-
Psyr_4632	-	-	2.66
Psyr_4640	2.00	-	-
Psyr_4660	1.96	-	2.78
Psyr_4665	2.31	-	2.23
Psyr_4666	2.23	-	2.50
Psyr_4667	1.96	-	2.27
Psyr_4674	2.15	-	2.82
Psyr_4675	2.11	-	2.37
Psyr_4677	-	-	2.45
Psyr_4678	-	-	2.91
Psyr_4681	-	-	2.27
Psyr_4685	-	-	2.22
Psyr_4686	1.99	-	2.42
Psyr_4730	2.37	-	-
Psyr_4769	1.98	-	-
Psyr_4775	2.58	2.98	2.47
Psyr_4782	-	2.72	-
Psyr_4812	-	-	2.59
Psyr_4835	1.93	2.92	-

Psyr_4836	-	2.74	-
Psyr_4837	-	2.57	-
Psyr_5093	-	-	2.63
Psyr_5107	-	-	2.27
			2.725268558

Yologroup 2d populations from the two isolation sources. Genes recombining
r which recombination hotness H_i value for at least one SNP was in the top
e top 1% interval that are shared by populations from the two isolation

Description	Ortholog in <i>Pto</i> DC3000
N-acetyltransferase GCN5	PSPTO_0245
hypothetical protein	PSPTO_2534
Mg chelatase-like protein	PSPTO_0042
Outer membrane autotransporter barrel	PSPTO_5200
acriflavin resistance protein	PSPTO_5191
hypothetical protein	PSPTO_5149
hypothetical protein	PSPTO_5149
hypothetical protein	PSPTO_5142
ATP-dependent protease ATP-binding subunit HslU	PSPTO_5141
trifunctional transcriptional regulator/proline	PSPTO_5017
dehydrogenase/pyrroline-5-carboxylate dehydrogenase	
acyltransferase 3	
histidine kinase, HAMP region: chemotaxis sensory transducer	
alkaline phosphatase	PSPTO_1010
N-acetyltransferase GCN5	
biopolymer transport protein ExbD/TolR	PSPTO_1012
hypothetical protein	PSPTO_1013
hypothetical protein	PSPTO_1038
aldose 1-epimerase	PSPTO_1047
outer membrane porin	PSPTO_1054
histidine kinase, HAMP region: chemotaxis sensory transducer	PSPTO_1066
glutamate racemase	PSPTO_1112
uracil phosphoribosyltransferase	PSPTO_1130
homocysteine S-methyltransferase	PSPTO_1135
malate:quinone oxidoreductase	PSPTO_1136
peptidase M48, Ste24p	PSPTO_1144
type III effector HopM1	
type III secretion cytoplasmic ATPase HrcN	PSPTO_1400
type III secretion protein HrpQ	
Type III secretion protein HrcV	PSPTO_1402
type III helper protein HrpK1	PSPTO_1405
2-isopropylmalate synthase	PSPTO_1444
phosphoribosylglycinamide formyltransferase 2	PSPTO_1468

citrate-proton symport	PSPTO_1469
DNA mismatch repair protein MutS	PSPTO_4058
N-acetyltransferase GCN5	
ribosomal protein S12 methylthiotransferase	PSPTO_4019
hypothetical protein	PSPTO_3994
biopolymer transport protein ExbD/TolR	PSPTO_3974
hypothetical protein	
amino acid adenylation	
hypothetical protein	PSPTO_3760
moxR protein	PSPTO_3759
bifunctional 5,10-methylene-tetrahydrofolate dehydrogenase/ 5,10-methylene-tetrahydrofolate cyclohydrolase	PSPTO_3733
Beta-glucosidase	PSPTO_3694
amino acid adenylation	
hypothetical protein	PSPTO_3682
histidine kinase, HAMP region: chemotaxis sensory transducer	PSPTO_3680
hypothetical protein	PSPTO_3679
regulatory protein LysR	PSPTO_3678
1-aminocyclopropane-1-carboxylate deaminase	PSPTO_3675
cell division protein ZipA	PSPTO_3657
NAD-dependent DNA ligase LigA	PSPTO_3656
tRNA-dihydrouridine synthase A	PSPTO_2126
amino acid adenylation	PSPTO_2147
peptidase S45, penicillin amidase	PSPTO_2161
glutamyl-tRNA synthetase	PSPTO_2166
PAS:GGDEF	PSPTO_2171
regulatory protein LysR	PSPTO_2172
regulatory proteins, IclR	PSPTO_2181
zinc-containing alcohol dehydrogenase superfamily protein	PSPTO_2182
electron transfer flavoprotein subunit beta	PSPTO_2185
hypothetical protein	PSPTO_2368
FecR protein	PSPTO_2369
threonyl-tRNA synthetase	PSPTO_2475
binding-protein dependent transport system inner membrane protein	PSPTO_2557
transcriptional regulator GntR	PSPTO_2558
phosphonate metabolism PhnG	PSPTO_2559
phosphonate metabolism protein	PSPTO_2561
amidase	PSPTO_2569
glutamine amidotransferase, class-II	PSPTO_2583
ferredoxin-dependent glutamate synthase	PSPTO_2585
hydrophobe/amphiphile efflux-1 HAE1	PSPTO_2592

glycoside hydrolase family protein	PSPTO_2493
hypothetical protein	
hypothetical protein	PSPTO_2528
hypothetical protein	PSPTO_2693
FAD dependent oxidoreductase	PSPTO_2694
hypothetical protein	PSPTO_2700
response regulator receiver	PSPTO_2712
quinoprotein	
lysine exporter protein LysE/YggA	PSPTO_2958
hypothetical protein	
L-lysine 6-monooxygenase	PSPTO_2961
extracellular ligand-binding receptor	PSPTO_2962
FAD dependent oxidoreductase	PSPTO_2963
methionyl-tRNA synthetase	PSPTO_2964
Phage integrase:Phage integrase, N-terminal SAM-like	PSPTO_2966
hypothetical protein	PSPTO_2967
regulatory protein, TetR	PSPTO_2969
2-nitropropane dioxygenase	PSPTO_2972
molybdenum ABC transporter periplasmic-binding protein	PSPTO_2973
binding-protein dependent transport system inner membrane protein	PSPTO_2974
hypothetical protein	
hydrophobe/amphiphile efflux-1 HAE1	
extracellular solute-binding protein	PSPTO_3029
binding-protein dependent transport system inner membrane protein	PSPTO_3030
binding-protein dependent transport system inner membrane protein	PSPTO_3031
ABC transporter	PSPTO_3032
hypothetical protein	PSPTO_4279
NADH:flavin oxidoreductase	PSPTO_3080
regulatory protein LysR	PSPTO_3081
diguanylate cyclase	
NADP oxidoreductase, coenzyme F420-dependent	
catalytic LigB subunit of aromatic ring-opening dioxygenase	
phospholipase/carboxylesterase	
Surfeit locus 4-related	
zinc-containing alcohol dehydrogenase superfamily protein	
glutathione-dependent formaldehyde-activating protein	
extracellular solute-binding protein	PSPTO_3088
cointegrate resolution protein T	
glutathione reductase	PSPTO_3112
6-phosphogluconate dehydrogenase	PSPTO_3122
precorrin-4 C(11)-methyltransferase	PSPTO_3154

indolepyruvate ferredoxin oxidoreductase	PSPTO_3155
hypothetical protein	
regulatory protein LysR	PSPTO_3265
major facilitator transporter	PSPTO_3284
secretion protein HlyD	PSPTO_3300
secretion protein HlyD	
acriflavin resistance protein	PSPTO_3302
major facilitator transporter	
type II secretion system protein E	PSPTO_3317
regulatory protein LysR	PSPTO_3322
aldehyde dehydrogenase	PSPTO_3323
N-methyltryptophan oxidase	PSPTO_3327
Type I secretion outer membrane protein, TolC	PSPTO_3328
Type I secretion membrane fusion protein, HlyD	PSPTO_3329
Type I secretion system ATPase, PrtD	PSPTO_3330
citrate-proton symport	PSPTO_3336
endoribonuclease L-PSP	PSPTO_3640
FAD dependent oxidoreductase	PSPTO_3642
sensor histidine kinase	PSPTO_1803
K ⁺ transporter Trk	PSPTO_1801
regulatory protein LysR	PSPTO_1794
hypothetical protein	PSPTO_1793
beta-ketoacyl synthase	
peptidoglycan glycosyltransferase	PSPTO_4825
penicillin-binding protein 1C	PSPTO_4847
hypothetical protein	PSPTO_0546
DNA primase	PSPTO_0538
formate/nitrite transporter	
hypothetical protein	PSPTO_0520
helix-turn-helix, Fis-type	PSPTO_0519
tryptophan synthase subunit alpha	PSPTO_0518
pyrroloquinoline quinone biosynthesis protein PqqE	PSPTO_0509
peptidase S9, prolyl oligopeptidase active site region	PSPTO_0508
acyl-CoA dehydrogenase	PSPTO_0506
GGDEF	PSPTO_0505
acyl-CoA dehydrogenase	PSPTO_0500
bioH protein	PSPTO_0496
8-amino-7-oxononanoate synthase	PSPTO_0495
FecR protein	PSPTO_0445
major facilitator transporter	PSPTO_0407
ferredoxin:oxidoreductase FAD/NAD(P)-binding subunit	PSPTO_0402
NADH:flavin oxidoreductase	PSPTO_0396
extracellular solute-binding protein	PSPTO_0364
imidazolonepropionate	PSPTO_5277

N-formylglutamate deformylase	PSPTO_5278
heavy metal-(Cd/Co/Hg/Pb/Zn)-translocating P-type ATPase	PSPTO_5279
histidine kinase, HAMP region: chemotaxis sensory transducer	PSPTO_5554
glycosyl transferase family protein	PSPTO_5587

Table S5C- List of *Pto* DC3000 / *Psy* B728a genes associated to hot regions for recombination in both strains

<i>Pto</i> DC3000 genes	Annotation in <i>Pto</i> DC3000
Pto_PSPTO_5191	AcrB/AcrD/AcrF family protein
Pto_PSPTO_1047	hypothetical protein
Pto_PSPTO_1054	outer membrane porin, OprD family
Pto_murl	glutamate racemase
Pto_rbsA_1	ribose ABC transporter ATP-binding protein
Pto_PSPTO_2569	amidase family protein
Pto_PSPTO_2583	glutamine amidotransferase, class-II protein
Pto_PSPTO_2962	hypothetical protein
Pto_PSPTO_3029	peptide ABC transporter periplasmic peptide-binding protein
Pto_PSPTO_3030	peptide ABC transporter permease
Pto_PSPTO_3031	peptide ABC transporter permease
Pto_PSPTO_3032	peptide ABC transporter ATP-binding protein
Pto_PSPTO_3284	major facilitator family transporter
Pto_PSPTO_3302	AcrB/AcrD/AcrF family protein
Pto_PSPTO_3336	citrate transporter
Pto_PSPTO_0364	sugar ABC transporter periplasmic sugar-binding protein
Pto_cadA_1	cadmium-translocating P-type ATPase
Pto_PSPTO_5554	methyl-accepting chemotaxis protein

Ortholog phylogroups.

Ortholog genes in Annotation in *Psy* B728a

Psyr_0346	acriflavin resistance protein
Psyr_0896	aldose 1-epimerase
Psyr_0903	outer membrane porin
Psyr_0952	glutamate racemase
Psyr_2152	hypothetical protein
Psyr_2260	amidase
Psyr_2274	glutamine amidotransferase, class-II
Psyr_2747	extracellular ligand-binding receptor
Psyr_2902	extracellular solute-binding protein
Psyr_2903	binding-protein dependent transport system inner membrane protein
Psyr_2904	binding-protein dependent transport system inner membrane protein
Psyr_2905	ABC transporter
Psyr_3119	major facilitator transporter
Psyr_3132	acriflavin resistance protein
Psyr_3167	citrate-proton symport
Psyr_4812	extracellular solute-binding protein
Psyr_4837	heavy metal-(Cd/Co/Hg/Pb/Zn)-translocating P-type ATPase
Psyr_5093	histidine kinase, HAMP region: chemotaxis sensory transducer
