1 **Genomic and gene-expression comparisons among phage-resistant type-IV pilus mutants of**

2 ***Pseudomonas syringae* pathovar *phaseolicola***

3 Mark Sistrom[*], Derek Park[§], Heath E. O'Brien[†], Zheng Wang[§], David S. Guttman[†,‡], Jeffrey P.

4 Townsend[§,**,§§,††], and Paul E. Turner[§,††]

5 * School of Natural Sciences, University of California Merced, Merced, 95343, CA, USA

6 [§] Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520,

7 USA

8 † Department of Cell and Systems Biology, University of Toronto, Toronto, Ontario, M5S 3B2,

9 Canada

10 ‡ Centre for the Analysis of Genome Evolution & Function, University of Toronto, Ontario,

11 M5S 3B2, Canada

12 ** Department of Biostatistics, Yale School of Public Health, New Haven, CT 06520, USA

13 [§§] Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT

14 06520, USA

15 †† Program in Microbiology, Yale University, New Haven, CT 06520, USA

16

17

18 Current addresses:

19   1. Derek Park, Mathematical Institute, University of Oxford, Oxford, OX2 6GG, United

20   Kingdom

21   2. Heath O'Brien, School of Biological Sciences, University of Bristol, Bristol, BS8 1TQ,

22   United Kingdom

23

24 **Running title:** *Genomics of type IV pilus P. phaseolicola mutants*

25

26 # Corresponding author's contact details

27 School of Natural Sciences, University of California Merced, Merced, 95343, CA, USA

28 msistrom@ucmerced.edu

29

30 Keywords: genomics, gene expression, *Pseudomonas syringe*, bacteriophages, φ6 bacteriophage

31

3

32   **Abstract**

33   *Pseudomonas syringae* pv. *phaseolicola* (Pph) is a significant bacterial pathogen of

34   agricultural crops, and phage φ6 and other members of the dsRNA virus family *Cystoviridae*

35   undergo lytic (virulent) infection of Pph, using the type IV pilus as the initial site of cellular

36   attachment. Despite the popularity of Pph/phage φ6 as a model system in evolutionary biology,

37   Pph resistance to phage φ6 remains poorly characterized. To investigate differences between

38   phage φ6 resistant Pph strains, we examined genomic and gene expression variation among

39   three bacterial genotypes that differ in the number of type IV pili expressed per cell: ordinary

40   (wild-type), non-piliated, and super-piliated. Genome sequencing of non-piliated and super-

41   piliated Pph identified few mutations that separate these genotypes from wild type Pph – and

42   none present in genes known to be directly involved in type IV pilus expression. Expression

43   analysis revealed that 81.1% of GO terms up-regulated in the non-piliated strain were down-

44   regulated in the super-piliated strain. This differential expression is particularly prevalent in

45   genes associated with respiration — specifically genes in the tricarboxylic acid cycle (TCA) cycle,

46   aerobic respiration, and acetyl-CoA metabolism. The expression patterns of the TCA pathway

47   appear to be generally up and down-regulated, in non-piliated and super-piliated Pph

48   respectively. As pilus retraction is mediated by an ATP motor, loss of retraction ability might

49   lead to a lower energy draw on the bacterial cell, leading to a different energy balance than

50   wild type. The lower metabolic rate of the super-piliated strain is potentially a result of its loss

51   of ability to retract.

52

53 **Introduction**

54   Halo blight is an economically significant disease of leguminous agricultural crops

55 caused by the Gram-negative bacterial pathogen *Pseudomonas syringae* pv. *phaseolicola* (Pph)

56 (Romantschuk and Bamford 1986). The genome of Pph comprises a 5.93 Mb chromosome and

57 two plasmids, 132 kb and 52 kb in size (Joardar *et al.* 2005). Virulence in Pph requires a

58 pathogenicity island located on the larger plasmid (Jackson *et al.* 1999). However, the ability for

59 Pph and other Pseudomonads to attach to plants is via chromosomally-encoded type IV protein

60 pili, which allow the bacterial cells to adhere to external structures (e.g., leaf surfaces) and to

61 resist environmental perturbations such as rain splatter and wind, which could disrupt bacterial

62 invasion of plant tissues (Romantschuk *et al.* 1993; Roine *et al.* 1996; Burdman *et al.* 2011). One

63 of the key mechanisms enabling motility of plant-pathogenic Pseudomonads — fundamental to

64 their interactions with host plants — occurs through twitching motility (see review by Mattick

65 2002), in which type IV pili are repeatedly extended and retracted to facilitate movement,

66 analogous to grappling hooks. Spontaneous mutants of Pph that lack type IV pili show greatly

67 reduced adherence to leaf surfaces and lower incidence of halo blight disease, whereas

68 virulence of wildtype and pilus-deficient Pph mutants is equivalent when these strains are

69 injected directly into bean leaves (Romantschuk and Bamford 1986). Therefore, adherence

70 ability via type IV pili can be considered a conditional virulence factor directly involved in

71 epiphytic colonization by *P. syringae* pathovars such as Pph.

72   Because type IV pili in Pseudomonads are critical for host plant attachment, it is

73 unsurprising that some bacteriophages have evolved to exploit these conserved structures to

74 facilitate viral infection of bacterial cells (Romantschuk and Bamford 1985). In particular, phage

75   φ6 and other members of the dsRNA virus family *Cystoviridae* undergo virulent (lytic) infection

76   of Pph, using the type IV pilus as the initial site of cellular attachment (Roine et al. 1996; see

77   review by Mindich 2006). Cystoviruses are characterized as having three dsRNA strands

78   (segments) per particle and a lipid envelope. Attachment of phage φ6 to the type IV pili of Pph

79   is mediated by the viral attachment protein P3 (Gottlieb *et al.* 1988; Duffy *et al.* 2006). Once

80   attached, it is believed that retraction of the pilus moves one or more φ6 particles in close

81   proximity to the bacterial cell membrane, allowing membrane fusion (Bamford *et al.* 1987) and

82   ultimately phage entry into the cell via the viral lytic enzyme P5 (Gottlieb *et al.* 1988; Dessau *et*

83   *al.* 2012). Phage φ6 then uses the host cell's metabolism to cause a typical lytic infection, where

84   cell lysis liberates mature virions capable of infecting host cells expressing the type IV pilus

85   (Bamford et al. 1976) (see review by Mindich 2006). In addition to the popularity of Pph as a

86   model for studying generalized plant-pathogen interactions (Arnold *et al.* 2011), the infection

87   dynamic between Pph and phage φ6 can also be readily examined *in vitro,* facilitating studies of

88   the evolution of host-parasite interactions (Turner and Chao 1999; Lythgoe and Chao 2003;

89   Dennehy, Abedon, *et al.* 2007).

90   Despite the growing popularity of Pph and phage φ6 as a study system in evolutionary

91   biology (Dennehy 2009), the evolution of Pph resistance to phage φ6 remains poorly

92   characterized. Laboratory culture of Pph in the presence of phage φ6 can be used to select for

93   spontaneous bacterial mutants that are partially or fully resistant to phage infection

94   (Romantschuk and Bamford 1985; Lythgoe and Chao 2003). We are unaware of any study that

95   has characterized the full spectrum of mechanisms that allow Pph resistance to phage φ6.

96   However, prior *in vitro* experiments show that evolved resistance generally coincides with

97   changes in the number of type IV pili produced by the bacteria (Romantschuk and Bamford

98   1985; Lythgoe and Chao 2003), suggesting that type IV pili alteration may be the primary

99   resistance mechanism. Notably, two of these resistant forms differ markedly in pili number:

100   strains lacking type IV pili altogether, and strains with super-abundant type IV pili

101   (Romantschuk and Bamford 1985; Lythgoe and Chao 2003). Resistance in non-piliated strains is

102   due to the elimination of the initial attachment site for phage φ6, although some low-level

103   infection of these hosts occurs, presumably due to virus particles colliding directly with the cell

104   surface (Dennehy, Abedon, *et al.* 2007). Super-piliated host strains have elevated numbers of

105   pili compared to wild type Pph, which suggests that these bacteria should be even more

106   sensitive to phage φ6 infection. Instead, super-piliated bacteria adsorb relatively larger

107   numbers of phage particles as expected, but remain uninfected apparently because they do not

108   retract their pili (Romantschuk and Bamford 1985; Lythgoe and Chao 2003); this inability to

109   retract type IV pili appears similar to the mechanism described in *P. aeruginosa* (Johnson and

110   Lory 1987). Thus, super-piliated Pph strains 'soak up' virus particles in culture, but restrict

111   phage infection because the non-retractable pili prevent virus particles from encountering the

112   cell surface (Lythgoe and Chao 2003; Dennehy, Abedon, *et al.* 2007). Wild-type, non-piliated

113   and super-piliated Pph strains show similar maximal growth rates, stationary-phase densities

114   and relative fitness abilities (i.e., competition for limited nutrients) in liquid culture in the

115   laboratory (Dennehy, Friedenberg, *et al.* 2007). More importantly, super-piliated mutants of

116   Pph are similar to wild-type bacteria in terms of leaf adherence, and initiation of halo blight

117   disease (Romantschuk and Bamford 1986) This observation suggests that super-piliated phage-

118   resistant mutants should be selectively favored over non-piliated (plant-adherence deficient)

119   resistant mutants in the wild; however, we are unaware of any field studies that have directly

120   compared the success of super-piliated versus non-piliated resistance mutants in natural

121   populations of Pph, when exposed to phage attack.

122   To investigate differences between phage φ6 resistant Pph strains and wild-type

123   bacteria, we compared gene-expression profiles and whole-genome sequences of the two

124   major types of resistant strains and wild-type bacteria. In particular, we examined genomic and

125   gene expression variation between ordinary (wild-type) Pph and two spontaneous mutants of

126   Pph that differed in the number of type IV pili expressed per cell: non-piliated, and super-

127   piliated.

128   **Materials and Methods**

129   *Strains*

130   Wild-type (WT) Pph was obtained from American Type Culture Collection (ATCC strain

131   #21781, a.k.a. strain HB10Y). Non-piliated (P–) Pph strain #LM2509 is a spontaneous mutant of

132   WT Pph, kindly provided by L. Mindich (Public Health Research Institute, Newark, NJ, USA).

133   Super-piliated (P+) Pph strain #KL6 is also a spontaneous mutant of WT Pph, kindly provided by

134   L. Chao (University of California, San Diego, La Jolla, CA).

135   *Culture conditions*

136   All strains were grown, diluted and plated at 25°C in Luria broth (LB; 10 g NaCl, 10 g

137   Bacto® tryptone [Becton, Dickinson and Co., Sparks, MD], and 5 g Bacto® yeast extract per L), as

138   described in Dennehy et al. (Dennehy, Abedon, *et al.* 2007). Cultures were initiated by placing a

139   single colony grown on LB agar into a flask with 10 mL of LB medium. Flasks were shaken at 120

140   rpm with incubation for 24 hrs, which was sufficient for bacteria to achieve stationary-phase

141   density (~1.65 × 10^9 cells/ml).

142   **_Genome sequencing and analysis_**

143   Sequencing followed the methods described in O'Brien et al. (O'Brien _et al._ 2011).

144   Briefly, DNA was isolated from 1 mL of stationary-phase culture using a Puregene Genomic DNA

145   Purification Kit (Qiagen Canada, Toronto, ON) with the Gram-negative bacterial culture protocol

146   with double volumes of each reagent, repeating the protein precipitation step twice, and

147   spooling the DNA during the precipitation step. DNA was sheared to 200 base pairs (bp) using a

148   Covaris S-series sample preparation system and paired-end sequencing libraries were prepared

149   using sample preparation kits from Illumina (San Diego, CA). Libraries were multiplexed and run

150   on a single lane of an Illumina GA IIx sequencer for 80 cycles per paired-end. Reads were

151   mapped to the genome sequence of wild-type Pph strain HB10Y (Guttman unpublished data)

152   using the CLC Genomics Workbench (Århus, Denmark), and polymorphisms were detected with

153   the CLC Probabilistic Variant Caller. Annotations were copied from the fully sequenced

154   reference strain Pph 1448A (Joardar _et al._ 2005) using a Mauve whole genome alignment

155   (Darling _et al._ 2004).

156   We evaluated mutational changes in four genes known to be directly involved in type IV

157   pilus formation – type IV pilus assembly protein (PilM), type IV pilus biogenesis protein (PilN),

158   type IV pilus biogenesis protein (PilO) and type IV pilus biogenesis protein (PilP) (Roine _et al._

159   1996, 1998; Taguchi and Ichinose 2011; Nguyen _et al._ 2012) in P+ and P- compared to wild type

160   in order to asses any direct genotype – phenotype explanations for the observed phenotypic

161   characteristics of these two strains.

162   For a well-characterized *P. syringae* pathovar distantly related to Pph (*P. syringae* pv.

163   *tomato* strain #DC3000), putative sigma70 binding sites inferred from the RegulonDB database

164   (Huerta *et al.* 1998) were downloaded from

165   http://www.ccg.unam.mx/Computational_Genomics/PromoterTools and homologous positions

166   in Pph HB10Y were determined from a whole genome alignment made with Mauve (Darling *et*

167   *al.* 2004). Mutations occurring within 50bp upstream of the inferred transcription start sites

168   were identified using BEDTools (Quinlan and Hall 2010). Pph HB10Y homologs were identified

169   for 5121 out of 6716 putative sigma70 binding sites in *P. tomato* DC3000, and mutations were

170   present within 50 bp of the putative transcription start sites for 8 (7 in P+ strain KL6, 1 in P–

171   strain LM2509).

172   **Gene expression analysis**

173   Two biological replicates of each strain were cultured. For RNA extraction we used a

174   standard Trizol approach. The bacterial cells (roughly 10^7 to 10^8 cells per ml of Trizol) were lysed

175   in Trizol (Life Technology) and homogenized with a micropestle (Eppendorf), and RNAs were

176   suspended in DEPC treated water. RNA quality was checked via spectrophotometry

177   (NanoDrop). Reverse transcription of the RNA into cDNA was performed using the Invitrogen

178   SuperScript Double-Stranded cDNA Synthesis kit (Thermo Fisher) using random primers. For

179   each biological replicate, a technical replicate of the reverse transcription was performed using

180   multi-targeted primers (Adomas *et al.* 2010) designed for the Pph genome, instead of random

181   primers, for first strand synthesis. All cDNA samples were labeled using a NimbleGen One-Color

182   (Cy3) DNA Labeling Kit (Madison, WI, USA). Hybridization of cDNA to NimbleGen 4 x 72k arrays

183   (090326_P_syringe_B728a_expr; NimbleGen design ID #9510; Madison, WI, USA) according to

184   the manufacturer guidelines was conducted by the Yale Center for Genome Analysis (New

185   Haven, Connecticut, USA). Microarray slides were scanned with a GenePix 4000B (Axon

186   Instruments, Foster City, CA). Spots were located and expression was quantified following

187   NimbleGen Arrays User's Guide, and spots with unusual morphology or with erratic signal

188   intensity distribution were excluded from all analyses. Raw fluorescence data were normalized

189   as in Townsend (Townsend 2004), and normalized data were then statistically analyzed using

190   Bayesian Analysis of Gene Expression Levels (BAGEL, (Townsend and Hartl 2002; Townsend

191   2004)). Genes were considered expressed and deemed well measured only when the median

192   spot foreground exceeded the median background plus two standard deviations of the

193   background intensity. Well measured genes were considered as significantly differentially

194   expressed when P ≤ 0.05. Significantly up and down regulated genes with Gene Ontology (GO)

195   terms were used to determine significantly enriched GO terms using AmiGO 1.8 (Carbon *et al.*

196   2009) with a significance level of 0.01 and a minimum of two genes for each enriched term. The

197   JCVI CMR (Davidsen *et al.* 2009) database was used for background filtering and results were

198   visualized using the AmiGO GO graph tool (Carbon *et al.* 2009). Visualization of gene expression

199   differences among strains was done by using the clustergram function of the bioinformatics

200   toolbox of the MatLab computational mathematics suite (Mathworks, Natick, MA, USA). The

201   data were first mean-normalized and then transformed so that gene expression values were

202   represented as standard deviations above or below the mean. Data were not sorted for

203   statistical significance beforehand although later conclusions did take specific *P* values between

204   strains into account.

205   **Results**

206   ***Strain isolation***

207   The super-piliated and non-piliated mutants in the current study are spontaneous

208   mutants of WT Pph strain HB10Y, obtained via laboratory assays that isolated spontaneous

209   resistance mutants in bacterial populations subjected to phage φ6 attack. This design facilitated

210   genomic and gene-expression comparisons among strains, which may have been complicated if

211   the strains were more distantly related, such as drawn from disparate natural sources and/or

212   geographic locations. We noted that all three strains achieved similar stationary-phase

213   densities (~1.65×10^9 cells/ml) and maximal growth rates in Luria broth (Dennehy et al. 2007),

214   indicating that average number of type IV pili expressed per cell constituted the major

215   differences between WT Pph and each of the two spontaneous mutants.

216   ***Genomics of super-piliated and non-piliated bacterial strains***

217   Illumina-based genome sequencing using 80 base pair (bp) paired-end reads of the

218   super-piliated (P+) and non-piliated (P−) strains produced 1.9 and 3.6 million reads respectively.

219   Subsequent referenced alignment to the wildtype (WT) *P. syringae* pv. *phaseolicola* strain

220   HB10Y resulted in a mean coverage of 14.0× and 24.4× for P+ and P−.

221   Alignment of the P+ strain identified 97 mutations separating it from WT (Table 1, Table

222   S1), consisting of 26 large-scale deletions, five deletion-insertion polymorphisms (DIPs), seven

223   multi-nucleotide variants (MNVs) and 59 single nucleotide polymorphisms (SNPs). A notably

224   large deletion of 1039 bp occurred in a retrotransposon hot spot (*rhs*) family gene, with the

225   remainder averaging 120bp (54–269 bp). Fourteen of the deletions occurred in inter-genic

226   regions, six in characterized genes and six in hypothetical proteins of unknown function. All five

227   DIPs in P+ occurred in intergenic regions; three DIPs were in consecutive positions 158 bp

228   downstream of the gene encoding the type III effector HopQ1. The remaining two DIPs were

229   upstream of a prophage and fructose-1, 6-bisphosphate aldolase gene. Each observed MNV

230   consisted of a 2–3bp change, six of which occurred in intergenic regions and the remaining

231   MNV located in the *tolQ2* ORF, causing a threonine to leucine amino-acid substitution at the

232   121st position. Of the 59 SNPs detected in P+, 15 occurred in ORFs, of which six were non-

233   synonymous amino acid changes.

234   Alignment of the P– strain identified 82 total mutations separating it from WT (Table 1,

235   Table S2) consisting of four large-scale deletions, eight DIPs, eight MNVs, and 62 SNPs. A single

236   large deletion of 2116 bp occurred in the gene encoding the *rulA* protein, with the remainder

237   averaging 80 bp (53–112 bp). Four of the eight observed DIPs occurred in a GntR family

238   transcriptional regulator, causing frameshift mutations at amino acids 131 and 132. Two DIPs

239   occurred in a putative H+ antiporter subunit C causing a frameshift at Phe_{50}, and one DIP was

240   present in the *PilX* ORF causing a frameshift mutation at Ile_{192}. The remaining DIP was present

241   in an intergenic region. The eight observed MNVs ranged from 2–5bp in length. Five of the eight

242   MNVs were in intergenic regions. However, the remaining three MNVs were located within the

243   H+ antiporter, a GntR family transcriptional regulator and a TolQ protein causing frameshifts at

244   Phe_{50}, at Asp_{131} and an amino acid change from Thr_{121} to Leu, respectively. Of the 62 SNPs

245   observed in P–, 34 were located in intergenic regions, and of the 28 SNPs occurring in

246   confirmed ORFs, 16 represented non-synonymous changes.

247   None of the observed mutations were present in any of the genes known to be directly

248   involved in type IV pilus formation (i.e., PilM, PilN, PilO and PilP) (Roine *et al.* 1996, 1998;

249   Taguchi and Ichinose 2011; Nguyen *et al.* 2012) in P+ or P-.

250   Putative phage φ6 binding site mutations are shown in Table 2. One was detected in P−,

251   with seven detected in P+. The single binding site mutation in P− represents a SNP (G to C)

252   upstream of a citrate transporter gene, and downstream of a hypothetical protein. The binding

253   site mutations in P+ include an MNV upstream of type III effector *HopQ2* which was also

254   significantly upregulated in comparison with WT, along with six deletions ranging in size from

255   64–269 bp. Of the genes potentially affected by these deletions, a *LuxR* family D--binding

256   response regulator and an iron(III) dicitrate transport protein fecA were significantly

257   downregulated in P+, and a ribosomal subunit interface protein, a periplasmic amino acid-

258   binding protein, and a hypothetical protein were significantly upregulated in P+.

259   ***Transcriptome Analysis***

260   The gene expression (GE) profiles of the 3 strains were compared using a clustergram

261   and heatmap (Figure 1) analysis applied to all genes ($N = 5069$), and applied to genes that

262   showed a significant difference ($P < 0.05$) between strains ($N = 1009$). Both analyses showed

263   that the GE profile of P+ differed more substantially from WT than did P− (Figure 1) with higher

264   significance when restricted to significantly different genes. Two major clusters of genes were

265   observed: one cluster contained genes highly expressed in P+ compared to WT and P−, and

266   another contained genes highly expressed in WT and P− in comparison with P+.

267   Of the 5069 genes measured on the microarray, 954 genes (18.8%) showed significant

268   expression-level differences between WT and P+, with 503 genes upregulated and 451 genes

269   downregulated (Figure 2, Table 1, Table S3). To further examine genome-wide differences in

270   GE, genes upregulated in P+ were analyzed for GO term enrichment (Figure S1). We observed

271   that 10 terms were significantly enriched, all of which pertained to processes of transposition

272   and DNA metabolism. An enrichment analysis of genes down-regulated in P+ showed significant

273   enrichment for 48 different GO terms. Visualizing the organization of these GO terms revealed

274   that a few key biological processes were enriched: gene expression and translation, nucleotide

275   and ribonucleotide biosynthesis, and cellular respiration (Figure S2). Of the 5069 genes

276   measured on the microarray, 132 (2.6%) genes significantly differed in gene expression levels

277   between WT and P−. Of these 132 genes, 104 genes were up-regulated and 28 genes were

278   down-regulated (Table S1). We found that 49 GO terms were enriched for P− up-regulation and

279   31 GO terms were enriched for down-regulation. Visualization of these GO terms showed

280   particularly high enrichment for intracellular organelles and ribosomal structure in genes up-

281   regulated in P− (Figure S3), and genes associated with conjugation as well as protein transport

282   and secretion showed high levels of down-regulation in P− (Figure S4).

283   **Discussion**

284   ***Genomic differences between P+ and P− phage resistant* P. syringae *pv.* phaseolicola**

285   Genome sequencing of P+ and P− strains identified 97 and 82 mutations that differed

286   from WT strain, respectively. These mutations consisted of 26 deletions, five DIPs, seven MNVs

287   and 59 single SNPs in P+, and four deletions, eight DIPs, eight MNVs, and 62 SNPs in P− (Table

288   1). Of these differences, 20 and 27 mutations resulted in non-synonymous changes to protein

289   coding genes in P+ and P− strains. These observations constitute a relatively very small number

290   of protein coding changes — approximately one mutation per 300 kb and 220 kb respectively,

291   indicating that the observed differences in pilus number associated with P+ and P− stem from a

292   small number of large effect mutational changes. Genes affected by protein coding changes in

293   P+ included SNPs located in a general virulence surface antigen protein, and a type III effector

294   AvrB4-2—both loci are associated with plant-disease virulence in Pph (Arnold *et al.* 2011).

295   Genes affected by protein coding changes in P– were a large deletion in the RulA protein, which

296   confers UV resistance in Pph (Sundin *et al.* 2000). Both P+ and P– had MNVs in the *tolQ2* gene—

297   including identical mutations from threonine to leucine at the 121st position. *tolQRA* genes

298   encode membrane proteins that determine phage sensitivity versus resistance in *Escherichia*

299   *coli* bacteria, where *tolQRA* mutants were resistant to phage infection (Picken and Beacham

300   1977). It is therefore possible that mutations in TolQ membrane proteins play a role in both

301   entry of phage ϕ6 into Pph cells, as well as resistance of Pph to phage ϕ6 attack. The specifics of

302   TolQ mutations associated with phage resistance deserve further investigation.

303   Interestingly, we did not find any mutations in genes known to be required for pilus expression

304   (i.e. PilM, PilN, PilO and PilP) (Roine *et al.* 1996, 1998; Taguchi and Ichinose 2011; Nguyen *et al.*

305   2012). This indicates that the link between the P+ and P- phenotypes and their corresponding

306   genotypes is likely to be complex and possibly regulatory in nature.

307   ***Interactions between mutation and gene expression***

308   Of the 20 non-synonymous mutations in P+, six were located in genes with significant

309   expression differences compared to WT; four were up-regulated and two were down-regulated.

310   Of the four genes which were up-regulated, two were annotated as hypothetical proteins of

311   unknown function. However, the two remaining upregulated genes were related to Pph

312   virulence. The first is AvrB4-2, a type III effector protein associated with Pph virulence in

313   common bean (*Phaseolus vulgaris*) plants (Zumaquero *et al.* 2010). P+ had a SNP mutation in

314   *AvrB4-2* that caused an amino acid change from aspartic acid to asparagine at the 240th

315   position. *Avr* genes in *P. syringae* are involved in host-specific resistance mechanisms between

316   the bacterium and its plant hosts (Lee *et al.* 2004). The second SNP was initially identified in a

317   search of 19 *P. syringae* genome sequences for genes related to bacterial pathogenicity (Baltrus

318   *et al.* 2011). A valine at the third amino acid position of the gene was changed to a leucine.

319   In P−, only one of the 26 genes containing non-synonymous mutations showed a

320   significant expression difference compared to WT: a phage integrase-family recombinase. In

321   this gene, two mutations were observed: a change from leucine to proline at position 216, and

322   a change from methionine to leucine at position 219. Also of interest is the combination of a

323   binding site mutation and up-regulation of a type III effector *hopQ2* gene in P−. Type III

324   virulence effector proteins are important for *P. syringae* to infect and grow in plant hosts, and

325   thus these mutations may yield insights into the virulence of P− strains of *P. syringae* (i.e., aside

326   from the expected reduced conditional virulence on plant hosts, due to the lack of type IV pili in

327   P− strains).

328   ***Gene expression differences between P+ and P− phage resistant* P. syringae *pv.* phaseolicola**

329   GE Profiling demonstrated that P+ had significantly more genes both up and down-

330   regulated in comparison to WT than P−: compared to WT, 503 and 104 genes were significantly

331   up-regulated and 451 and 28 genes were downregulated in P+ and P−. GO term enrichment

332   analysis revealed four and 40 GO terms enriched for up-regulation in P+ and P−, and 38 and one

333   GO terms enriched for down-regulation in P+ and P−.

334   The exceptionally low number of GO terms enriched for up-regulation in P+ suggests

335   that transcription and DNA metabolism are likely to be key processes in P+ resistance to phage

336   infection, and the genes annotated for each term represent candidates for further, gene

337   specific analysis. Down-regulation of genes associated with gene expression and translation,

338   nucleotide and ribonucleotide biosynthesis, and cellular respiration in P+ (Figure 2) indicates

339   the potential importance of these pathways in P+ resistance to phage infection.

340   The up-regulation of genes involved in gene expression and ribosome production in P−

341   suggests the up-regulation of protein production in general. Additionally, the up-regulation of

342   genes associated with cellular respiration and ATP synthesis indicates that energy production in

343   general increased in P−. The observed down-regulation of pili-related genes in P− is

344   unsurprising given that this strain lacks pili, however the specific functions of the proteins

345   within this down-regulated set of genes have yet to be investigated.

346   Perhaps of most interest is the observation that 81.1% of GO terms up-regulated in P−

347   were shown to be down-regulated in P+ (Figure 1, Table S3). This result indicates that

348   regulatory changes in similar pathways have led to divergent, independent mechanisms of

349   phage resistance in Pph. A number of genes associated with respiration – specifically TCA cycle,

350   aerobic respiration and acetyl-CoA metabolism genes – were up-regulated in P− and

351   downregulated in P+.

352   There were 10 and seven genes annotated for all three of these GO terms differentially

353   expressed in comparison with WT in P+ and P− respectively, of which four were common to

354   both. The four genes were succinate dehydrogenase, alpha-ketoglutarate dehydrogenase,

355   isocitrate dehydrogenase, and succinyl-CoA synthetase – all encoding key enzymes in the TCA

356   cycle. Plotting the expression levels of these genes reveals a similar pattern of greatest

357   expression in P− and lowest expression in P+ (Figure 2). Viewing these genes in the context of

358   their location in the TCA cycle (Figure 2) reveals that the enzymes they encode are responsible

359   for catalyzing energy-yielding reactions, either in terms of a nucleoside triphosphate or a

360   reduced cofactor. In general, kinetic and allosteric feedback loops between metabolites and

361   enzymes in a biological pathway generally maintain a degree of stability, and the increases or

362   reductions in the activity of one gene versus another does not significantly change flux through

363   the pathway. However, if all of the enzymes are up-regulated/down-regulated in concert, such

364   as in the case of P+ and P−, the whole pathway would be affected. Thus, the expression pattern

365   of these four genes suggests that the TCA pathway is generally up-regulated/down-regulated in

366   P+ and P− respectively, and that its pathway flux relates to pilus regulation

367   The down-regulation of the TCA pathway could indicate lower metabolic rate of the P+

368   strain, potentially a result of its loss of retraction ability. Pilus retraction is mediated by an ATP

369   motor, so loss of retraction ability could lead to a draw of ATP on the bacterial system (Helaine

370   *et al.* 2007). These conclusions are speculative, however, and require further experiments for

371   confirmation.

372   The divergent expression patterns of these TCA cycle genes in P− versus P+ reflect the

373   general reciprocal GE profiles of the two strains. Overall, genes up-regulated in P− are down-

374   regulated in P+ and vice versa (Table S3). The only exception to these reciprocal GE profiles is

375   that transposition is up-regulated in both P− and P+. The fact that in P− the transposition genes

376   are up-regulated, while pili genes are down-regulated may suggest a mutation causing a

377   fundamental disconnect between these two otherwise co-regulated processes. This

378   observation will be a continued area of study, since the exact regulatory mechanisms behind

379   transposition in Pph are not well known.

380   The results of our study are useful for general research efforts harnessing phage φ6 and

381   Pph. In particular, these microbial strains have been used in prior studies in experimental

382   ecology and evolutionary biology, especially to conduct experimental evolution that addressed

383   fundamental questions in these fields (Turner and Chao 1999; Montville *et al.* 2005; Dessau *et*

384   *al.* 2012; Goldhill and Turner 2014). However, these prior studies have not focused on co-

385   evolutionary interactions between phage φ6 and Pph, perhaps because resulting genetic

386   changes in the host bacteria could not be compared to existing full genome sequences of the

387   founding strains. Our study addresses this limitation and lends more power to the phage

388   φ6/Pph model system, because such work can now harness affordable re-sequencing

389   approaches to compare ancestral and derived strains of bacteria. Also, recent studies show that

390   Cystoviruses are readily found in the terrestrial phyllosphere in temperate environments,

391   especially in association with bacteria residing on or within common legumes such as white

392   clover plants (Mindich *et al.* 1999; Silander *et al.* 2005; O'Keefe *et al.* 2010; Díaz-Muñoz *et al.*

393   2013). These studies emphasize the prevalence of RNA phages in terrestrial biomes (Mindich *et*

394   *al.* 1999; Silander *et al.* 2005; O'Keefe *et al.* 2010; Díaz-Muñoz *et al.* 2013), and to our

395   knowledge Cystoviruses constitute the only non-marine RNA phages that are actively

396   researched in the wild (Culley *et al.* 2006). Whereas these earlier studies were used to infer

397   population structure and isolation-by-distance in Cystoviruses, our study is useful for future

398   empirical work that manipulates Cystoviruses and Pph strains on legumes in field experiments

399   or in the greenhouse. For example, the bacteria characterized in our study could be used to

400   examine the effects of phage infection on the community dynamics of bacteria that differ in

401   plant pathogenicity, under realistic conditions such as on leaf surfaces. Overall, characterization

402   of Pph pili mutants in the current study support the crucial need to further develop emerging

403   models in phage/bacteria interactions, which should broaden the understanding of symbiotic

404   interactions between the most prevalent entities in the biosphere (Hendrix 2002; Wasik and

405   Turner 2013).

406

407   **Acknowledgements**

412   **References**

413   Adomas, A. B., F. Lopez-Giraldez, T. A. Clark, Z. Wang, and J. P. Townsend, 2010 Multi-targeted
414   priming for genome-wide gene expression assays. BMC Genomics 11: 477.

415   Arnold, D. L., H. C. Lovell, R. W. Jackson, and J. W. Mansfield, 2011 Pseudomonas syringae pv.
416   phaseolicola: From "has bean" to supermodel. Mol. Plant Pathol. 12: 617–627.

417   Baltrus, D. A., M. T. Nishimura, A. Romanchuk, J. H. Chang, M. S. Mukhtar *et al.*, 2011 Dynamic
418   evolution of pathogenicity revealed by sequencing and comparative genomics of 19
419   pseudomonas syringae isolates. PLoS Pathog. 7.:

420   Bamford, D. H., M. Romantschuk, and P. J. Somerharju, 1987 Membrane fusion in prokaryotes:
421   bacteriophage phi 6 membrane fuses with the Pseudomonas syringae outer membrane.
422   EMBO J. 6: 1467–1473.

423   Burdman, S., O. Bahar, J. K. Parker, and L. de la Fuente, 2011 Involvement of type IV pili in
424   pathogenicity of plant pathogenic bacteria. Genes (Basel). 2: 706–735.

425   Carbon, S., A. Ireland, C. J. Mungall, S. Shu, B. Marshall *et al.*, 2009 AmiGO: Online access to
426   ontology and annotation data. Bioinformatics 25: 288–289.

427   Culley, A. I., A. S. Lang, and C. A. Suttle, 2006 Metagenomic analysis of coastal RNA virus
428   communities. Science 312: 1795–1798.

429   Darling, A. C. E., B. Mau, F. R. Blattner, and N. T. Perna, 2004 Mauve: Multiple alignment of
430   conserved genomic sequence with rearrangements. Genome Res. 14: 1394–1403.

431   Davidsen, T., E. Beck, A. Ganapathy, R. Montgomery, N. Zafar *et al.*, 2009 The comprehensive
432   microbial resource. Nucleic Acids Res. 38.:

433   Dennehy, J. J., 2009 Bacteriophages as model organisms for virus emergence research. Trends
434   Microbiol. 17: 450–457.

435   Dennehy, J. J., S. T. Abedon, and P. E. Turner, 2007 Host density impacts relative fitness of
436   bacteriophage Φ6 genotypes in structured habitats. Evolution (N. Y). 61: 2516–2527.

437   Dennehy, J. J., N. A. Friedenberg, Y. W. Yang, and P. E. Turner, 2007 Virus population extinction
438   via ecological traps. Ecol. Lett. 10: 230–240.

439   Dessau, M., D. Goldhill, R. L. McBride, P. E. Turner, and Y. Modis, 2012 Selective Pressure
440   Causes an RNA Virus to Trade Reproductive Fitness for Increased Structural and Thermal
441   Stability of a Viral Enzyme. PLoS Genet. 8.:

442   Díaz-Muñoz, S. L., O. Tenaillon, D. Goldhill, K. Brao, P. E. Turner *et al.*, 2013 Electrophoretic
443   mobility confirms reassortment bias among geographic isolates of segmented RNA phages.
444   BMC Evol. Biol. 13: 206.

445   Duffy, S., P. E. Turner, and C. L. Burch, 2006 Pleiotropic costs of niche expansion in the RNA
446   bacteriophage Φ6. Genetics 172: 751–757.

447   Goldhill, D. H., and P. E. Turner, 2014 The evolution of life history trade-offs in viruses. Curr.
448   Opin. Virol. 8: 79–84.

449   Gottlieb, P., S. Metzger, M. Romantschuk, J. Carton, J. Strassman *et al.*, 1988 Nucleotide
450   sequence of the middle dsRNA segment of bacteriophage phi 6: placement of the genes of
451   membrane-associated proteins. Virology 163: 183–190.

452   Helaine, S., D. H. Dyer, X. Nassif, V. Pelicic, and K. T. Forest, 2007 3D structure/function analysis
453   of PilX reveals how minor pilins can modulate the virulence properties of type IV pili. Proc.
454   Natl. Acad. Sci. U. S. A. 104: 15888–15893.

455   Hendrix, R. W., 2002 Bacteriophages: evolution of the majority. Theor. Popul. Biol. 61: 471–480.

456   Huerta, A. M., H. Salgado, D. Thieffry, and J. Collado-Vides, 1998 RegulonDB: a database on
457   transcriptional regulation in Escherichia coli. Nucleic Acids Res. 26: 55–59.

458   Jackson, R. W., E. Athanassopoulos, G. Tsiamis, J. W. Mansfield, A. Sesma *et al.*, 1999
459   Identification of a pathogenicity island, which contains genes for virulence and avirulence,

460   on a large native plasmid in the bean pathogen Pseudomonas syringae pathovar
461   phaseolicola. Proc. Natl. Acad. Sci. U. S. A. 96: 10875–10880.

462 Joardar, V., M. Lindeberg, R. W. Jackson, J. Selengut, R. Dodson *et al.*, 2005 Whole-genome
463   sequence analysis of Pseudomonas syringae pv. phaseolicola 1448A reveals divergence
464   among pathovars in genes involved in virulence and transposition. J. Bacteriol. 187: 6488–
465   6498.

466 Johnson, K., and S. Lory, 1987 Characterization of Pseudomonas aeruginosa mutants with
467   altered piliation. J. Bacteriol. 169: 5663–7.

468 Lee, C. C., M. D. Wood, K. Ng, C. B. Andersen, Y. Liu *et al.*, 2004 Crystal structure of the type III
469   effector AvrB from Pseudomonas syringae. Structure 12: 487–494.

470 Lythgoe, K. A., and L. Chao, 2003 Mechanisms of coexistence of a bacteria and a bacteriophage
471   in a spatially homogeneous environment. Ecol. Lett. 6: 326–334.

472 Mattick, J. S., 2002 Type IV pili and twitching motility. Annu. Rev. Microbiol. 56: 289–314.

473 Mindich, L., 2006 Phages with segmented double-stranded RNA genomes, pp. 197–210 in *The*
474   *Bacteriophages*, edited by C. R. and A. ST. Oxford University Press, New York.

475 Mindich, L., X. Qiao, J. Qiao, S. Onodera, M. Romantschuk *et al.*, 1999 Isolation of additional
476   bacteriophages with genomes of segmented double- stranded RNA. J. Bacteriol. 181:
477   4505–4508.

478 Montville, R., R. Froissart, S. K. Remold, O. Tenaillon, and P. E. Turner, 2005 Evolution of
479   mutational robustness in an RNA virus. PLoS Biol. 3: 1939–1945.

480 Nguyen, L. C., F. Taguchi, Q. M. Tran, K. Naito, M. Yamamoto *et al.*, 2012 Type IV pilin is
481   glycosylated in Pseudomonas syringae pv. tabaci 6605 and is required for surface motility
482   and virulence. Mol. Plant Pathol. 13: 764–774.

483 O'Brien, H. E., Y. Gong, P. Fung, P. W. Wang, and D. S. Guttman, 2011 Use of low-coverage,
484   large-insert, short-read data for rapid and accurate generation of enhanced-quality draft
485   Pseudomonas genome sequences. PLoS One 6.:

486 O'Keefe, K. J., O. K. Silander, H. McCreery, D. M. Weinreich, K. M. Wright *et al.*, 2010
487   Geographic differences in sexual reassortment in RNA phage. Evolution (N. Y). 64: 3010–
488   3023.

489 Picken, R. N., and I. R. Beacham, 1977 Bacteriophage-resistant mutants of Escherichia coli K12.
490   Location of receptors within the lipopolysaccharide. J. Gen. Microbiol. 102: 305–318.

491   Quinlan, A. R., and I. M. Hall, 2010 BEDTools: A flexible suite of utilities for comparing genomic
492   features. Bioinformatics 26: 841–842.

493   Roine, E., D. N. Nunn, L. Paulin, and M. Romantschuk, 1996 Characterization of genes required
494   for pilus expression in Pseudomonas syringae pathovar phaseolicola. J. Bacteriol. 178:
495   410–417.

496   Roine, E., D. M. Raineri, M. Romantschuk, M. Wilson, and D. N. Nunn, 1998 Characterization of
497   type IV pilus genes in Pseudomonas syringae pv. tomato DC3000. Mol. Plant. Microbe.
498   Interact. 11: 1048–1056.

499   Romantschuk, M., and D. H. Bamford, 1985 Function of Pili in Bacteriophage-Phi-6 Penetration.
500   J. Gen. Virol. 66: 2461–2469.

501   Romantschuk, M., and D. H. Bamford, 1986 The causal agent of halo blight in bean,
502   Pseudomonas syringae pv. phaseolicola, attaches to stomata via its pili. Microb. Pathog. 1:
503   139–148.

504   Romantschuk, M., E.-L. Nurmiaho-Lassila, E. Roine, and A. Suoniemi, 1993 Pilus-mediated
505   adsorption of Pseudomonas syringae to the surface of host and non-host plant leaves. J.
506   Gen. Microbiol. 139: 2251–2260.

507   Silander, O. K., D. M. Weinreich, K. M. Wright, K. J. O'Keefe, C. U. Rang et al., 2005 Widespread
508   genetic exchange among terrestrial bacteriophages. Proc. Natl. Acad. Sci. U. S. A. 102:
509   19009–19014.

510   Sundin, G. W., J. L. Jacobs, and J. Murillo, 2000 Sequence diversity of rulA among natural
511   isolates of Pseudomonas syringae and effect on function of rulAB-mediated UV radiation
512   tolerance. Appl. Environ. Microbiol. 66: 5167–5173.

513   Taguchi, F., and Y. Ichinose, 2011 Role of type IV pili in virulence of Pseudomonas syringae pv.
514   tabaci 6605: correlation of motility, multidrug resistance, and HR-inducing activity on a
515   nonhost plant. Mol. Plant. Microbe. Interact. 24: 1001–1011.

516   Townsend, J. P., 2004 Resolution of large and small differences in gene expression using models
517   for the Bayesian analysis of gene expression levels and spotted DNA microarrays. BMC
518   Bioinformatics 5: 54.

519   Townsend, J. P., and D. L. Hartl, 2002 Bayesian analysis of gene expression levels: statistical
520   quantification of relative mRNA level across multiple strains or treatments. Genome Biol.
521   3: RESEARCH0071.

522   Turner, P. E., and L. Chao, 1999 Prisoner's dilemma in an RNA virus. Nature 398: 441–443.

523   Wasik, B. R., and P. E. Turner, 2013 On the biological success of viruses. Annu. Rev. Microbiol.
524   67: 519–41.

525   Zumaquero, A., A. P. Macho, J. S. Rufián, and C. R. Beuzón, 2010 Analysis of the role of the type
526   III effector inventory of Pseudomonas syringae pv. phaseolicola 1448a in interaction with
527   the plant. J. Bacteriol. 192: 4474–4488.

528   **Figure Legends**

529   **Figure 1:** Clustergram of standardized values of all genes (left) and significantly ($P < 0.05$)

530   differentially expressed genes (right). Significantly differentially expressed genes are those that

531   are significantly up-regulated (red) /down-regulated (green) in P+ or P− relative to WT.

532   **Figure 2:** Expression levels of key enzymes of the TCA cycle in P− (yellow), WT (blue), and P+

533   (red) strains.

534

535   **Supplementary Material**

536   **Table S1:** Genomic mutations present in super-piliated Pph in comparison to WT.

537   **Table S2:** Genomic mutations present in non-piliated Pph in comparison to WT.

538   **Table S3:** Gene expression differences between super and non-piliated Pph in comparison to

539   WT. The number 1 indicates genes which are significantly up or down regulated, 0 indicates

540   genes which do not show a significant difference.

541   **Figure S1:** Visualization of the GO terms enriched in the genes up-regulated in P+. Shading

542   corresponds to level of significance. The vertical axis generally indicates greater or lesser

543   specificity of the GO term because of the hierarchical definitions of the terms. The three major

544   biological functions that are enriched are those of Gene expression and translation, nucleotide

545   and ribonucelotide biosynthesis, and respiration. The majority of unboxed terms are mostly

546   general terms lacking a unifying biological function.

25

547 **Figure S2:** Visualization of the GO terms enriched in the genes down-regulated in P+. Shading

548 corresponds to level of significance. The vertical axis generally indicates greater or lesser

549 specificity of the GO term because of the hierarchical definitions of the terms. The three major

550 biological functions that are enriched are those of gene expression and translation, nucleotide

551 and ribonucelotide biosynthesis, and respiration. The majority of unboxed terms are mostly

552 general terms lacking a unifying biological function.

553 **Figure S3:** Visualization of the GO terms enriched in the genes upregulated in P–. Shading

554 corresponds to level of significance. The vertical axis generally indicates greater or lesser

555 specificity of the GO term because of the hierarchical definitions of the terms. The three major

556 biological functions that are enriched are those of gene expression and translation, nucleotide

557 and ribonucelotide biosynthesis, and respiration. The majority of unboxed terms are mostly

558 general terms lacking a unifying biological function.

559 **Figure S4:** Visualization of the GO terms enriched in the genes downregulated in P–. Shading

560 corresponds to level of significance. The vertical axis generally indicates greater or lesser

561 specificity of the GO term because of the hierarchical definitions of the terms. The three major

562 biological functions that are enriched are those of gene expression and translation, nucleotide

563 and ribonucelotide biosynthesis, and respiration. The majority of unboxed terms are mostly

564 general terms lacking a unifying biological function.

565

566 **Table 1:** A summary of genomic and gene expression differences between super-piliated and

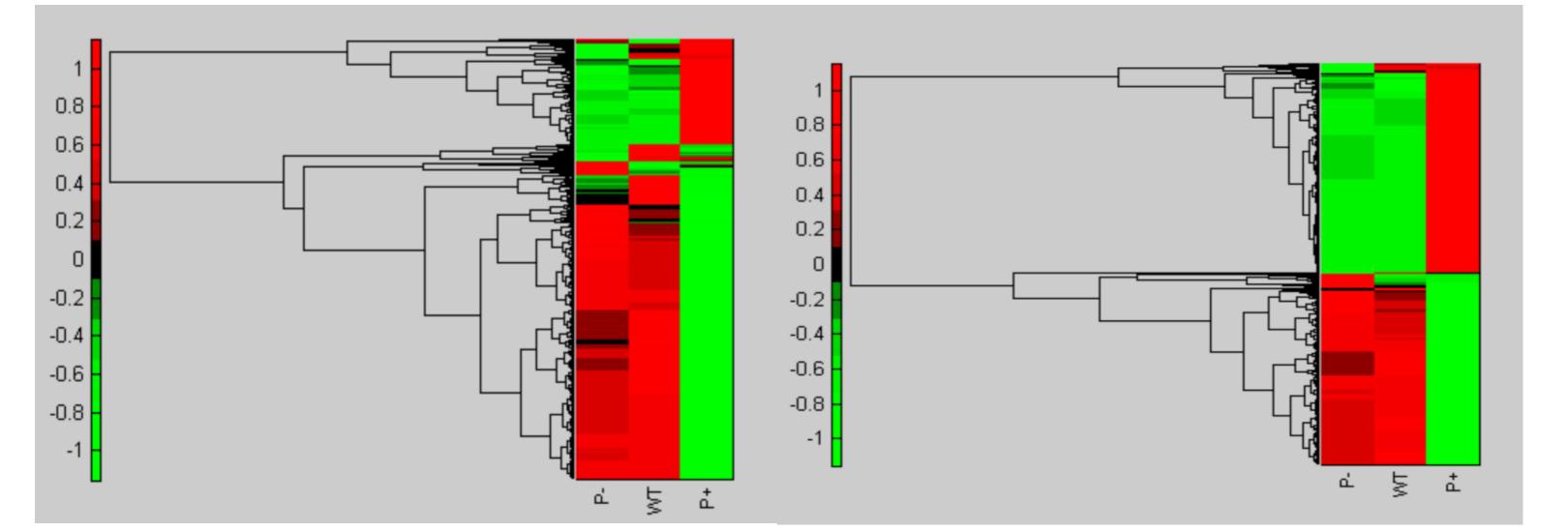567 non-piliated Pph.

Mutation Type	Strain

	P+	P−
Large Scale deletions	26	4
Deletion-Insertion Polymorphisms	5	8
Multi-nucleotide Variants	7	8
Single Nucleotide Polymorphisms	59	62
Total	97	82
Expression Differences		
Genes UP–regulated	503	104
Genes Down-regulated	451	28
Gene Ontology Terms UP–regulated	4	3
Gene Ontology Terms Down-regulated	38	1

8   **Table 2:** Details putative sigma70 binding sites inferred from the RegulonDB database for which significant regulatory changes were detected.

Reference Position	Consensus Position	Variation Type	Length	Reference	Allele Variations	Overlapping Annotations	Upstream annotation	Upstream distance	Upstream orientation	Downstream annotation
4059	4059	MNV	3	---	GCT		PSPPH_A0012/hopQ1 (type III effector HopQ1)	158	−1	
546554	546628	Deletion	75				PSPPH_0642 (LuxR family D--binding response regulator)	53	1	"PSPPH_0643 (polyamine ABC transporter, ATP-binding protein)"
101598	101722	Deletion	125				"PSPPH_1090 (ribosomal subunit interface protein, putative)"	135	−1	PSPPH_1091 (iron(III) dicitrate transport protein fecA)
272415	272478	Deletion	64				"PSPPH_1387 (amino acid ABC transporter, permease protein)"	12	−1	"PSPPH_1388 (amino acid ABC transporter, periplasmic amino acid-binding protein)"
141653	141921	Deletion	269			PSPPH_4466 (hypothetical protein)				
259583	259844	Deletion	262			PSPPH_0408 (hypothetical protein)				

2	127	Deletion	126			PSPPH_4902 (molybdenum-pterin binding domain-containing protein)			
14090	14090	SNP	1	G	C	PSPPH_0186 (citrate transporter)	159	−1	PSPPH_0187 (hypothetical protein)

9