Genome-Wide Association to Study the Host-Specificity Determinants of *Xanthomonas perforans*

Newberry, E. A.[1], Minsavage, G.V.[2], Holland, A. [1], Jones, J.B.[2], and Potnis, N.[1*]

[1]Department of Entomology and Plant Pathology, Auburn University, AL 36849

[2]Department of Plant Pathology, University of Florida, FL 32611.

*Corresponding author: Neha Potnis, nzp0024@auburn.edu

Abstract

Xanthomonas perforans and *X. euvesicatoria* are the causal agents of bacterial spot disease of tomato and pepper, endemic to the Southeastern United States. Although very closely related, the two bacterial species differ in host-specificity, where *X. perforans* is the dominant pathogen of tomato and *X. euvesicatoria* of pepper. This is in part due to the activity of avirulence proteins that are secreted by *X. perforans* strains and elicit an effector-triggered immunity (ETI) in pepper leaves, thereby restricting pathogen growth. In recent years, the emergence of several pepper-pathogenic *X. perforans* lineages has revealed variability within the bacterial species to multiply and cause disease in pepper, even in the absence of avirulence gene activity. Here, we investigated the basal evolutionary processes underlying the host-range of this species using multiple genome-wide association analyses. Surprisingly, we identified two novel gene-candidates that were significantly associated with pepper-pathogenic *X. perforans* and *X. euvesicatoria.* Both candidates were predicted to be involved in the transport/acquisition of nutrients common to plant-cell wall or apoplast and included a TonB-dependent receptor, which was disrupted through independent mutations within the *X. perforans* lineage. The other included a symporter of protons/glutamate, *gltP*, enriched with pepper-associated mutations near the promoter and start

codon of the gene. Functional analysis of these candidates revealed that only the TonB-dependent receptor had a minor effect on the symptom development and growth of *X. perforans* in pepper leaves, indicating that pathogenicity to this host may have evolved independently within the bacterial species and is likely a complex, multigenic trait.

Keywords: Genome-wide association study, host range determinant, *Xanthomonas perforans*, avirulence genes

Introduction

Xanthomonas is a genus of Gram-negative, Gammaproteobacteria and is responsible for numerous plant diseases of economic importance worldwide (Hayward 1993). A characteristic feature of the genus is the high degree of host specificity exhibited by the individual species and/or sub-species found within it, which are often limited in host-range to a particular botanical family (An et al. 2019). As such, most pathogenic *Xanthomonas* species are not commonly encountered with high abundance outside of the host environment (Jacques et al. 2016) and exhibit a highly clonal population structure (Ferreira et al. 2019; Mhedbi-Hajri et al. 2013; An et al. 2019; Dhakal et al. 2019), with horizontal gene transfer and homologous recombination serving as the primary factors driving diversification at the intraspecific level (Huang et al. 2015; Jibrin et al. 2018; Aritua et al. 2015). Thus, the divergence of *Xanthomonas* spp. is widely considered to be driven by ecological isolation associated with the adaptation to a particular host (Huang et al. 2015; Mhedbi-Hajri et al. 2013). These characteristics have enabled this group of pathogens to emerge as attractive model systems for the investigation of host-range evolution among plant pathogenic bacteria.

Xanthomonads employ a type 3 secretion system to translocate effector proteins, termed type 3 secreted effectors (T3SEs), into the plant cell and cause disease. These proteins serve to facilitate

the infection process by suppressing the plant immune response and in modulating the physiology of the host (White et al. 2009). Many T3SEs (more specifically referred to as avirulence proteins) have also gained attention for their role in mediating genotype-specific interactions and may serve as a host-limiting factors when recognized by a corresponding plant-immune receptor (Stall et al. 2009; Minsavage 1990). This 'gene-for-gene' interaction leads to the induction of a localized, programmed host-cell death, referred to as an effector-triggered immunity (ETI; Jones and Dangl 2006). However, under selection pressure, such elicitors of ETI (*i.e.* avirulence proteins) may be rapidly pseudogenized or lost, often resulting in a compatible interaction and host-range expansion of the pathogen (Kousik and Ritchie 1996; Lovell et al. 2011; Swords et al. 1996).

In some cases, the loss of a single avirulence gene does not always lead to plant-pathogen compatibility (Schwartz et al. 2015; Hajri et al. 2009). To account for this observation, a repertoire-for-repertoire hypothesis was proposed (Hajri et al. 2009), which postulated that a network of T3SEs and their interaction with a corresponding set of host factors were responsible for determining the host-range spectra of *Xanthomonas* spp. Within the framework of this model, host adaptation is expected to result from the acquisition, loss, or genomic rearrangement of multiple T3SEs as well as their functional redundancy. Most evidence supporting this hypothesis can be found in the form of transcription activation-like effectors, which often display host-associations due to their role in manipulating the expression of specific plant genes via DNA binding activity, irrespective of their relative importance in determining disease outcome (Ruh et al. 2017; Chen et al. 2018; Hutin et al. 2015). With the advance of the genomics era, it has become increasingly recognized that an array of factors beyond T3SEs including, but not limited to the evasion of microbe-associated molecular pattern recognition, environmental sensory and chemotaxis, as well

as specific nutritional requirements influence the host-range of *Xanthomonas* spp. (Jacques et al. 2016).

Bacterial leaf spot disease is a major limiting factor of tomato (*Solanum lycopersicum*) and pepper (*Capsicum annuum*) production worldwide and is endemic in the Southeast United States, where *X. perforans* and *X. euvesicatoria* are the dominant pathogens of these two crops respectively (Potnis et al. 2015; Jones et al. 2004). Although the two bacterial species are very closely related, sharing approximately 98% average nucleotide identity over 80% of their genomes (Barak et al. 2016), they display marked differences in host specificity. While the host range of *X. euvesicatoria* is reported to encompass several solanaceous plant species including tomato and pepper, the pathogen is more aggressive on the latter host (Areas et al. 2015; Osdaghi et al. 2016; Roach et al. 2018). In contrast, the host range of *X. perforans* is widely considered to be restricted to tomato, in part due to the activity of several avirulence proteins (either AvrBsT, AvrXv3, and/or XopJ6) that are secreted by strains of the pathogen and elicit an ETI in most pepper varieties (Minsavage 1990; Astua-Monge et al. 2000; Iruegas-Bocardo et al. 2018). Strikingly, mutation of the *avrBsT* and *avrXv3* genes revealed a differential ability among individual *X. perforans* strains to multiply and cause disease when inoculated into pepper leaves (Schwartz et al. 2015). Interestingly, the pathogenicity of these strains correlated with their underlying genetic background, indicating the heritability of this trait and that other genetic factors may contribute to the host range expansion of this pathogen.

In recent years, significant changes have been observed in the *X. perforans* population recovered from tomato and pepper fields in the Southeastern United States. These include a shift in race associated with a decline in populations carrying a functional copy of the avirulence gene *avrXv3* (Timilsina et al. 2016). They also include diversification into at least six distinct phylogenetic

groups as the result of genome-wide homologous recombination events derived from *X. euvesicatoria* and related pathogens found within the *X. euvesicatoria* species complex (Jibrin et al. 2018; Timilsina et al. 2019; Newberry et al. 2019), which is comprised of a heterogenous group of bacteria that are responsible for diseases of diverse woody and herbaceous plant species (Parkinson et al. 2009; Bansal et al. 2018; Constantin et al. 2016). Concurrent with the recent diversification of *X. perforans*, two separate lineages of the pathogen have arisen that display an expanded natural host range to include both tomato and pepper. Our previous analysis of these lineages did not yield evidence for a common horizontal transfer or recombination event that would readily explain such a shift in host range, suggesting that independent evolutionary processes may have given rise to this trait (Newberry et al. 2019).

Genome-wide association (GWAS) is a relatively new technique in the field of microbial genomics that has been used to link genetic variants to phenotype (San et al. 2020). This approach has been successful in the identification of genes and/or mutations responsible for traits that are under strong selection pressure, such as antibiotic resistance (Suzuki et al. 2016; Lees et al. 2016; Hicks et al. 2019). It has also proven useful in the identification of genetic variants responsible for more complex traits such as virulence and host specificity among human and animal pathogenic bacteria (Sheppard et al. 2013; Méric et al. 2018; Berthenet et al. 2018). Here, we leveraged the high degree of host-specificity exhibited by *Xanthomonas* species/pathovars and the diversity of related pathogens found within the larger *X. euvesicatoria* species complex to test for genetic signatures that may indicate the ability of *X. perforans* to infect pepper plants, independent of an ETI. Using multiple GWAS approaches, we identified two novel gene-candidates, each located in a previously described recombination hot-spot (Newberry et al. 2019), that were significantly associated with

112 pepper pathogenic *X. perforans* and *X. euvesicatoria* and confirmed that the host range of *X.*
113 *perforans* is likely a complex, multigenic trait.

114 **Results**

115 **Pathogenicity phenotyping supports a correlation between host-range and population**
116 **structure within *X. perforans* and confirms the presence of three pepper-pathogenic lineages**
117 **of the pathogen.**

118 Aside from the *X. perforans* strains previously examined by Schwartz et al. (2015), 17 additional
119 strains that did not elicit a hypersensitive response (HR) in pepper cv. Early CalWonder (ECW)
120 were identified (Fig. 1). When infiltrated into pepper leaves using a low concentration of bacterial
121 cells (10^3 CFU ml^{-1}), all of these strains induced circular, water-soaked lesions approximately two
122 weeks post-inoculation and were scored as pathogenic (Fig. 1B). As expected, the genomes of
123 these strains lacked evidence of any previously described avirulence gene. Moreover, the presence
124 of either *avrBsT*, *avrXv3*, and/or *xopJ6* was confirmed for those strains that were HR positive
125 based on BLASTn analysis (Evalue < $10e^{-5}$; Fig. 1C).

126 When analyzed along with 53 additional *X. euvesicatoria* genomes, a maximum likelihood
127 reconstruction of 93,380 high quality, core genome single-nucleotide polymorphisms (SNPs),
128 coupled with a hierarchical Bayesian analysis of population structure (Tonkin-Hill et al. 2019)
129 revealed a phylogeny (Fig. 1C) like that described in previous studies (Timilsina et al. 2019;
130 Newberry et al. 2019). This showed that related *X. euvesicatoria* pathovars comprised a
131 paraphyletic collection of lineages and branched separately from pepper-pathogenic *X.*
132 *euvesicatoria*. Likewise, the *X. perforans* lineage was divided into multiple sub-clades, referred to
133 here as sequence clusters (SCs). Pepper-pathogenic *X. perforans* strains were dispersed among

clusters SC3, SC4, and SC6. Among these, clusters SC3 and SC6 harbored strains that were isolated from naturally infected pepper leaves. While the SC4 strains have only been isolated from symptomatic tomato plants to date, they were also capable of producing typical water-soaked lesions in pepper (Fig. 1B). Finally, our previous analysis of the *X. perforans* core genome revealed six distinct genetic groups (Newberry et al. 2019). Here, analysis with the FastBAPS algorithm collapsed strains from clusters SC1 and SC2 into a single group, designated SC1/SC2, and included only strains that were non-pathogenic to pepper plants in the absence of an effector-triggered immunity (ETI).

Due to the scarcity of *X. perforans* strains that did not induce an ETI in pepper, we performed two rounds of association testing for the results described below. The first included a conservative dataset ($n = 71$), which included only *X. perforans* strains for which phenotypic data was available and pepper-pathogenic *X. euvesicatoria* strains that were isolated from naturally infected pepper leaves and/or were previously tested for pathogenicity (Schwartz et al. 2015; Roach et al. 2018). We also conducted an exploratory association analyses ($n = 131$), which was based on the hypothesis that host-range is associated with genetic background within the *X. perforans* lineage and that related *X. euvesicatoria* pathovars are non-pathogenic to pepper plants, independent of an ETI. It should be noted that the pathogenicity of strains assigned to many of these pathovars including *X. euvesicatoria* pv. *commiphorae* LMG26789, *X. euvesicatoria* pv. *citrumelonis* F1, *X. euvesicatoria* pv. *alfalfae* CFBP3836, and the rose pathogen (GEV-Rose-07) was previously assessed (Yaripour et al. 2018; Huang et al. 2013; Bansal et al. 2018; Samanta et al. 2013), confirming the inability of these strains to cause typical disease symptoms in pepper leaves.

Pan-genome association analysis identifies a TonB-dependent receptor (TBDR) associated with pepper-pathogenic *X. perforans* and *X. euvesicatoria.*

Pan-genome analysis of the *X. euvesicatoria* species complex with Roary (Page et al. 2015) revealed a complement of 11,562 orthologous groups (OGs), of which 3,430 comprised the core and soft-core genomes (present in > 95% of strains), while 2,133 OGs were more broadly distributed between 15 and 95% of strains, and 5,999 were more sparsely distributed between 1 and 15% of strains. Association analysis of these groups produced similar results for both the conservative and exploratory datasets, yielding a single OG (a predicted TonB-dependent receptor) that was significantly associated with pepper-pathogenic *X. perforans* and *X. euvesicatoria* strains (Table 1). Among the conservative genome sample, this group displayed a nearly perfect association with pepper-pathogenicity (99% sensitivity, 100% specificity) and was the only group with a probability < 0.01 after Bonferroni correction. Among the exploratory genome sample, this remained the most strongly associated group (Bonferroni $p = 1.14E^{-17}$), with an overall diagnostic accuracy of 92%. While a number of other OGs displayed potentially significant associations (Bonferroni $p \leq 3.90E^{-12}$) in the exploratory sample, the diagnostic accuracy of these groups dropped off significantly (< 75%), indicating that these were likely spurious associations resulting from a strong lineage effect. Further investigation based on a quantile-quantile (QQ) plot of the $-\log_{10}$ transformed *p*-values confirmed these suspicions (Supplementary Fig. 1).

Downstream analysis of the pepper-associated TBDR indicated that the gene was not actually unique to pepper-pathogenic *X. perforans* and *X. euvesicatoria* strains, but that it was a predicted pseudogene, while most non-pathogenic strains possessed an intact copy of the coding sequence (Fig. 2A and 2B). Notably, with a sensitivity of 87% among the exploratory dataset, four strains from non-pathogenic clades (LMG26789, CFBP3836, CFBP6369, and LH3) were also predicted to carry this pseudogene. Although, downstream analysis of the TBDR sequences with ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder/) confirmed this only for *X. euvesicatoria* pv. *alii*

CFBP6369 and *X. perforans* LH3, both strains for which pathogenicity to pepper plants in the absence of an ETI remains undetermined. Moreover, strain LH3 contained a 180 bp truncation at the beginning of the coding sequence, which was unlikely to affect the expression of the gene (Fig 2A).

Examination of the TBDR alleles among pepper-pathogenic strains revealed that the coding sequences had indeed been disrupted through independent mutations. Among these, *X. perforans* clusters SC3, SC4, and *X. euvesicatoria* carried nearly identical alleles (≥ 99.9% pairwise identity) with an approximately 130 bp deletion near the center of the gene (positions 1152 to 1280), while the SC6 allele contained a point mutation (C1734G) and was more closely related to that found in non-pathogenic *X. perforans* strains of cluster SC1/SC2. Interestingly, *X. perforans* cluster SC5 strains, for which pathogenicity could not be assessed independent of an effector triggered immunity, carried the same allele as other non-pathogenic strains. Examination of the genomic context of the pepper-associated TBDR indicated that it was part of a previously described carbohydrate utilization locus and was flanked on either side by operons involved in the metabolism of galactose or xylan/xylose (Fig. 2B).

Single nucleotide polymorphism (SNP) and unitig-based tests identify pepper-associated mutations in a symporter of protons/glutamate, *gltP*, and its promoter region.

We first tested for associations between 93,380 core-genome SNPs, identified within the *X. euvesicatoria* species complex, and pepper pathogenicity using the conservative sample of genomes. This analysis revealed 45 SNPs that displayed significant associations ($p = 0.0032$ and 100% diagnostic accuracy) with pepper pathogenic strains and mapped to four distinct chromosomal regions encompassing nine coding sequences (Fig. 3A). The pepper associated variants were found among genes that encoded for a predicted membrane protein, two sugar/amino

203 acid transporters, two type VI secretion system proteins, a type II secretion system protein, a
204 predicted Fe-S oxidoreductase, a Hpr(Ser) kinase/phosphatase, and a hypothetical protein of
205 unknown function (Supplementary Table 3). However, due to the limited power to test for
206 associations with this more conservative sample size, we also conducted association tests of
207 genome-wide unitigs (which are assembled k-mers) using the exploratory genome sample
208 described above (Fig. 3B).

209 As expected, the QQ-plot of $-\log_{10}$ transformed *p*-values suggested that these associations suffered
210 from inflated *p*-values due to poorly controlled population structure (Supplementary Fig. 1).
211 Nonetheless, we identified 22 significantly associated unitigs ($p > 40$ after $-\log_{10}$ transformation
212 and diagnostic accuracy ≥ 98%), located in two of the four genomic regions identified with the
213 SNP-based associations. All but three of these unitigs mapped to the proton/glutamate-aspartate
214 symporter, *gltP* (BJD13_RS22165), and corresponded to two missense and three synonymous
215 mutations near the start codon of the gene, as well as several SNPs and one insertion/deletion site
216 within 200 bp upstream of the coding sequence (Fig. 3C). The remaining significant unitigs
217 mapped to a predicted major facilitator superfamily (MFS) encoding gene (BJD13_RS08055).
218 Together, these revealed a complex variant (Fig. 3D) characterized by an approximately 50 bp
219 deletion that led to an early stop codon, ancestral to the pepper-pathogenic *X. euvesicatoria* lineage
220 and *X. perforans* clusters SC3 and SC4. However, this insertion/deletion site overlapped a
221 synonymous mutation that distinguished the pepper-pathogenic strains of cluster SC6 from those
222 located in clades that contain non-pathogenic strains, which both carried intact coding sequences.

223 **Functional analysis of the pepper-associated variants provides limited evidence for their role**
224 **in mediating the host range of *X. perforans*.**

To assess the function of the pepper associated TBDR, we performed allele exchange experiments and replaced the disrupted coding sequence found in pepper pathogenic *X. perforans* strains AL65 (cluster SC6) and GEV2393 (cluster SC3) with that of an intact coding sequence (2,937 bp) derived from non-pathogenic strain Xp5-6 (cluster SC1/SC2). Pathogenicity analysis revealed that the AL65 ($TBDR_{2937}$) allele exchange mutant displayed attenuated water-soaking in pepper leaves at the onset of symptom development, four days post-infiltration, while no discernable differences were observed in tomato leaves (Fig. 4A). In preliminary experiments, a similar phenotype was observed for the GEV2393 (TBDR2937) allele exchange mutant (data not shown). However, as these differences in symptom development were more subtle, strain AL65 and its derivatives were selected for further analysis.

This attenuated symptom development was accompanied by a mean difference in CFU/cm^2 of approximately 0.33 log units at four-days post-infiltration in pepper leaves. While a similar trend in mean bacterial growth was also observed between the allele exchange mutant and wild type strain in tomato leaves, these differences were not statistically significant, based on a mixed model analysis of variance ($\alpha = 0.05$). When the SC1/SC2 TBDR allele was complemented back with that of the wildtype in strain AL65, we observed a partial recovery of water-soaking symptoms and although not statistically significant, an upward trend in bacterial growth relative to the allele exchange mutant (Fig. 4B). This phenotype was further investigated on pepper plants using a dip inoculation technique, which revealed a similar trend but greater reduction in bacterial growth and symptom development (Fig. 5). Overall, the differences in mean CFU/cm^2 between wildtype and mutant strains were the most pronounced at eight days post inoculation (0.81 log units). While the growth of the AL65 ($TBDR_{1734}$) wildtype complement strain did not significantly differ from that of the AL65 ($TBDR_{2937}$) allele exchange mutant at this time point, an upward trend in growth and

248 significant differences in disease severity were observed, suggesting a partial recovery of the
249 wildtype phenotype.

250 The function of the pepper-associated TBDR was also assessed in the background of non-
251 pathogenic strain Xp5-6 through introduction of a frameshift mutation in the coding sequence.
252 This analysis revealed no discernable differences in symptom development or bacterial growth
253 between the wildtype and mutant (Supplementary Fig. 2). Likewise, we also investigated the
254 pepper associated variants found in *gltP* and its promoter region through topoisomerase insertion
255 of the endogenous coding sequence in the AL65 and Xp5-6 TBDR mutants, followed by trans-
256 complementation with the alternative *gltP* allele among the two strains. Initial characterization of
257 the resulting transformants revealed that the pUFR047-*gltP* construct was not stable in the AL65
258 ($TBDR_{2937}$) allele exchange mutant, which displayed an approximately 6.0% plating efficiency on
259 media amended with gentamicin (data not shown). While the same vector was found to be stable
260 in the Xp5-6ΔTBDR mutant (~95% plating efficiency), no significant differences in bacterial
261 growth or symptom development were observed when complemented with the *gltP* allele derived
262 from strain AL65 (Supplementary Fig. 2).

263 **_In-vitro_ growth experiments indicate that the pepper apoplasm is not a limiting factor of non-**
264 **pathogenic _X. perforans_ and that the pathogen response to polygalacturonic acid is associated**
265 **with pathogenicity phenotype.**

266 As the association analyses described above suggested that nutrient acquisition/transport may play
267 a role in determining the host range of *X. perforans*, we extracted apoplastic wash fluid from
268 pepper leaves to examine whether this environment might be a limiting factor for non-pathogenic
269 *X. perforans* strains. No significant differences in the growth of pepper-pathogenic strain AL65
270 and non-pathogenic strain 91-118 were observed in these experiments (Supplementary Fig. 3).

271 Furthermore, since the growth of these strains in apoplastic wash fluid was like that of XVM2
272 medium, the latter was utilized for further *in-vitro* growth experiments.

273 Previously, Blanvillain et al. (2007) reported that a homologue (XCC1749) of the pepper-
274 associated TBDR carried by *X. campestris* pv. *campestris* was induced in the presence of
275 polygalacturonic acid (PGA) and arabinose. Therefore, we assessed the growth of strain AL65 and
276 its derivatives, as well as several wild-type strains of differing genetic backgrounds and host-range
277 in standard XVM2 (which contains sucrose and fructose as the sole carbon source) and modified
278 XVM2 medium (mXVM2) amended with 0.25% PGA in place of sucrose and fructose (Fig. 6). A
279 mixed model analysis of variance with repeated measures confirmed that the TBDR allele did not
280 affect the growth of strain AL65 in either medium, with probabilities of 0.7195 and 0.8645 for
281 XVM2 and mXVM2, respectively. Similar results were obtained in modified XVM2 amended
282 with arabinose, xylose, and several other carbon sources (Supplementary Fig. 4).

283 When examining the wildtype strains, the overall growth of non-pathogenic strains Xp4B and
284 Xp5-6 was significantly lower than that of the others in XVM2 ($F = 7.66$, $p = 0.0019$). However,
285 no differences were observed over time ($F = 1.10$, $p = 0.3888$), with all strains generally exhibiting
286 the same growth trend. In contrast, the interaction between strain and time was significant in
287 mXVM2 ($F = 4.23$, $p < 0.0001$), where linear contrasts confirmed that the mean optical densities
288 of the three pepper-pathogenic strains were significantly greater than that of the three non-
289 pathogenic strains at the 38, 48, and 62 h time points ($p < 0.01$). These differences were more
290 pronounced for strains 91-118 and Xp5-6; however, all three non-pathogenic generally displayed
291 an approximately 24 h delay in entering the exponential growth phase in relation to the three
292 pepper-pathogenic strains (Fig. 6). Finally, we also investigated the growth of the wildtype strains

293 in XVM2 media under acidic and hypertonic conditions and did not see a similar correlation
294 (Supplementary Fig. 4).

295 **Discussion**

296 An understanding of the underlying factors that determine the host-specificity of plant pathogenic
297 bacteria remains a complex and open question, especially in the case of *Xanthomonas* spp., for
298 which abundant evidence of host adaptation is available. Here, we attempted to uncover new
299 insights into this question using genome-wide association (GWAS) to probe the host-specificity
300 determinants of *X. perforans* and related pathogens found within the *X. euvesicatoria* species
301 complex, which differ in their ability to infect pepper leaves independent of an effector-triggered
302 immunity (ETI; Fig. 1). Strikingly, we identified two novel gene-candidates that were significantly
303 associated with pepper-pathogenic *X. perforans* and *X. euvesicatoria*. Given that *X. perforans* was
304 previously reported to display an open pan-genome (Timilsina et al. 2019; Jibrin et al. 2018), it
305 was surprising to find that both gene-candidates were conserved among most members of the
306 genus and were predicted to play a role in the transport/acquisition of nutrients commonly
307 encountered in the plant-cell wall or apoplast (Fatima and Senthil-Kumar 2015).

308 One of these pepper-associated genes encoded for a TonB-dependent receptor (TBDR; Table 1),
309 which is a class of outer membrane transporter that spans the bacterial cell-wall and binds with
310 high affinity to a variety of molecules that are poorly permeable through porin channels or are
311 environmentally scarce (Schauer et al. 2008). A multiple sequence alignment of the TBDR in
312 question revealed that its significant association was an artifact of genome annotation, where
313 pepper-pathogenic *X. perforans* and *X. euvesicatoria* carried a pseudogene (divided into two
314 predicted coding sequences), while non-pathogenic lineages possessed an intact copy of the coding
315 sequence (Fig. 2). It is worth noting that because this TBDR was disrupted through independent

316 mutations within the *X. perforans* lineage, its significant association was only uncovered through
317 association tests of pan-genome orthologous groups, rather than single nucleotide polymorphisms
318 (SNPs) or unitigs, which are often considered more sensitive variants to employ in association
319 testing. This observation highlights the necessity of utilizing multiple strategies when conducting
320 a GWAS analysis.

321 A distinctive feature of the *Xanthomonas* genus is the high number of TBDR encoding genes found
322 within a given genome, which ranges from approximately 30 to 70 among various species and are
323 distributed across loci that serve diverse cellular functions (Blanvillain et al. 2007; Fatima and
324 Senthil-Kumar 2015). Likewise, the relative importance of a given TBDR in influencing the
325 pathogenesis of xanthomonads is variable and likely dependent upon genetic background
326 (Blanvillain et al. 2007). Therefore, we generated allele exchange mutants and replaced the
327 disrupted TBDR coding sequence carried by pepper-pathogenic strain AL65 with the intact allele
328 derived from strain Xp5-6. This analysis revealed that expression of the TBDR was associated
329 with attenuated symptom development and a minor reduction of bacterial growth when infiltrated
330 into pepper leaves, relative to the near-isogenic wildtype strain, while no discernable differences
331 were observed in tomato leaves (Fig. 4). Dip inoculation experiments revealed a more pronounced
332 reduction in growth of the allele exchange mutant, indicating that the effect of this gene towards
333 disease severity on pepper is not just limited to apoplastic colonization, but also epiphytic
334 colonization (Fig. 5). In contrast, when an early stop codon was introduced into the functional
335 TBDR gene of pepper-non-pathogenic strain Xp5-6, no benefit in bacterial growth or symptom
336 development was observed (Supplementary Fig. 2). Together, these data suggest that the pepper
337 associated TBDR is likely functionally redundant and while its pseudogenization may represent a

338 step in the process leading to the adaptation of *X. perforans* to *Capsicum*, other factors are clearly
339 involved.

340 Nonetheless, there are several mechanisms in which this class of transporter may be implicated in
341 the plant-pathogen interaction and contribute to the attenuated symptom development observed in
342 the pathogenic *X. perforans* background. For example, the C-terminal, membrane-spanning barrel
343 domain of the protein has the potential to serve as a pathogen-associated molecular pattern and
344 elicit the innate plant-defense response (Schauer et al. 2008). Alternatively, we noted that the gene
345 was located in a previously described carbohydrate utilization locus (Déjean et al. 2013;
346 Blanvillain et al. 2007) and was flanked on either side by operons involved in the metabolism of
347 galactose or xylan/xylose, as well as *gamR* (Fig. 2), which is a LysR-type regulator of galactose
348 metabolism that is involved in the regulation of *hrp* gene expression in a galactose independent
349 manner (Rashid et al. 2016). Indeed, expression analysis of a homologous TBDR (XCC1749) from
350 *X. campestris* pv. *campestris* indicated that the gene was induced in the presence of
351 polygalacturonic acid and arabinose, suggesting that the membrane receptor may be involved in
352 the transport of oligosaccharides derived from pectin or hemicellulose (Blanvillain et al. 2007),
353 two major components of the plant cell-wall.

354 To investigate this possibility in *X. perforans*, we assessed the growth of strains in an apoplast
355 mimicking medium (modified XVM2) amended with polygalacturonic acid as the sole carbon
356 source. While no significant differences in the growth were noted of the AL65 ($TBDR_{2937}$) allele
357 exchange mutant and wildtype strain, we observed an interesting correlation between the lag
358 growth period of several wildtype strains of differing genetic backgrounds and ability to infect
359 pepper leaves. Specifically, non-pathogenic strains exhibited an approximately 24 h longer lag
360 growth period in relation to pepper-pathogenic strains, while no significant differences in growth

over time between the two populations were observed in standard XVM2 medium (Fig. 6). Although the specific substrate of the pepper-associated TBDR remains unclear, these results point towards the pathogen-specific response to polygalacturonides as playing a role in mediating the adaptation of *X. perforans* to the pepper niche. Further supporting this hypothesis, we previously observed that other pathways involved in the secretion of plant-cell wall degrading enzymes and the recognition of pectin-derived oligosaccharides, including the core components of TonB transduction system (Vorhölter et al. 2012), were frequently recombining within the *X. perforans* lineage (Newberry et al. 2019), suggesting their ecological significance.

Association tests based on core-genome SNPs and genome-wide unitigs revealed that the proton/glutamate-aspartate symporter, *gltP,* was enriched with pepper-associated SNPs and insertion/deletion sites near the start-codon and promoter region and thus, may have the potential to influence its expression (Fig. 3C). This gene is highly conserved among the Proteobacteria and is essential for uptake and chemotaxis of bacterial cells to glutamate (Jacobs et al. 1995), which is an abundant nutrient in the tomato apoplast (Rico and Preston 2008). Furthermore, glutamate metabolism is essential for *Xanthomonas oryzae* pv. *oryzae* virulence in rice, suggesting that the availability of this amino acid in rice xylem vessels is a limiting factor in the absence of the glutamate synthase pathway (Pandey et al. 2014). As little data is available regarding the nutrient composition of the pepper apoplasm, we assessed the growth of strains in apoplastic wash fluid extracted from pepper leaves and found no evidence to indicate that this environment might be a limiting factor of non-pathogenic *X. perforans* (Supplementary Fig. 3). Glutamate is also an osmolyte that contributes to cellular homeostasis in Gram-negative bacteria and is thought to play an important role in the niche adaptation of several plant pathogenic bacterial species such as *Erwinia chrysanthemi* (Gouesbet et al. 1995), *Burkholderia glumae* (Kang and Hwang 2018), and

384 *Pseudomonas savastanoi* (Matas et al. 2012). Therefore, we investigated the effect of the pepper-
385 associated mutations in *gltP* and its promoter region *in-planta* and found no evidence that they
386 provided an adaptive advantage to non-pathogenic *X. perforans* strain Xp5-6 (Supplementary Fig.
387 2). While we were unable to assess these mutations in the pepper-pathogenic background of strain
388 AL65, which did not provide a stable host to our pUFR047-*gltP* construct, we noted that a
389 topoisomerase insertion in the *gltP* coding sequence of this strain incurred no discernable effect
390 on bacterial growth or symptom development (data not shown).

391 Over the course of this study, we encountered several limitations that may hinder the application
392 of similar GWAS analyses to *Xanthomonas* spp. For example, while formally recognized as
393 discrete bacterial species, *X. perforans* and *X. euvesicatoria* each displayed a genetically
394 monomorphic population structure that was more consistent with that of a clonal complex (Fig. 1).
395 As homologous recombination originating from primarily a single donor lineage (*X. euvesicatoria*)
396 was the principle factor driving the diversification of *X. perforans* (Jibrin et al. 2018; Timilsina et
397 al. 2019), this had the unfortunate side-effect of limiting the gene pool available to test for
398 associations (and the resulting power of statistical inference) and likely masked signatures of
399 shared selection pressure associated with host-adaptation (Skwark et al. 2017), despite
400 implementing appropriate controls for population structure. Moreover, the scarcity of strains that
401 did not exhibit evidence of a gene-for gene interaction in pepper leaves rendered our dataset biased
402 towards pathogenicity (Fig. 1C). While it might be possible to overcome this limitation through
403 more extensive sampling of *X. perforans* and *X. euvesicatoria* in an unbiased manner, variability
404 in the amenability of *X. perforans* to genetic manipulations would likely render this approach
405 unsuccessful as there is no correlation between host-range and avirulence gene content, after the
406 activity of said avirulence gene has been abolished through mutagenesis (Schwartz et al. 2015).

Indeed, the two gene candidates identified in this study were each located in a previously described recombination hot spot (Newberry et al. 2019), with the pepper-pathogenic *X. perforans* strains of clusters SC3 and SC4 carrying alleles nearly identical to that found in *X. euvesicatoria,* while the SC6 alleles both displayed distinct evolutionary histories (Fig. 2 and Supplementary Fig. 6). Likewise, other candidates that displayed marginal associations in the SNP-based analysis (Fig. 3A) also mapped to presumed recombination events in which primarily a single allele, ancestral to the *X. euvesicatoria* lineage was identified. Overall, these dynamics raise the likelihood that multi-locus epistatic interactions drive the adaptative evolution of *X. perforans* (Arnold et al. 2018). As such, analysis of genome-wide patterns of co-evolution may serve as a fruitful alternative to reconstructing the evolutionary events leading to the host-range expansion of this pathogen (Skwark et al. 2017). Alternatively, it is possible that virulence to pepper leaves has evolved independently within the bacterial species and would need to be considered for each lineage separately. Regardless, a more detailed understanding of the non-pathogenic *X. perforans* interaction with pepper plants will be crucial to provide clues as to the specific functional pathways with biological relevance in mediating this interaction.

Materials and Methods

Bacterial strain collection. Seventy eight *Xanthomonas perforans* strains and corresponding genome sequences described in several previous studies were analyzed here (Timilsina et al. 2019; Newberry et al. 2019; Abrahamian et al. 2019; Schwartz et al. 2015; Potnis et al. 2011). These strains were isolated from symptomatic tomato (n = 73) and pepper (n = 5) leaves between 1998 and 2017 and represent the diversity of the bacterial species documented in the Southeast United States. Over the course of this study, strains were routinely cultured on nutrient agar (NA) plates at 28℃ and kept at -80℃ in a sterile 30% glycerol solution for long term storage. The National

Center for Biotechnology (NCBI) GenBank accession numbers and associated metadata for the xanthomonads analyzed here are provided in Supplementary Table 2.

Pathogenicity phenotyping. *Xanthomonas perforans* strains were initially screened for the ability to induce a hypersensitive response (HR) in the susceptible pepper cv. Early CalWonder (ECW). Pepper seeds were sown in individual pots filled with commercial potting soil amended with Osmocote (The Scotts Company, Marysville, OH) and grown under greenhouse conditions with a mean temperature of 28 ± 5℃ and relative humidity between 70 and 90%. Using overnight cultures, a bacterial suspension was adjusted to a concentration of 5 x 10^8 CFU ml^{-1} spectrophotometrically (OD_{600nm} = 0.3) in a sterile $MgSO_4$•$7H_2O$ (50 mM) solution. The resulting cell suspension was then infiltrated into the leaves of pepper plants at the four- to five-week-old stage using a needle-less syringe. The appearance of collapsed, necrotic tissue at the site of inoculation after 24 to 48 h was considered a positive result, while leaves infiltrated with a sterile $MgSO_4$•$7H_2O$ solution served as a negative control. All strains that did not elicit a HR were then subjected to further pathogenicity analysis. For these assays, plants were prepared as described above and transferred to a Percival I-36VL growth chamber (Percival Scientific, Perry, IA), set at 25℃ with a 12 h photoperiod, 24 h prior to inoculation. The abaxial side of leaves were infiltrated (two to three strains per leaf) using a bacterial suspension diluted to 10^3 CFU ml^{-1} with 10-fold serial dilutions. Strains were qualitatively scored for pathogenicity based on the development of circular, water-soaked lesions, approximately 10 to 14 days post inoculation. In addition to $MgSO_4$•$7H_2O$, three strains (Xp5-6, Xp4BΔ*avrBsT*, 91-118Δ*avrXv3*) described as being HR negative and non-pathogenic to pepper plants by Schwartz et al. (2015) were included as negative controls. It is important to note that these strains occasionally induced resistant lesions that could be confused with a susceptible reaction. These generally appeared as necrotic specks, were not

accompanied by water-soaking, and their development was variable among individual leaves (Supplementary Fig. 7). Therefore, a strain was scored as pathogenic if all three replicates developed water-soaked lesions in a at least two independent experiments.

Avirulence gene prediction. To corroborate the results of the pathogenicity phenotyping, the *X. perforans* genome assemblies were screened for the presence of three avirulence genes known to elicit an effector triggered immunity (ETI) in the susceptible pepper cv. ECW using BLASTn (Evalue < $10e^{-5}$). These included *avrBsT*, *avrXv3,* and *xopJ6* (Minsavage 1990; Astua-Monge et al. 2000; Iruegas-Bocardo et al. 2018), using the following GenBank accession numbers for refence sequences respectively: AF156163.1, AF190120.1, and XFF6992_110003. As the activity of *avrXv3* was reported to be disrupted through frameshift mutation or transposon insertion in many *X. perforans* strains (Timilsina et al. 2016), alignments with the curated reference sequence were used to assess for the presence of an early stop-codon or other putative disruption in the coding sequence of the gene.

Phylogenetic reconstruction. A core-genome phylogeny was constructed utilizing the genomes of 78 *X. perforans* strains analyzed here along two additional *X. perforans*, 43 pepper-pathogenic *X. euvesiatoria*, and the genomes of 10 related *X. euvesiatoria* pathovars available from NCBI Genbank (Supplementary Table 2) using the program Parsnp (v1.2; Treangen et al. 2014). Parsnp was run under the default settings using the completed genome of *X. perforans* strain LH3 as the alignment reference. High-quality, core genome single-nucleotide polymorphisms (SNPs) were extracted from the alignment using the Harvest Tools suite that accompanies Parsnp and used to construct a Maximum Likelihood phylogenetic tree with iQTree (v.1.6.4; Nguyen et al. 2015) using the Generalized Time-Reversible nucleotide substitution model (Nei and Kumar 2000). This concatenated, core-SNP alignment was also used as input for analysis with Fastbaps software

476 (Tonkin-Hill et al. 2019) with the optimized prior set to "type = hc" to ensure that the inferred
477 population structure was consistent with that described in previous studies. For individual gene
478 phylogenies, BLASTn was used the extract the pepper associated TonB-dependent receptor (locus
479 tag IS_RS10330) and *gltP* (locus tag IS_RS18650) sequences from the genome assemblies.
480 Multiple sequence alignments were constructed for each gene individually with the webPRANK
481 server (Löytynoja and Goldman 2010) and used to infer Neighbor-Joining phylogenetic trees
482 (Saitou and Nei 1987) based on the Jukes-Cantor model of nucleotide substitution (Jukes and
483 Cantor 1969) with MEGA7 software (Kumar et al. 2016). Phylogenetic trees were annotated using
484 the Interactive Tree of Life (iTOL) web server (Letunic and Bork 2007).

485 **Genome-wide association analyses.** We first tested for associations between the
486 presence/absence of pan-genome orthologous groups and core genome SNPs using the program
487 Scoary (Brynildsrud et al. 2016). For this analysis, the *X. euvesicatoria* species complex pan-
488 genome was estimated using the program Roary (v.3.8.2; Page et al. 2015) under the default
489 settings with the same 133 genomes utilized for phylogenetic analysis as input. Likewise, SNPs
490 were extracted from the *X. euvesicatoria* species complex core genome using the Harvest Tools
491 suite, as described above. The input files for Scoary consisted of a comma-separated list of
492 bacterial strains in which the phenotype was coded as a binary variable as well as a variant file,
493 which consisted of either a pan-genome presence/absence matrix or a variant call file (VCF) with
494 SNP data. Finally, the maximum likelihood phylogenetic tree constructed from core-genome SNPs
495 was used to control for population structure. A third association analysis was also conducted the
496 using the DBGWAS pipeline (Jaillard et al. 2018) with the minor allele frequency set to 0.15 using
497 the "-maf" tag. This approach tests for associations between phenotype and assembled kmers
498 (termed unitigs) using the bugwas method, which employs the GEMMA library (Zhou and

499 Stephens 2012) to implement a linear mixed model while controlling for population structure by
500 estimating the effect of lineage on a given phenotype (Earle et al. 2016).

501 Initial association tests were carried out using a conservative dataset (n = 71). For this, we only
502 considered *X. perforans* strains for which pathogenicity independent of an ETI could be
503 experimentally verified and pepper-pathogenic *X. euvesicatoria* strains that were either isolated
504 from naturally infected pepper leaves and/or were previously tested for pathogenicity (Schwartz
505 et al. 2015; Roach et al. 2019). Strains that did not meet these criteria were considered missing
506 data and scored as 'NA' in the phenotype file. However, due to the limited power to test for
507 associations with this smaller sample size, we also performed exploratory association analyses
508 utilizing the entire collection of genomes. This analysis was based on the hypothesis that host
509 range is associated with genetic background within the *X. perforans* lineage, while related *X.*
510 *euvesicatoria* pathovars are non-pathogenic to pepper plants, independent of an ETI. For this, we
511 attempted to generate deletion mutant constructs of *avrXv3* and *xopJ6* to assess the pathogenicity
512 of *X. perforans* strains from cluster SC5 but found that these strains were not amenable to genetic
513 manipulations such as transformation via electroporation or conjugation and were therefore,
514 considered missing data, leaving the final sample size at 131 genomes. Finally, variants with a p
515 < 0.01 after Bonferroni correction (or Cochran's Q statistic in the case of unitigs) and a diagnostic
516 accuracy of $\geq 90\%$ for orthologous groups or 98% for SNPs and unitigs were considered significant
517 in these analyses.

518 **Generation of allele exchange mutants.** To assess the function of the pepper associated TonB-
519 dependent receptor (TBDR), allele exchange mutants were generated using Gateway cloning
520 technology (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. For these
521 experiments, the TBDR coding sequence and endogenous promoter region (approximately 440 bp

upstream of the coding sequence) carried by pepper non-pathogenic *X. perforans* strain, Xp5-6, was replaced via double-crossover homologous recombination with the coding sequence derived from the pathogenic strain GEV2393 (1,166 bp), which contains a disruptive early stop codon. Likewise, the disrupted TBDR alleles carried by strains GEV2393 and AL65 were replaced with the intact coding sequence (2,937 bp) derived from strain Xp5-6. To serve as a control in functional characterization, the AL65 ($TBDR_{2937}$) allele exchange mutant was complemented back with the original, truncated allele (1,734 bp), derived from the wildtype strain as described above.

We also attempted to use a similar approach assess the pepper-associated variants found in *gltP* and its promoter region; however, were not successful in obtaining double-crossover mutants after several months of sub-culturing. We therefore employed an alternative approach, which involved the generation of *gltP* knockout mutants in the AL65 ($TBDR_{2937}$) and the Xp5-6ΔTBDR backgrounds using a topoisomerase insertion targeting the middle of the *gltP* coding sequence (positions 236-741). The resulting double *gltP* and TBDR allele exchange mutants were then complemented with an intact *gltP* coding sequence and endogenous promoter region (approximately 500 bp intergenic region upstream of the coding sequence) derived from the alternative strain, using a pUFR047-*gltP* construct. The methodologies and cloning procedures employed in these experiments are detailed in the supplementary materials along with a complete list of plasmids and bacterial strains, presented in Supplementary Table 1.

***In-planta* growth assays.** To assess the population dynamics of the *X. perforans* mutants generated in this study, plants were incubated under growth chamber conditions with the temperature set at 28℃ and a 12 h photoperiod. The middle leaves of pepper (two strains per leaf) and tomato (one strain per leaf) plants were syringe-infiltrated with bacterial suspensions adjusted to 10^4 CFU ml^{-1} as described above. Leaves were harvested from plants (n = 3 per treatment) at 0

and 4 days post-inoculation (DPI) for strain AL65 and its derivatives, while strain Xp5-6 and its derivatives were sampled at 0, 4, and 8 DPI. Three leaf-discs per experimental replicate were excised from the inoculated tissue using a flame-sterilized cork-borer (1 cm^2), then macerated in 1 ml of a sterile $MgSO_4 \bullet 7H_2O$ solution. Ten-fold serial dilutions were conducted to enumerate the CFUs per cm^2 of tissue after 48 to 72 h of incubation on NA amended with either rifamycin (50 $\mu g\ ml^{-1}$) or gentamicin (10 $\mu g\ ml^{-1}$), where appropriate. These data were log transformed to obtain a normal distribution and subjected to a mixed model analysis of variance for each host individually. The random effects in this model included the experiment and residual error term, while fixed effects included strain.

Dip Inoculation Experiments. The pathogenicity of *X. perforans* AL65 and its derivatives was further assessed using a dip inoculation technique. Briefly, bacterial suspensions were adjusted to 10^7 CFU ml^{-1} in a $MgSO_4 \bullet 7H_2O$ solution amended with 0.0025% (vol/vol) Silwet 77. Four- to five-week-old pepper plants (cv. ECW) were inoculated by inverting pots and submerging plants into bacterial inoculum for 30 sec so that all leaves were covered. After drying, plants were then placed into plastic bins for 48 hrs. to maintain high humidity during disease establishment, after which they were removed from the bins and maintained in a greenhouse for the duration of the experiment. Each treatment contained three replicates and the experiment was repeated once. Plants dipped in a sterile $MgSO_4 \bullet 7H_2O$ solution amended with Silwet 77 served as a negative control. Bacterial populations were enumerated as described above at 0, 4, 8, and 12 DPI. Similarly, disease severity ratings were assigned using a modified Horsfall-Barratt scale where 0 = no symptoms, 1 = 1 to 3, 2 = 3 to 6, 3 = 6 to 12, 4 = 12 to 25, 5 = 25 to 50, 6 = 50 to 75, 7 = 75 to 87, and 8 = 87 to 100% symptomatic (Horsfall and Barratt 1945). The log_{10} transformed

population data and raw disease severity scores and subjected to a mixed model analysis of variance as described above.

Growth experiments in apoplastic wash fluid and XVM2 media. To extract apoplastic was fluid (AWF) from pepper leaves for the growth assays, we used the infiltration-centrifugation technique described by O'Leary et al. (2014). Likewise, XVM2 and modified XVM2 (mXVM2), amended with either 0.25% polygalacturonic acid or 10 mM of various sugars (xylose, arabinose, maltose, glucose) as the sole carbon source, were prepared as described previously (Jiang et al. 2014). To inoculate the media, bacterial strains were cultured overnight on NA plates, suspended in sterile deionized (sDI) H_2O using a pipette tip, centrifuged briefly, and washed twice in sDI H_2O. Cell suspensions were adjusted to 10^6 CFU ml^{-1} using ten-fold serial dilutions and then 50 µl of the resulting cell suspension was added to an Eppendorf tube containing 1 ml of medium or AWF. Media inoculated with 50 µl of sDI H_2O served as a negative control in these experiments. The inoculated medium was then vortexed briefly and 200 µl aliquots were transferred to a 96 well plate and incubated at 28℃ with 250 rpm shaking. Plates were read at 600nm every 10 to 14 h using a UV-Vis spectrophotometer (GeneSys 180; Daly City, CA). To test for differences in the growth of strains in XVM2 and mXVM2 media, the optical density values were subjected to a square root transformation to obtain a normal distribution and used to conduct a repeated-measures mixed model analysis of variance for each medium individually. When significant differences were observed, a linear contrasts were used to test for overall differences between pepper-pathogenic and non-pathogenic strains within each time-point ($p < 0.05$). All statistical analyses described above were performed using JMP Pro 13 software (SAS Institute, Cary, NC).

Acknowledgments

We thank Dr. David Young and the staff at the Alabama Supercomputer Authority for providing the computational resources necessary to conduct this work.

References

Abrahamian, P., Timilsina, S., Minsavage, G. V., Potnis, N., Jones, J. B., Goss, E. M., et al. 2019. Molecular Epidemiology of *Xanthomonas perforans* Outbreaks in Tomato Plants from Transplant to Field as Determined by Single-Nucleotide Polymorphism Analysis. Appl. Environ. Microbiol. 85.

An, S.-Q., Potnis, N., Dow, M., Vorhölter, F.-J., He, Y.-Q., Becker, A., et al. 2019. Mechanistic insights into host adaptation, virulence and epidemiology of the phytopathogen *Xanthomonas*. FEMS Microbiology Reviews.

Areas, M. S., Gonçalves, R. M., Soman, J. M., Sakate, R. K., Gioria, R., Júnior, T. A. F. da S., et al. 2015. Prevalence of *Xanthomonas euvesicatoria* on Pepper in Brazil. Journal of Phytopathology. 163:1050–1054.

Aritua, V., Harrison, J., Sapp, M., Buruchara, R., Smith, J., and Studholme, D. J. 2015. Genome sequencing reveals a new lineage associated with lablab bean and genetic exchange between *Xanthomonas axonopodis* pv. *phaseoli* and *Xanthomonas fuscans* subsp. *fuscans*. Front. Microbiol. 6.

Astua-Monge, G., Minsavage, G. V., Stall, R. E., Davis, M. J., Bonas, U., and Jones, J. B. 2000. Resistance of Tomato and Pepper to T3 Strains of *Xanthomonas campestris* pv. *vesicatoria* Is Specified by a Plant-Inducible Avirulence Gene. MPMI. 13:911–921.

Bansal, K., Kumar, S., and Patil, P. B. 2018. Complete Genome Sequence Reveals Evolutionary Dynamics of an Emerging and Variant Pathovar of *Xanthomonas euvesicatoria*. Genome Biol Evol. 10:3104–3109.

Barak, J. D., Vancheva, T., Lefeuvre, P., Jones, J. B., Timilsina, S., Minsavage, G. V., et al. 2016. Whole-Genome Sequences of *Xanthomonas euvesicatoria* Strains Clarify Taxonomy and Reveal a Stepwise Erosion of Type 3 Effectors. Front. Plant Sci. 7.

Berthenet, E., Yahara, K., Thorell, K., Pascoe, B., Meric, G., Mikhail, J. M., et al. 2018. A GWAS on *Helicobacter pylori* strains points to genetic variants associated with gastric cancer risk. BMC Biology. 16:84.

Blanvillain, S., Meyer, D., Boulanger, A., Lautier, M., Guynet, C., Denancé, N., et al. 2007. Plant Carbohydrate Scavenging through TonB-Dependent Receptors: A Feature Shared by Phytopathogenic and Aquatic Bacteria. PLOS ONE. 2:e224.

Brynildsrud, O., Bohlin, J., Scheffer, L., and Eldholm, V. 2016. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. Genome Biol. 17.

623 Constantin, E. C., Cleenwerck, I., Maes, M., Baeyen, S., Malderghem, C. V., Vos, P. D., et al.
624 2016. Genetic characterization of strains named as *Xanthomonas axonopodis* pv. *dieffenbachiae*
625 leads to a taxonomic revision of the X. axonopodis species complex. Plant Pathology. 65:792–
626 806.

627 Déjean, G., Blanvillain-Baufumé, S., Boulanger, A., Darrasse, A., Bernonville, T. D. de, Girard,
628 A.-L., et al. 2013. The xylan utilization system of the plant pathogen *Xanthomonas campestris*
629 pv *campestris* controls epiphytic life and reveals common features with oligotrophic bacteria and
630 animal gut symbionts. New Phytologist. 198:899–915.

631 Dhakal, U., Dobhal, S., Alvarez, A. M., and Arif, M. 2019. Phylogenetic Analyses of
632 Xanthomonads Causing Bacterial Leaf Spot of Tomato and Pepper: *Xanthomonas euvesicatoria*
633 Revealed Homologous Populations Despite Distant Geographical Distribution. Microorganisms.
634 7.

635 Earle, S. G., Wu, C.-H., Charlesworth, J., Stoesser, N., Gordon, N. C., Walker, T. M., et al. 2016.
636 Identifying lineage effects when controlling for population structure improves power in bacterial
637 association studies. Nat Microbiol. 1:16041.

638 Fatima, U., and Senthil-Kumar, M. 2015. Plant and pathogen nutrient acquisition strategies.
639 Front. Plant Sci. 6.

640 Feng, T., Yan, K.-P., Mikkelsen, M. D., Meyer, A. S., Schols, H. A., Westereng, B., et al. 2014.
641 Characterisation of a novel endo-xyloglucanase (XcXGHA) from *Xanthomonas* that
642 accommodates a xylosyl-substituted glucose at subsite −1. Appl Microbiol Biotechnol. 98:9667–
643 9679.

644 Ferreira, M. A. S. V., Bonneau, S., Briand, M., Cesbron, S., Portier, P., Darrasse, A., et al. 2019.
645 *Xanthomonas citri* pv. *viticola* Affecting Grapevine in Brazil: Emergence of a Successful
646 Monomorphic Pathogen. Front. Plant Sci. 10.

647 Gouesbet, G., Jebbar, M., Bonnassie, S., Hugouvieux-Cotte-Pattat, N., Himdi-Kabbab, S., and
648 Blanco, C. 1995. Erwinia chrysanthemi at high osmolarity: influence of osmoprotectants on
649 growth and pectate lyase production. Microbiology. 141:1407–1412.

650 Hajri, A., Brin, C., Hunault, G., Lardeux, F., Lemaire, C., Manceau, C., et al. 2009. A
651 «Repertoire for Repertoire» Hypothesis: Repertoires of Type Three Effectors are Candidate
652 Determinants of Host Specificity in *Xanthomonas*. PLOS ONE. 4:e6632.

653 Hayward, A. C. 1993. The hosts of *Xanthomonas*. In *Xanthomonas*, eds. J. G. Swings and E. L.
654 Civerolo. Dordrecht: Springer Netherlands, p. 1–119.

655 Hicks, N. D., Carey, A. F., Yang, J., Zhao, Y., and Fortune, S. M. 2019. Bacterial Genome-Wide
656 Association Identifies Novel Factors That Contribute to Ethionamide and Prothionamide
657 Susceptibility in Mycobacterium tuberculosis. mBio. 10.

Huang, C.-H., Vallad, G. E., Adkison, H., Summers, C., Margenthaler, E., Schneider, C., et al. 2013. A Novel *Xanthomonas* sp. Causes Bacterial Spot of Rose (*Rosa* spp.). Plant Disease. 97:1301–1307.

Huang, C.-L., Pu, P.-H., Huang, H.-J., Sung, H.-M., Liaw, H.-J., Chen, Y.-M., et al. 2015. Ecological genomics in *Xanthomonas*: the nature of genetic adaptation with homologous recombination and host shifts. BMC Genomics. 16:188.

Iruegas-Bocardo, F., Abrahamian, P., Minsavage, G. V., Potnis, N., Vallad, G. E., Jones, J. B., et al. 2018. XopJ6, a new member of the XopJ family of type III effectors, in *Xanthomonas perforans*. In Boston, USA.

Jacobs, M. H. J., Van Der Heide, T., Tolner, B., Driessen, A. J. M., and Konings, W. N. 1995. Expression of the *gltP* gene of *Escherichia coli* in a glutamate transport-deficient mutant of *Rhodobacter sphaeroides* restores chemotaxis to glutamate: Chemotaxis of glutamate in *Rhodobacter sphaeroides*. Molecular Microbiology. 18:641–647.

Jacques, M.-A., Arlat, M., Boulanger, A., Boureau, T., Carrère, S., Cesbron, S., et al. 2016. Using Ecology, Physiology, and Genomics to Understand Host Specificity in *Xanthomonas*. Annual Review of Phytopathology. 54:163–187.

Jaillard, M., Lima, L., Tournoud, M., Mahé, P., van Belkum, A., Lacroix, V., & Jacob, L. 2018. A fast and agnostic method for bacterial genome-wide association studies: Bridging the gap between k-mers and genetic events. PLoS genetics, 14(11), e1007758. https://doi.org/10.1371/journal.pgen.1007758

Jiang, G.-F., Jiang, B.-L., Yang, M., Liu, S., Liu, J., Liang, X.-X., et al. 2014. Establishment of an inducing medium for type III effector secretion in *Xanthomonas campestris* pv. *campestris*. Braz J Microbiol. 44:945–952.

Jibrin MO, Potnis N, Timilsina S, Minsavage GV, Vallad GE, Roberts PD, Jones JB, Goss EM. Genomic Inference of Recombination-Mediated Evolution in *Xanthomonas euvesicatoria* and *X. perforans*. Appl Environ Microbiol. 2018. 18:84(13):e00136-18. doi: 10.1128/AEM.00136-18.

Jones, J. B., Lacy, G. H., Bouzar, H., Stall, R. E., and Schaad, N. W. 2004. Reclassification of the Xanthomonads Associated with Bacterial Spot Disease of Tomato and Pepper. Systematic and Applied Microbiology. 27:755–762.

Jones, J. D. G., and Dangl, J. L. 2006. The plant immune system. Nature. 444:323–329.

Jukes, T. H., and Cantor, C. R. 1969. *Evolution of protein molecules*. New York: Academic Press.

Kang, Y., and Hwang, I. 2018. Glutamate uptake is important for osmoregulation and survival in the rice pathogen *Burkholderia glumae*. PLoS One. 13.

Kousik, C., and Ritchie, D. F. 1996. Race Shift in *Xanthomonas campestris* pv. *vesicatoria* Within a Season in Field-Grown Pepper. Phytopathology. 86:952.

694 Kumar, S., Stecher, G., and Tamura, K. 2016. MEGA7: Molecular Evolutionary Genetics
695 Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 33:1870–1874.

696 Lees, J. A., Vehkala, M., Välimäki, N., Harris, S. R., Chewapreecha, C., Croucher, N. J., et al.
697 2016. Sequence element enrichment analysis to determine the genetic basis of bacterial
698 phenotypes. Nature Communications. 7:1–8.

699 Letunic, I., and Bork, P. 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic
700 tree display and annotation. Bioinformatics. 23:127–128.

701 Lovell, H. C., Jackson, R. W., Mansfield, J. W., Godfrey, S. a. C., Hancock, J. T., Desikan, R., et
702 al. 2011. *In planta* conditions induce genomic changes in *Pseudomonas syringae* pv.
703 *phaseolicola*. Molecular Plant Pathology. 12:167–176.

704 Löytynoja, A., and Goldman, N. 2010. webPRANK: a phylogeny-aware multiple sequence
705 aligner with interactive alignment browser. BMC Bioinformatics. 11:579.

706 Matas, I. M., Lambertsen, L., Rodríguez-Moreno, L., and Ramos, C. 2012. Identification of
707 novel virulence genes and metabolic pathways required for full fitness of *Pseudomonas*
708 *savastanoi* pv. *savastanoi* in olive (*Olea europaea*) knots. New Phytologist. 196:1182–1196.

709 Méric, G., Mageiros, L., Pensar, J., Laabei, M., Yahara, K., Pascoe, B., et al. 2018. Disease-
710 associated genotypes of the commensal skin bacterium *Staphylococcus* epidermidis. Nature
711 Communications. 9:1–11.

712 Mhedbi-Hajri, N., Hajri, A., Boureau, T., Darrasse, A., Durand, K., Brin, C., et al. 2013.
713 Evolutionary History of the Plant Pathogenic Bacterium *Xanthomonas axonopodis*. PLoS One. 8.

714 Minsavage, G. V. 1990. Gene-For-Gene Relationships Specifying Disease Resistance in
715 *Xanthomonas campestris* pv. *vesicatoria* - Pepper Interactions. Molecular Plant-Microbe
716 Interactions. 3:41.

717 Nei, M., and Kumar, S. 2000. *Molecular Evolution and Phylogenetics*. Oxford, New York:
718 Oxford University Press.

719 Newberry, E. A., Bhandari, R., Minsavage, G. V., Timilsina, S., Jibrin, M. O., Kemble, J., et al.
720 2019. Independent Evolution with the Gene Flux Originating from Multiple *Xanthomonas*
721 Species Explains Genomic Heterogeneity in *Xanthomonas perforans*. Appl. Environ. Microbiol.
722 85.

723 O'Leary, B. M., Rico, A., McCraw, S., Fones, H. N., and Preston, G. M. 2014. The Infiltration-
724 centrifugation Technique for Extraction of Apoplastic Fluid from Plant Leaves Using Phaseolus
725 vulgaris as an Example. JoVE (Journal of Visualized Experiments). :e52113–e52113.

726 Osdaghi, E., Taghavi, S. M., Hamzehzarghani, H., and Lamichhane, J. R. 2016. Occurrence and
727 Characterization of the Bacterial Spot Pathogen *Xanthomonas euvesicatoria* on Pepper in Iran.
728 Journal of Phytopathology. 164:722–734.

729 Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. G., et al. 2015.
730 Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 31:3691–3693.

731 Pandey, A., Ray, S. K., Sonti, R. V., and Rajeshwari, R. 2014. *gltB*/*D* Mutants of *Xanthomonas*
732 oryzae pv. *oryzae* are Virulence Deficient. Curr. Microbiol. 68:105–112.

733 Parkinson, N., Cowie, C., Heeney, J., and Stead, D. 2009. Phylogenetic structure of
734 *Xanthomonas* determined by comparison of *gyrB* sequences. International Journal of Systematic
735 and Evolutionary Microbiology. 59:264–274.

736 Potnis, N., Krasileva, K., Chow, V., Almeida, N. F., Patil, P. B., Ryan, R. P., et al. 2011.
737 Comparative genomics reveals diversity among xanthomonads infecting tomato and pepper.
738 BMC Genomics. 12:146.

739 Potnis, N., Timilsina, S., Strayer, A., Shantharaj, D., Barak, J. D., Paret, M. L., et al. 2015.
740 Bacterial spot of tomato and pepper: diverse *Xanthomonas* species with a wide variety of
741 virulence factors posing a worldwide challenge. Molecular Plant Pathology. 16:907–920.

742 Rashid, M. M., Ikawa, Y., and Tsuge, S. 2016. GamR, the LysR-Type Galactose Metabolism
743 Regulator, Regulates *hrp* Gene Expression via Transcriptional Activation of Two Key *hrp*
744 Regulators, HrpG and HrpX, in *Xanthomonas oryzae* pv. *oryzae*. Appl. Environ. Microbiol.
745 82:3947–3958.

746 Rico, A., and Preston, G. M. 2008. *Pseudomonas syringae* pv. *tomato* DC3000 Uses Constitutive
747 and Apoplast-Induced Nutrient Assimilation Pathways to Catabolize Nutrients That Are
748 Abundant in the Tomato Apoplast. Molecular Plant-Microbe Interactions. 21:269–282.

749 Roach, R., Mann, R., Gambley, C. G., Chapman, T., Shivas, R. G., and Rodoni, B. 2019.
750 Genomic sequence analysis reveals diversity of Australian *Xanthomonas* species associated with
751 bacterial leaf spot of tomato, capsicum and chilli. BMC Genomics. 20:310.

752 Roach, R., Mann, R., Gambley, C. G., Shivas, R. G., and Rodoni, B. 2018. Identification of
753 *Xanthomonas* species associated with bacterial leaf spot of tomato, capsicum and chilli crops in
754 eastern Australia. Eur J Plant Pathol. 150:595–608.

755 Saitou, N., and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing
756 phylogenetic trees. Mol Biol Evol. 4:406–425.

757 Samanta, J. N., Mandal, K., and Maiti, S. 2013. A novel pathovar of *Xanthomonas axonopodis*
758 causes gumming of Guggal (*Commiphora wightii*). Eur J Plant Pathol. 135:115–125.

759 San, J. E., Baichoo, S., Kanzi, A., Moosa, Y., Lessells, R., Fonseca, V., et al. 2020. Current
760 Affairs of Microbial Genome-Wide Association Studies: Approaches, Bottlenecks and
761 Analytical Pitfalls. Front. Microbiol. 10.

762 Schauer, K., Rodionov, D. A., and de Reuse, H. 2008. New substrates for TonB-dependent
763 transport: do we only see the “tip of the iceberg”? Trends Biochem. Sci. 33:330–338.

764 Schwartz, A. R., Potnis, N., Timilsina, S., Wilson, M., Patané, J., Martins, J. J., et al. 2015.
765 Phylogenomics of *Xanthomonas* field strains infecting pepper and tomato reveals diversity in
766 effector repertoires and identifies determinants of host specificity. Front. Microbiol. 6.

767 Serrania, J., Vorhölter, F.-J., Niehaus, K., Pühler, A., and Becker, A. 2008. Identification of
768 *Xanthomonas campestris* pv. *campestris* galactose utilization genes from transcriptome data.
769 Journal of Biotechnology. 135:309–317.

770 Sheppard, S. K., Didelot, X., Meric, G., Torralbo, A., Jolley, K. A., Kelly, D. J., et al. 2013.
771 Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in
772 *Campylobacter*. PNAS. 110:11923–11927.

773 Skwark, M. J., Croucher, N. J., Puranen, S., Chewapreecha, C., Pesonen, M., Xu, Y. Y., et al.
774 2017. Interacting networks of resistance, virulence and core machinery genes identified by
775 genome-wide epistasis analysis. PLOS Genetics. 13:e1006508.

776 Stall, R. E., Jones, J. B., and Minsavage, G. V. 2009. Durability of Resistance in Tomato and
777 Pepper to Xanthomonads Causing Bacterial Spot. Annual Review of Phytopathology. 47:265–
778 284.

779 Suzuki, M., Shibayama, K., and Yahara, K. 2016. A genome-wide association study identifies a
780 horizontally transferred bacterial surface adhesin gene associated with antimicrobial resistant
781 strains. Scientific Reports. 6:1–10.

782 Swords, K. M., Dahlbeck, D., Kearney, B., Roy, M., and Staskawicz, B. J. 1996. Spontaneous
783 and induced mutations in a single open reading frame alter both virulence and avirulence in
784 *Xanthomonas campestris* pv. *vesicatoria avrBs2*. Journal of Bacteriology. 178:4661–4669.

785 Timilsina, S., Abrahamian, P., Potnis, N., Minsavage, G. V., White, F. F., Staskawicz, B. J., et al.
786 2016. Analysis of Sequenced Genomes of *Xanthomonas perforans* Identifies Candidate Targets
787 for Resistance Breeding in Tomato. Phytopathology. 106:1097–1104.

788 Timilsina, S., Pereira-Martin, J. A., Minsavage, G. V., Iruegas Bocardo, F., Abrahamian, P.,
789 Potnis, N., et al. 2019. Multiple recombination events drive the current genetic structure of
790 *Xanthomonas perforans* in Florida. Front. Microbiol. 10..

791 Tonkin-Hill, G., Lees, J. A., Bentley, S. D., Frost, S. D. W., and Corander, J. 2019. Fast
792 hierarchical Bayesian analysis of population structure. Nucleic Acids Res. 47:5539–5549.

793 Treangen, T. J., Ondov, B. D., Koren, S., and Phillippy, A. M. 2014. The Harvest suite for rapid
794 core-genome alignment and visualization of thousands of intraspecific microbial genomes.
795 Genome Biology. 15:524.

796 Vorhölter, F.-J., Wiggerich, H.-G., Scheidle, H., Mrozek, K., Pühler, A., and Niehaus, K. 2012.
797 Involvement of bacterial TonB-dependent signaling in the generation of an oligogalacturonide
798 damage-associated molecular pattern from plant cell walls exposed to *Xanthomonas campestris*
799 pv. *campestris* pectate lyases. BMC Microbiol. 12:239.

White, F. F., Potnis, N., Jones, J. B., and Koebnik, R. 2009. The type III effectors of *Xanthomonas*. Molecular Plant Pathology. 10:749–766.

Yaripour, Z., Mohsen Taghavi, S., Osdaghi, E., and Lamichhane, J. R. 2018. Host range and phylogenetic analysis of *Xanthomonas alfalfae* causing bacterial leaf spot of alfalfa in Iran. Eur J Plant Pathol. 150:267–274.

Zhou, X., and Stephens, M. 2012. Genome-wide Efficient Mixed Model Analysis for Association Studies. Nat Genet. 44:821–824.

Figure Legends

Figure 1. Population structure and pathogenicity analysis of *X. perforans*, *X. euvesicatoria,* and related pathogens. **A)** Hypersensitive response in pepper cv. ECW induced by *X. perforans* strains 91-118 and Xp4B, and lack of HR associated with mutagenesis of the avirulence genes *avrXv3* or *avrBsT* among the two strains, respectively. Leaves were photographed 24 h after inoculation with a bacterial suspension adjusted 10^8 CFU ml^{-1}. **B)** Pathogenicity phenotyping of *X. perforans*. Pepper leaves were infiltrated with a bacterial suspension adjusted to 10^3 CFU ml^{-1} and qualitatively scored for pathogenicity based on the development of circular, water-soaked lesions, approximately 10 to 14 days after inoculation. Representative pathogenic (right side of leaf) and non-pathogenic (left side of leaf) strains are shown, with the genetic group of the individual strain indicated in parentheses. Leaves were soaked in tap water for 20 min. prior to photographing to enhance water-soaking symptoms. **C)** Maximum likelihood phylogeny of the *X. euvesicatoria* species complex, which consists of *X. perforans* (*Xp*), *X. euvesicatoria* (*Xeu*), and related *Xeu* pathovars, based on 93,380 SNPs that were extracted from a core-genome alignment of 3.71 Mbp. The inner ring shows the distribution of avirulence gene content, the middle ring indicates the host of isolation, and the outer ring denotes the genetic groups inferred through first level of BAPS hierarchy (Tonkin-Hill et al. 2019). The sequence clusters (SCs) present within the *Xp* lineage, as described previously (Newberry et al. 2019), are labeled accordingly. Within the *Xp* lineage, pepper-pathogenic strains are highlighted in red, non-pathogenic strains in blue, and strains that induced a HR in pepper leaves are shown in black. Black dots indicate the strains included in the conservative dataset for association testing. A star indicates a strain for which mutational analysis of a given avirulence gene was used to verify pathogenicity by Schwartz et al. (2015), which were also included in the conservative dataset. The coloring of the branches shows the scoring of genetic

groups that were used to test for associations in the exploratory dataset, which was based on the hypothesis that host-range is associated with genetic background within the *Xp* lineage and that related *Xeu* pathovars are non-pathogenic to pepper plants, independent of an effector-triggered immunity. Groups that harbor pepper-pathogenic strains are shown in red, those with non-pathogenic strains in blue, and groups for which no phenotypic data are available in grey. The tree was drawn to scale and bootstrap values with < 95% support are shown at the nodes.

Figure 2. The coding sequence of a pepper-associated TonB-dependent receptor (TBDR) is disrupted through independent mutations among pepper-pathogenic *X. perforans* and *X. euvesicatoria*. **A)** Neighbor-joining phylogenetic tree (left) and multiple sequence alignment (right) of a TBDR associated with pepper-pathogenicity among members of the *X. euvesicatoria* species complex. The tree was constructed from the nucleotide sequences of 130 TBDRs extracted from *X. euvesicatora* and *X. perforans* genomes. The distribution of polymorphisms across the multiple sequence alignment is shaded within the grey boxes and yellow arrows below indicate the open reading frame of the sequence, as predicited with ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder/). Taxa that contain pepper-pathogenic *X. perforans and X. euvesicatoria* strains are highlited in bold with the number of identical alleles at each tip shown in parentheses. The tree was rooted on *X. citri* pv. citri strain 306 and branch support, shown at the nodes, was asessed using 1,000 bootstrap replicates. **B)** Genomic neighborhood of the pepper-associated TBDR (highlighted in red). Genes of the same color indicate orthologous groups. The operon involved in galactose metabolism as described by Serrania et al. (2008); regulator of galactose associated metabolism (*gamR*) as described by Rashid et al. (2016); xylan/xylose related metabolism (*xyl*) as described by Déjean et al. (2013) and Feng et al. (2014).

875 **Figure 3.** Associations between core-genome single nucleotide polymorphisms (SNPs) or
876 genome-wide unitigs and pepper-pathogenicity within the *X. euvesicatoria* species complex. **A)**
877 SNP-based associations derived from pairwise comparisons of probability after Bonferroni
878 correction using a dataset of *X. perforans* and *X. euvesicatoria* strains for which the phenotype
879 could be experimentally verified and/or were isolated from naturally infected pepper leaves (n =
880 71). Forty-five SNPs were significantly associated with pepper-pathogenicity ($p < 0.0032$, 100%
881 diagnostic accuracy) and mapped to four chromosomal regions encompassing nine coding
882 sequences. Genomic features located at pepper-associated variants are labeled accordingly; T2SS
883 – type 2 secretions system protein, T6SS – type 6 secretion system protein. **B)** Unitig-based
884 associations with an exploratory dataset (n = 131), which included related *X. euvesicatoria*
885 pathovars and assumes that host-range is conserved among *X. perforans* sequence clusters,
886 independent of an effector-triggered immunity. Unitigs with a probability > 40 after $-\log_{10}$
887 transformation and with a diagnostic accuracy ≥ 98% (shown in red) were considered significant
888 in this analysis. Variants were plotted against their corresponding positions in the *X. perforans*
889 LH3 genome. **C)** Annotation of pepper-associated variants located at *gltP* and its promoter region
890 and **D)** the major facilitator super-family gene. The P-values indicate the $-\log_{10}$ transformed
891 probability of the unitigs associated with a particular variant, based on the exploratory analysis.

892 **Figure 4**. Effect of a pepper-associated TonB-dependent receptor (TBDR) on *in-planta* growth
893 and pathogenicity of pepper-pathogenic *X. perforans* strain AL65 using a syringe infiltration
894 method. **A)** Representative pepper (top panel) and tomato (bottom two panels) leaves infiltrated
895 with wildtype AL65 (WT), AL65 ($TBDR_{2937}$) allele exchange mutant, and AL65 ($TBDR_{1734}$)
896 complemented with the wildtype allele, shown at the onset of symptoms, four days post-
897 inoculation with an initial inoculum adjusted to 10^4 CFU ml^{-1}. Leaves were soaked in tap water

for 20 min. prior to photographing to enhance water soaking. **B)** *In-planta* growth of *X. perforans* TBDR allele exchange mutants in pepper cv. Early CalWonder and tomato cv. FL47. Boxplots show Log_{10}-transformed CFU/cm^2 of leaf tissue from four independent experiments (n = 12 replicates per strain/host combination). Letters show signficant differences between treatments within days based on a mixed model analysis of variance followed by a Student's t-test for mean seperation ($p < 0.05$). Numbers under each boxplot indicate the mean.

Figure 5. Effect of a pepper-associated TonB-dependent receptor (TBDR) on *in-planta* growth and pathogenicity of pepper-pathogenic *X. perforans* strain AL65 using a dip inoculation technique. **A)** *In-planta* growth of *X. perforans* TBDR allele exchange mutants in pepper cv. Early CalWonder. The top panel shows the mean log_{10}-transformed CFU/cm^2 of leaf tissue from two independent experiments (n = 6 replicates per strain), while the bottom panel shows the mean disease severity rating. Letters above points indicate signficant differences between treatments within days based on a mixed model analysis of variance, followed by a Student's t-test for mean seperation. The error bar indiactes the standard error of the mean. **B)** Representative symptoms on pepper leaves induced by wildtype strain AL65 (WT), AL65 ($TBDR_{2937}$) allele exchange mutant, and AL65 ($TBDR_{1739}$) complemented with the wildtype allele, shown eight days post-inoculation with an initial inoculum adjusted to 10^7 CFU ml^{-1}.

Figure 6. Growth of *X. perforans* in the presence of polygalacturonic acid (PGA) is associated with pathogenicity phenotype and is not affected by a pepper-associated TonB-dependent receptor. The top panel shows the growth of wildtype *X. perforans* strains of differing genetic backgrounds (shown in parentheses) and pathogenicity phenotypes in modified XVM2 medium (mXVM2) amended with 0.25% PGA as the sole carbon source or standard XVM2, which contains 10 mM sucrose and fructose. The bottom panel shows the growth of pepper-pathogenic strain AL65 (WT)

921 (which carries a disruptive early stop codon in the pepper-associated TBDR), AL65 ($TBDR_{2937}$)
922 allele exchange mutant, and AL65 ($TBDR_{1734}$) allele exchange mutant, complemented with the
923 wildtype TBDR allele. Points show the mean of three independent experiments and error bars
924 indicate the standard error. Asterisks mark significant differences between pepper-pathogenic and
925 non-pathogenic strains based on a repeated measure mixed model analysis of variance.

Table 1. Top ten orthologous groups associated with the phenotype of pepper-pathogenicity within the *X. euvesicatoria* species complex, which consists of *X. perforans*, *X. euvesicatoria*, and related pathogens. Association tests were conducted with Scoary (Brynildsrud et al. 2016) using a conservative dataset, which included strains for which the phenotype could be experimentally verified and/or were isolated from naturally infected pepper leaves. They were also conducted using an exploratory dataset, based on the hypothesis that host range associated with genetic background (independent of an ETI) within the *X. perforans* lineage and that related *X. euvesicatoria* pathovars are non-pathogenic to pepper plants.

Conservative dataset (*n* = 71)				**Exploratory dataset (*n* = 131)**			
Gene/Annotation	**Locus tag[a]**	**Accuracy (%)[b]**	***p*[c]**	**Gene/Annotation**	**Locus tag[a]**	**Accuracy (%)[b]**	***p*[c]**
TonB-dependent receptor	IS_RS10330	97	2.98E^{-03}	TonB-dependent receptor	IS_RS10330	91	1.14E^{-17}
nuclear transport factor 2 family	IS_RS2339	91	0.4793	*tesB*	IS_RS07660	74	1.05E^{-15}
Polyvinylalchohol dehydrogenase	IS_RS22625	91	0.6386	dephospho-CoA kinase	IS_RS18050	73	3.23E^{-15}
Penicillin amidase	IS_RS15125	77	0.6386	pilin	IS_RS18035	71	9.48E^{-15}
dephospho-CoA kinase	IS_RS18050	77	0.9933	alcohol dehydrogenase family	IS_RS06450	68	7.29E^{-14}
tonB	IS_RS18605	74	0.9933	type II secretion system F family protein	IS_RS18040	73	1.90E^{-13}
type II secretion system F family protein	IS_RS18040	74	0.9933	*tonB*	IS_RS18605	73	1.90E^{-13}
tRNA-Leu(cag)	ND[d]	74	0.9933	hypothetical protein	IS_RS20965	73	1.90E^{-13}
hypothetical protein	IS_RS12015	74	0.9933	MFS transporter	IS_RS10010	73	4.89E^{-13}
NADPH quinone reductase	IS_RS06540	74	0.9333	ABC transporter	IS_RS10005	65	3.90E^{-12}

[a]Locus tag reported for *X. euvesicatoria* strain 85-10

[b]Diagnostic accuracy expressed as the sum of pepper-pathogenic strains in which a gene is present and the number of non-pathogenic strains in which a gene is absent (or vice versa), as a proportion of the total population

[c]Probability after Bonferroni correction

[d]Not determined

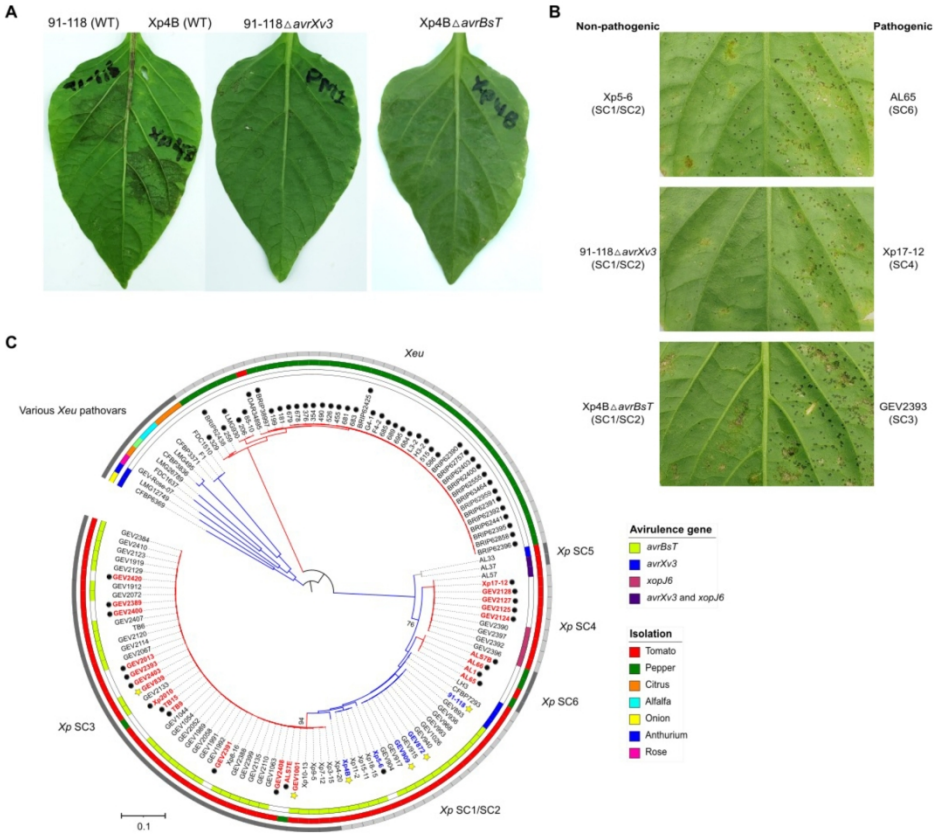


Figure 1. Population structure and pathogenicity analysis of *X. perforans*, *X. euvesicatoria*, and related pathogens. A) Hypersensitive response in pepper cv. ECW induced by *X. perforans* strains 91-118 and Xp4B, and lack of HR associated with mutagenesis of the avirulence genes *avrXv3* or *avrBsT* among the two strains, respectively. Leaves were photographed 24 h after inoculation with a bacterial suspension adjusted 10^8 CFU ml-1. B) Pathogenicity phenotyping of *X. perforans*. Pepper leaves were infiltrated with a bacterial suspension adjusted to 10^3 CFU ml-1 and qualitatively scored for pathogenicity based on the development of circular, water-soaked lesions, approximately 10 to 14 days after inoculation. Representative pathogenic (right side of leaf) and non-pathogenic (left side of leaf) strains are shown, with the genetic group of the individual strain indicated in parentheses. Leaves were soaked in tap water for 20 min. prior to photographing to enhance water-soaking symptoms. C) Maximum likelihood phylogeny of the *X. euvesicatoria* species complex, which consists of *X. perforans* (Xp), *X. euvesicatoria* (Xeu), and related Xeu pathovars, based on 93,380 SNPs that were extracted from a core-genome alignment of 3.71 Mbp. The inner ring shows the distribution of avirulence gene content, the middle ring indicates the host of isolation, and the outer ring denotes the genetic groups inferred through first level of BAPS hierarchy (Tonkin-Hill et al. 2019). The sequence clusters (SCs) present within the *Xp* lineage, as described previously (Newberry et al. 2019), are labeled accordingly. Within the *Xp* lineage, pepper-pathogenic strains are highlighted in red, non-pathogenic strains in blue, and strains that induced a HR in pepper leaves are shown in black. Black dots indicate the strains included in the conservative dataset for association testing. A star indicates a strain for which mutational analysis of a given avirulence gene was used to verify pathogenicity by Schwartz et al. (2015), which were also included in the conservative dataset.. The coloring of the branches shows the scoring of genetic groups that were used to test for associations in the exploratory dataset, which was based on the hypothesis that host-range is associated with genetic background within the *Xp* lineage and that related *Xeu* pathovars are non-pathogenic to pepper plants, independent of an effector-triggered immunity. Groups that harbor pepper-pathogenic strains are shown in red, those with non-pathogenic strains in blue,

and groups for which no phenotypic data are available in grey. The tree was drawn to scale and bootstrap values with < 95% support are shown at the nodes.

177x160mm (300 x 300 DPI)

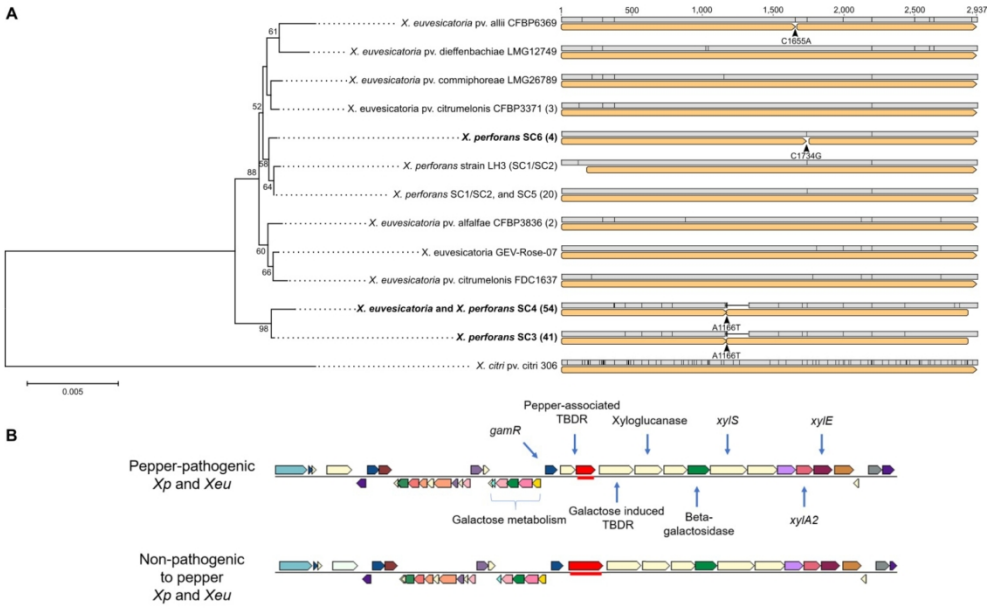


Figure 2. The coding sequence of a pepper-associated TonB-dependent receptor (TBDR) is disrupted through independent mutations among pepper-pathogenic *X. perforans* and *X. euvesicatoria*. A) Neighbor-joining phylogenetic tree (left) and multiple sequence alignment (right) of a TBDR associated with pepper-pathogenicity among members of the *X. euvesicatoria* species complex. The tree was constructed from the nucleotide sequences of 130 TBDRs extracted from *X. euvesicatora* and *X. perforans* genomes. The distribution of polymorphisms across the multiple sequence alignment is shaded within the grey boxes and yellow arrows below indicate the open reading frame of the sequence, as predicited with ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder/). Taxa that contain pepper-pathogenic *X. perforans* and *X. euvesicatoria* strains are highlighted in bold with the number of identical alleles at each tip shown in parentheses. The tree was rooted on *X. citri* pv. citri strain 306 and branch support, shown at the nodes, was assessed using 1,000 bootstrap replicates. B) Genomic neighborhood of the pepper-associated TBDR (highlighted in red). Genes of the same color indicate orthologous groups. The operon involved in galactose metabolism as described by Serrania et al. (2008); regulator of galactose associated metabolism (*gamR*) as described by Rashid et al. (2016); xylan/xylose related metabolism (*xyl*) as described by Déjean et al. (2013) and Feng et al. (2014).

177x120mm (300 x 300 DPI)

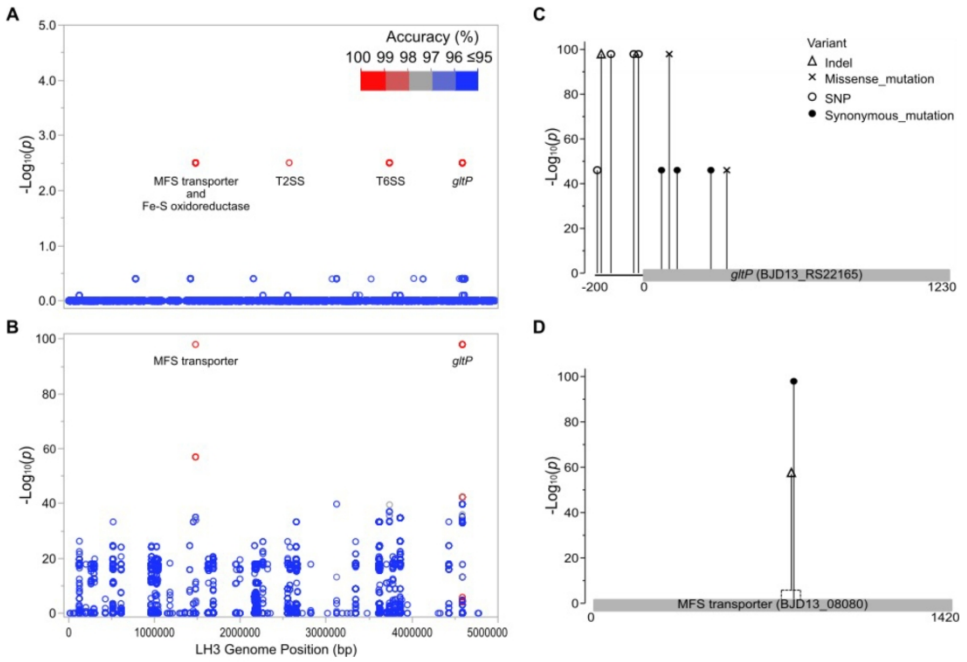


Figure 3. Associations between core-genome single nucleotide polymorphisms (SNPs) or genome-wide unitigs and pepper-pathogenicity within the *X. euvesicatoria* species complex. A) SNP-based associations derived from pairwise comparisons of probability after Bonferroni correction using a dataset of *X. perforans* and *X. euvesicatoria* strains for which the phenotype could be experimentally verified and/or were isolated from naturally infected pepper leaves (n = 71). Forty-five SNPs were significantly associated with pepper-pathogenicity ($p < 0.0032$, 100% diagnostic accuracy) and mapped to four chromosomal regions encompassing nine coding sequences. Genomic features located at pepper-associated variants are labeled accordingly; T2SS – type 2 secretions system protein, T6SS – type 6 secretion system protein. B) Unitig-based associations with an exploratory dataset (n = 131), which included related *X. euvesicatoria* pathovars and assumes that host-range is conserved among *X. perforans* sequence clusters, independent of an effector-triggered immunity. Unitigs with a probability > 40 after -log10 transformation and with a diagnostic accuracy ≥ 98% (shown in red) were considered significant in this analysis. Variants were plotted against their corresponding positions in the *X. perforans* LH3 genome. C) Annotation of pepper-associated variants located at *gltP* and its promoter region and D) the major facilitator super-family gene. The P-values indicate the -log10 transformed probability of the unitigs associated with a particular variant, based on the exploratory analysis.

177x126mm (300 x 300 DPI)

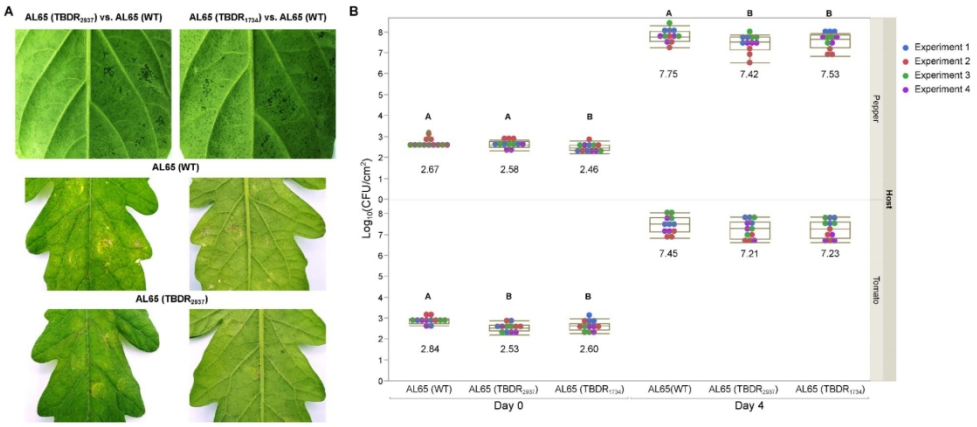


Figure 4. Effect of a pepper-associated TonB-dependent receptor (TBDR) on in-planta growth and pathogenicity of pepper-pathogenic *X. perforans* strain AL65 using infiltration method. A) Representative pepper (top panel) and tomato (bottom two panels) leaves infiltrated with wildtype AL65 (WT), AL65 (TBDR2937) allele exchange mutant, and AL65 (TBDR1734) complemented with the wildtype allele, shown at the onset of symptoms, four days post-inoculation with an initial inoculum adjusted to 10^4 CFU ml-1. Leaves were soaked in tap water for 20 min. prior to photographing to enhance water soaking. B) *In-planta* growth of *X. perforans* TBDR allele exchange mutants in pepper cv. Early CalWonder and tomato cv. FL47. Boxplots show Log10-transformed CFU/cm^2 of leaf tissue from four independent experiments (n = 12 replicates per strain/host combination). Letters show significant differences between treatments within days based on a mixed model analysis of variance followed by a Student's t-test for mean separation ($p < 0.05$). Numbers under each boxplot indicate the mean.

177x83mm (300 x 300 DPI)

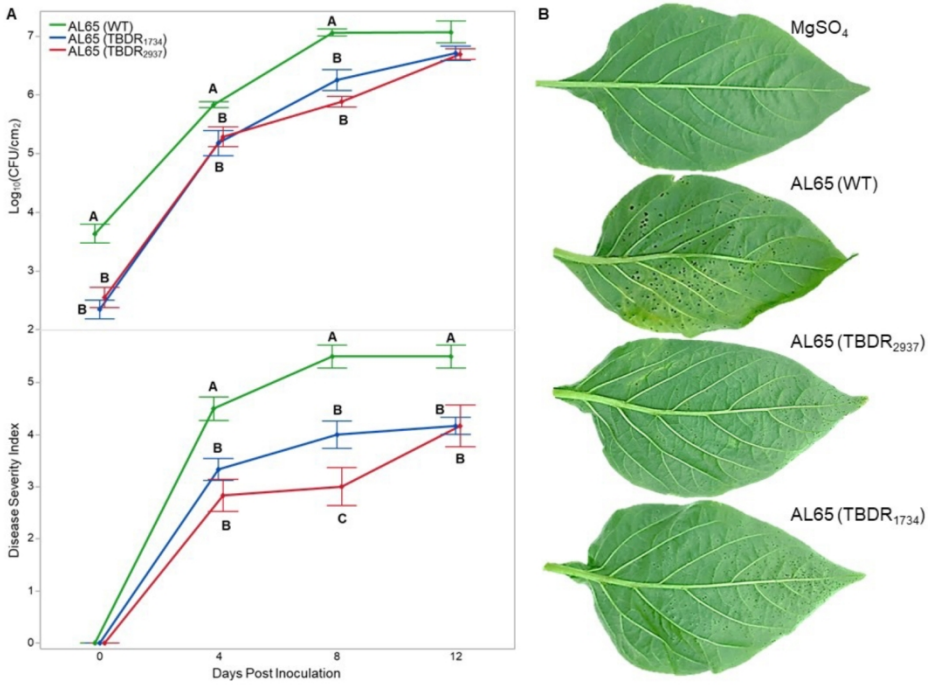


Figure 5. Effect of a pepper-associated TonB-dependent receptor (TBDR) on in-planta growth and pathogenicity of pepper-pathogenic *X. perforans* strain AL65 using a dip inoculation technique. A) *In-planta* growth of *X. perforans* TBDR allele exchange mutants in pepper cv. Early CalWonder. The top panel shows the mean log10-transformed CFU/cm^2 of leaf tissue from two independent experiments (n = 6 replicates per strain), while the bottom panel shows the mean disease severity rating. Letters above points indicate significant differences between treatments within days based on a mixed model analysis of variance, followed by a Student's t-test for mean separation. The error bar indicates the standard error of the mean. B) Representative symptoms on pepper leaves induced by wildtype strain AL65 (WT), AL65 (TBDR2937) allele exchange mutant, and AL65 (TBDR1739) complemented with the wildtype allele, shown eight days post-inoculation with an initial inoculum adjusted to 10^7 CFU ml^{-1}.

146x111mm (300 x 300 DPI)

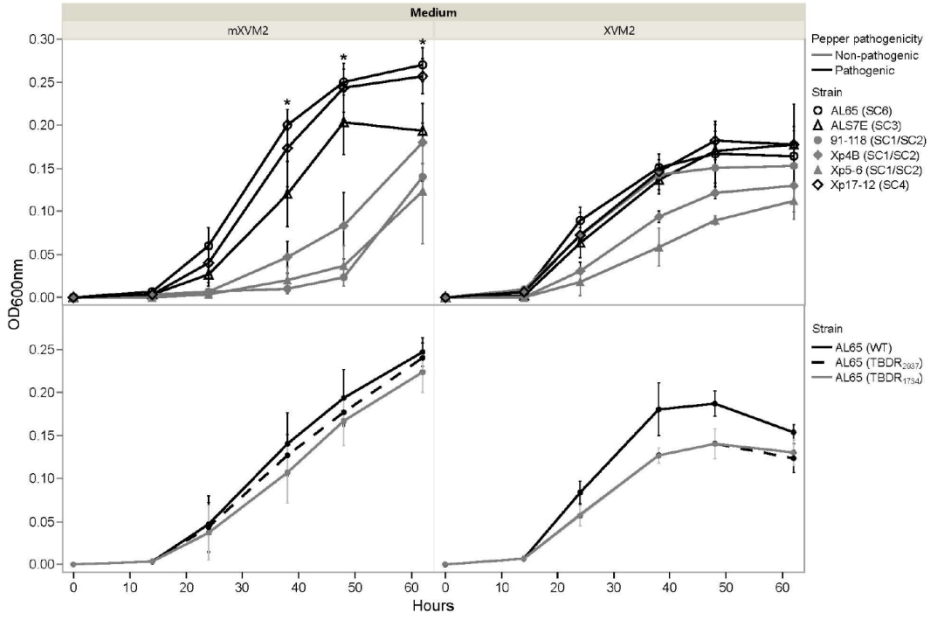


Figure 6. Growth of *X. perforans* in the presence of polygalacturonic acid (PGA) is associated with pathogenicity phenotype and is not affected by a pepper-associated TonB-dependent receptor. The top panel shows the growth of wildtype *X. perforans* strains of differing genetic backgrounds (shown in parentheses) and pathogenicity phenotypes in modified XVM2 medium (mXVM2) amended with 0.25% PGA as the sole carbon source or standard XVM2, which contains 10 mM sucrose and fructose. The bottom panel shows the growth of pepper-pathogenic strain AL65 (WT) (which carries a disruptive early stop codon in the pepper-associated TBDR), AL65 (TBDR2937) allele exchange mutant, and AL65 (TBDR1734) allele exchange mutant, complemented with the wildtype TBDR allele. Points show the mean of three independent experiments and error bars indicate the standard error. Asterisks mark significant differences between pepper-pathogenic and non-pathogenic strains based on a repeated measure mixed model analysis of variance.

147x100mm (300 x 300 DPI)

Supplementary Information

Genome-Wide Association to Study the Host-Specificity Determinants of *Xanthomonas perforans*

Newberry, E. A.[1], Minsavage, G.V.[2], Holland, A. [1], Jones, J.B.[2], and Potnis, N.[1*]

[1]Department of Entomology and Plant Pathology, Auburn University, AL 36849

[2]Department of Plant Pathology, University of Florida, FL 32611.

*Corresponding author: Neha Potnis, nzp0024@auburn.edu

Supplementary Material and Methods

Molecular Cloning Techniques. Standard laboratory procedures, as described by Sambrook et al. (1989), were employed for the extraction of genomic and plasmid DNA as well as agarose gel electrophoresis, while proprietary instructions were followed for all cloning procedures and enzymatic reactions. Plasmids were maintained in *Escherichia coli* DH5α (Intact Genomics Inc., St. Louis, MO), which was routinely cultured on lysogeny broth (LB) agar at 37℃. When appropriate, LB media was amended with filter-sterilized antibiotics at the following concentrations (μg ml^{-1}): kanamycin (Km), 50; tetracycline (Tet), 50; gentamicin (Gm), 10; and spectinomycin (Spc) 50. Plasmid vectors used in this study are listed in Supplementary Table 1.

Generation of TonB-dependent receptor (TBDR) allele exchange mutants. The TBDR coding sequences and endogenous promoter regions (approximately 440 bp of the upstream sequence) of *X. perforans* strains GEV2393, AL65, and Xp5-6 were PCR amplified with the following forward (5-GACGACGCGACTACATCCTG-3') and reverse (5'-AGATGCGGTGTAAACCCGAG-3') primers. The PCR contained a final volume of 20 μl and consisted of 1x HF buffer, 0.2 mM dNTPs, 0.6 μM of each primer, 0.4 U of Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA), and 1 μl of DNA. The PCRs were run on T100 thermal cycler (Bio-Rad, Hercules, CA) using an initial denaturation of 98°C for 30 sec, followed by 30 cycles of denaturation at 98°C for 10 s, annealing at 62°C for 30 s, extension at 72°C for 60 s, and a final extension of 72°C for 7 min. The resulting ~3.5 kbp fragments were each cloned into the Gateway

entry vector, pENTR, then transferred to the Gateway compatible suicide vector, pLVC18, via LR Clonase II reaction. The TBDR-pLVC18 constructs were conjugated into appropriate *X. perforans* strains via triparental mating using the pRK2073 helper plasmid to increase transformation efficiency. Matings were selected for growth on nutrient yeast-extract glycerol agar (NYGA) plates, supplemented with 100 μg ml^{-1} rifamycin and 50 μg ml^{-1} tetracycline, and subjected to several rounds of sub-culturing on NYGA without tetracycline selection. Serial dilutions were conducted to screen for putative transformants, where single colonies were replica-plated on media with and without tetracycline selection to identify tetracycline-sensitive double-crossover mutants. Verification that homologous recombination had occurred at the TBDR locus was obtained through PCR and Sanger sequencing of amplicons, as described above.

Generation of *gltP* allele exchange mutants. Initially, *gltP* topoisomerase insertion (TOPO) mutants were constructed using a ~500 bp region of the coding sequence (positions 236-741) that had been amplified from strain Xp5-6 using the following forward (5'-GCTATACCGTGGTGCTGTCC -3') and reverse (5'- AATGCCGTAACTGACCACCA -3') primers. To generate amplicons for the *gltP* 500 bp fragment and the coding sequences described below, the PCR reactions contained a final volume of 25 μl and consisted of 1x PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.6 μM of each primer, 0.625 U of *Taq* DNA polymerase (Promega, Madison, WI), and 2 μl of DNA. Cycling conditions consisted of 24 cycles of denaturation at 95°C for 30 s, annealing at 56°C for 15 s, extension at 72°C for 30 s, and a final extension of 72°C for 5 min. The resulting amplicon was cloned into pCR8 and the purified plasmid was subsequently transformed into the AL65-TBDR allele exchange mutant and Xp5-6ΔTBDR mutant using the electroporation protocol described by Amaral et al. (2005). In parallel, the entire *gltP* open reading frame and ~500 bp intergenic region upstream of the coding sequence (approximately 1.90 kbp in

total) was amplified from strains AL65 and Xp5-6 with the following forward (5'-CACACCTAACAGTGGCTCGT -3') and reverse (5'- CAGCGTTAGGGGAATCGG -3') primers. Cycling conditions for the PCR included 95°C for 3 min, followed by 34 cycles of denaturation at 95°C for 30 s, annealing at 56°C for 30 s, extension at 72°C for 60 s, and a final extension of 72°C for 5 min. The amplicons were each cloned in pCR8 and then transferred to pUFR047 using *Eco*RI digestion (New England Biolabs, Ipswich, MA) followed by T4 DNA ligation (Promega, Madison, WI). The resulting pUFR047-*gltP* constructs were subsequently conjugated into AL65 and Xp5-6 derived *gltP* knockout mutants using the tri-parental mating protocol described above, using kanamycin and gentamicin to select for transformants.

Supplementary Table 1. Bacterial strains and plasmids used in functional analysis of pepper associated variants.

Strains and plasmids	**Source/Reference**	**Notes**
Xanthomonas perforans		
AL65; Rif^r	Newberry et al. 2019	Pepper-virulent strain from cluster SC6
GEV2393; Rif^r	Abrahamian et al. 2019	Pepper-virulent strain from cluster SC3
Xp5-6; Rif^r	Schwartz et al. 2015	Pepper-avirulent strain from cluster SC2 that does not induce a hypersensitive response in pepper cv. ECW
Xp4BΔ*avrBsT*	Schwartz et al. 2015	Pepper-avirulent strain with a marker-less, in-frame deletion of *avrBsT*
91-118Δ*avrXv3*	Schwartz et al. 2015	Pepper-avirulent strain with a transposon insertion in *avrXv3*; Km^r
Escherichia coli		
DH5α	Intact Genomics Inc., St. Louis, MO	F - *mcrA* Δ(mrr-hsdRMS-mcrBC) *endA1 recA1 φ80dlacZΔM15* Δ*lacX74 araD139* Δ(ara, leu)7697 *galU galK rpsL* (Str^R) *nupG* λ-
AL65 derivatives		
AL65 (TBDR$_{2937}$); Rif^r	This study	Unmarked TBDR* allele exchange mutant with an intact coding sequence derived from cluster SC1/SC2 *X. perforans* strain, Xp5-6
AL65 (TBDR$_{1739}$); Rif^r	This study	Unmarked TBDR allele exchange mutant complemented with the wildtype allele, which contains a disruptive early stop codon
Xp5-6 derivatives		
Xp5-6ΔTBDR; Rif^r	This study	Unmarked TBDR allele exchange mutant, derived from SC3 *X. perforans* strain GEV2393, which contains a disruptive early stop codon
Xp5-6ΔTBDR, Δ*gltP;* Rif^r, Km^r	This study	TBDR allele exchange mutant with a disruptive early stop codon and topoisomerase insertion targeting *gltP*

Xp5-6ΔTBDR, Δ*gltP, gltP* (SC6); Rif^r, Km^r, Gm^r	This study	TBDR allele exchange mutant with a disruptive early stop codon, topoisomerase insertion targeting *gltP*, and *gltP* allele from pepper-virulent strain AL65 in pUFR47
Xp5-6ΔTBDR, Δ*gltP, gltP* (WT); Rif^r, Km^r, Gm^r	This study	TBDR allele exchange mutant with a disruptive early stop codon, topoisomerase insertion targeting *gltP*, and wildtype *gltP* allele in pUFR47
GEV2393 derivatives		
GEV2393 (TBDR$_{2937}$); Rif^r	This study	Unmarked TBDR* allele exchange mutant with an intact coding sequence derived from cluster SC2 *X. perforans* strain, Xp5-6
Plasmids		
pENTR/D-TOPO	Invitrogen, Carlsbad, CA	Gateway entry vector; Km^r
pCR8/GW/TOPO	Invitrogen, Carlsbad, CA	Gateway entry vector; Spc^r
pLVC18-RfC	Roden et al. 2004	Gateway destination suicide vector; Tet^r
pUFR047	DeFeyter et al. 1990	IncW, Mob+, lacZα+, Par+, Gm^r
pRK2073	Figurski and Helinski 1979	Helper plasmid; Spc^r, Strep^r
pENTR derivatives		
pENTR (TBDR-GEV2393)	This study	TBDR coding sequence and promoter region from SC3 strain GEV2393
pENTR (TBDR-Xp5-6)	This study	TBDR coding sequence and promoter region from SC2 strain GEV2393
pENTR (TBDR-AL65)	This study	TBDR coding sequence and promoter region from SC6 strain AL65
pCR8 derivatives		
pCR (*gltP*-Topo)	This study	Approximately 500 bp fragment (positions 236-741) of *gltP* for topoisomerase insertion
pCR8 (*gltP*-AL65)	This study	Coding sequence and promoter of *gltP* from strain AL65
pCR8 (*gltP*-Xp5-6)	This study	Coding sequence and promoter of *gltP* from strain Xp5-6
pLVC18 derivatives		
pLVC18 (TBDR-GEV2393)	This study	TBDR coding sequence and promoter from strain GEV2393
pLVC18 (TBDR-Xp5-6)	This study	TBDR coding sequence and promoter from strain Xp5-6
pLVC18 (TBDR-AL65)	This study	TBDR coding sequence and promoter from strain AL65

pUFR047 derivatives		
pUFR047 (*gltP*-Xp5-6)	This study	EcoRI fragment with *gltP* coding sequence and promoter from strain Xp5-6
pUFR047 (*gltP*-AL65)	This study	EcoRI fragment with *gltP* coding sequence and promoter from strain AL65

*Pepper associated TonB-dependent receptor (TBDR; locus tag IS_RS10330)

Supplementary Table 2. National Center for Biotechnology GenBank accession numbers and associated metadata for the *Xanthomonas* strains/genomes analyzed in this study.

Strain	Accession	Isolation	Pepper pathogenicity[a]	Avirulence gene[b]	Reference[a]
X. euvesicatoria:					
181*	GCA_001010095.1	Pepper	NT	-	-
199*	GCA_001008975.1	Pepper	NT	-	-
206*	GCA_001008815.1	Pepper	NT	-	-
259*	GCA_001008965.1	Pepper	NT	-	-
329*	GCA_001008805.1	Pepper	NT	-	-
354*	GCA_001008995.1	Pepper	NT	-	-
376*	GCA_001009045.1	Pepper	NT	-	-
455*	GCA_001009055.1	Pepper	NT	-	-
490*	GCA_001009075.1	Pepper	NT	-	-
515*	GCA_001008825.1	Pepper	NT	-	-
526*	GCA_001008835.1	Pepper	NT	-	-
586*	GCA_001008885.1	Pepper	NT	-	-
678*	GCA_001008915.1	Pepper	NT	-	-
679*	GCA_001008895.1	Pepper	NT	-	-
681*	GCA_001008905.1	Pepper	NT	-	-
683*	GCA_001009095.1	Pepper	NT	-	-
684*	GCA_001009125.1	Pepper	NT	-	-
685*	GCA_001009135.1	Pepper	NT	-	-
689*	GCA_001009205.1	Pepper	NT	-	-
695*	GCA_001009215.1	Pepper	NT	-	-
85-10*	GCA_000009165.1	Pepper	+	-	Potnis et al. 2011
BRIP38997*	GCA_003993175.1	Pepper	+	-	Roach et al. 2018
BRIP62390*	GCA_003993345.1	Pepper	+	-	Roach et al. 2018
BRIP62391*	GCA_003993335.1	Pepper	+	-	Roach et al. 2018
BRIP62392*	GCA_003993725.1	Pepper	+	-	Roach et al. 2018
BRIP62395*	GCA_003993315.1	Pepper	+	-	Roach et al. 2018
BRIP62396*	GCA_003993275.1	Pepper	+	-	Roach et al. 2018

BRIP62400*	GCA_003993685.1	Pepper	+	-	Roach et al. 2018
BRIP62403*	GCA_003993675.1	Pepper	+	-	Roach et al. 2018
BRIP62425*	GCA_003993265.1	Pepper	+	-	Roach et al. 2018
BRIP62438*	GCA_003993255.1	Pepper	+	-	Roach et al. 2018
BRIP62441*	GCA_003993655.1	Pepper	+	-	Roach et al. 2018
BRIP62555*	GCA_003993605.1	Pepper	+	-	Roach et al. 2018
BRIP62757*	GCA_003993615.1	Pepper	+	-	Roach et al. 2018
BRIP62858*	GCA_003993595.1	Pepper	+	-	Roach et al. 2018
BRIP62959*	GCA_003993225.1	Pepper	+	-	Roach et al. 2018
BRIP63464*	GCA_003993185.1	Pepper	+	-	Roach et al. 2018
DAR34895*	GCA_003992805.1	Pepper	+	-	Roach et al. 2018
F4-2*	GCA_001009165.1	Pepper	NT	-	-
G4-1*	GCA_001009245.1	Pepper	NT	-	-
H3-2*	GCA_001009175.1	Pepper	NT	-	-
L3-2*	GCA_001009255.1	Pepper	NT	-	-
LMG930*	GCA_001908795.1	Pepper	NT	-	-
X. perforans:					
91-118*	GCA_000192045.3	Tomato	-[c]	*avrXv3*	Schwartz et al. 2015
AL1*	GCA_007714105.1	Tomato	+	-	This study
AL33	GCA_007713965.1	Tomato	HR	*avrXv3*	This study
AL37	GCA_007713985.1	Tomato	HR	*avrXv3, xopJ6*	This study
AL66*	GCA_007714065.1	Pepper	+	-	This study
ALS7B*	GCA_007714045.1	Pepper	+	-	This study
ALS7E*	GCA_007714075.1	Pepper	+	-	This study
AL65*	GCA_007714115.1	Pepper	+	-	This study
AL57	GCA_007713955.1	Tomato	HR	*avrXv3, xopJ6*	This study
CFBP7293	GCA_001976075.1	Tomato	NT	*avrXv3*	-
GEV1001*	GCA_001010025.1	Tomato	+[c]	*avrBsT*	Schwartz et al. 2015
GEV1026	GCA_001010035.1	Tomato	HR	*avrBsT*	This study
GEV1044	GCA_001009935.1	Tomato	HR	*avrBsT*	This study

GEV1054	GCA_001009925.1	Tomato	HR	*avrBsT*	This study
GEV1063	GCA_001010085.1	Tomato	HR	*avrBsT*	This study
GEV1912	GCA_006980315.1	Tomato	HR	*avrBsT*	This study
GEV1919	GCA_006980475.1	Tomato	HR	*avrBsT*	This study
GEV1989	GCA_004102325.1	Tomato	HR	*avrBsT*	This study
GEV1991	GCA_006980665.1	Tomato	HR	*avrBsT*	This study
GEV1992	GCA_006980655.1	Tomato	HR	*avrBsT*	This study
GEV2013*	GCA_006980585.1	Tomato	+	-	This study
GEV2052	GCA_006980525.1	Tomato	HR	*avrBsT*	This study
GEV2058	GCA_006980385.1	Tomato	HR	*avrBsT*	This study
GEV2067	GCA_006980175.1	Tomato	HR	*avrBsT*	This study
GEV2072	GCA_006980115.1	Tomato	HR	*avrBsT*	This study
GEV2110	GCA_006980465.1	Tomato	HR	*avrBsT*	This study
GEV2114	GCA_006980395.1	Tomato	HR	*avrBsT*	This study
GEV2120	GCA_004102205.1	Tomato	HR	*avrBsT*	This study
GEV2123	GCA_006979915.1	Tomato	HR	*avrBsT*	This study
GEV2124*	GCA_004102415.1	Tomato	+	-	This study
GEV2125*	GCA_004102085.1	Tomato	+	-	This study
GEV2127*	GCA_004102065.1	Tomato	+	-	This study
GEV2128*	GCA_006979875.1	Tomato	+	-	This study
GEV2129	GCA_004102425.1	Tomato	HR	*avrBsT*	This study
GEV2133	GCA_006979855.1	Tomato	HR	*avrBsT*	This study
GEV2135	GCA_004102335.1	Tomato	HR	*avrBsT*	This study
GEV2384	GCA_006979835.1	Tomato	HR	*avrBsT*	This study
GEV2388	GCA_006979815.1	Tomato	HR	*avrBsT*	This study
GEV2389*	GCA_006979805.1	Tomato	+	-	This study
GEV2390	GCA_006979785.1	Tomato	HR	*xopJ6*	This study
GEV2391*	GCA_006979715.1	Tomato	+	-	This study
GEV2392	GCA_006979735.1	Tomato	HR	*xopJ6*	This study
GEV2393*	GCA_006979725.1	Tomato	+	-	This study
GEV2396	GCA_006979645.1	Tomato	HR	*xopJ6*	This study
GEV2397	GCA_006979675.1	Tomato	HR	*xopJ6*	This study
GEV2399	GCA_006979635.1	Tomato	HR	*avrBsT*	This study

GEV2400*	GCA_006979625.1	Tomato	+	-	This study
GEV2403*	GCA_006979535.1	Tomato	+	-	This study
GEV2407	GCA_006979525.1	Tomato	HR	*avrBsT*	This study
GEV2408*	GCA_006979555.1	Tomato	+	-	This study
GEV2410	GCA_006979565.1	Tomato	HR	*avrBsT*	This study
GEV2420*	GCA_006979545.1	Tomato	+	-	This study
GEV839*	GCA_001009475.1	Tomato	+[c]	*avrBsT*	Schwartz et al. 2015
GEV872*	GCA_001009485.1	Tomato	-[c]	*avrBsT*	Schwartz et al. 2015
GEV893	GCA_001009545.1	Tomato	HR	*avrBsT*	This study
GEV904	GCA_001009795.1	Tomato	HR	*avrBsT*	This study
GEV909*	GCA_001009825.1	Tomato	-[c]	*avrBsT*	Schwartz et al. 2015
GEV915	GCA_001009855.1	Tomato	HR	*avrBsT*	This study
GEV917	GCA_001009865.1	Tomato	HR	*avrBsT*	This study
GEV936	GCA_001009845.1	Tomato	HR	*avrBsT*	This study
GEV940	GCA_001009885.1	Tomato	HR	*avrBsT*	This study
GEV968	GCA_001010015.1	Tomato	HR	*avrBsT*	This study
GEV993	GCA_001010005.1	Tomato	HR	*avrBsT*	This study
LH3	GCA_001908855.1	Tomato	NT	*avrXv3*	-
TB15*	GCA_001010105.1	Tomato	+	-	Schwartz et al. 2015
TB6	GCA_001009945.1	Tomato	HR	*avrBsT*	This study
TB9*	GCA_001009955.1	Tomato	+	-	Schwartz et al. 2015
Xp10-13	GCA_001009405.1	Tomato	HR	*avrBsT*	This study
Xp11-2	GCA_001009445.1	Tomato	HR	*avrBsT*	This study
Xp15-11	GCA_001009465.1	Tomato	HR	*avrBsT*	This study
Xp17-12*	GCA_001009745.1	Tomato	+	-	Schwartz et al. 2015
Xp18-15	GCA_001009765.1	Tomato	HR	*avrBsT*	This study
Xp2010*	GCA_001009785.1	Pepper	+	-	Schwartz et al. 2015
Xp3-15	GCA_001009675.1	Tomato	HR	*avrBsT*	This study
Xp4-20	GCA_001009705.1	Tomato	HR	*avrBsT*	This study
Xp4B*	GCA_001009665.1	Tomato	-[c]	*avrBsT*	Schwartz et al. 2015
Xp5-6*	GCA_001009365.1	Tomato	-	-	Schwartz et al. 2015
Xp7-12	GCA_001009385.1	Tomato	HR	*avrBsT*	This study
Xp8-16	GCA_001009685.1	Tomato	HR	*avrBsT*	This study

Xp9-5	GCA_001009395.1	Tomato	HR	*avrBsT*	This study
Various *X. euvesicatoria* pathovars:					
CFBP3371	GCA_002939715.1	Citrus	NT	-	-
CFBP6369	GCA_000730305.1	Onion	NT	*avrXv3*	-
F1	GCA_000225915.1	Citrus	-[d]	-	Huang et al. 2013
FDC1510	GCA_005059785.1	Citrus	NT	-	-
FDC1637	GCA_005059795.1	Citrus	NT	-	-
LMG12749	GCA_001401675.2	Anthurium	NT	*avrXv3*	-
LMG26789	GCA_003698225.1	Comiphora	HR[e]	-	Bansal et al. 2018; Samanta et al. 2013
LMG495	GCA_001401625.1	Alfalfa	NT	-	-
CFBP3836	GCA_000488955.1	Alfalfa	-[d]	-	Yaripour et al. 2018
GEV-Rose-07	GCA_001855615.1	Rose	-[d]	-	Huang et al. 2013

[a] Pathogenicity phenotype in pepper leaves and study in which the strain was tested; (+) = pathogenic, (-) = non-pathogenic in the absence of an effector-triggered immunity, HR = hypersensitive response, NT = not tested

[b] Avirulence genes recognized by pepper cv. ECW were predicted through BLASTn analysis

[c] Mutational analysis was conducted by Schwartz et al. (2015) to assess the pathogenicity of the strain in the absence of an effector-triggered immunity

[d] Strain was reported to be non-pathogenic but was not tested for HR

[e] Strain was reported to induce a HR but no evidence of a canonical avirulence gene was detected in its genome

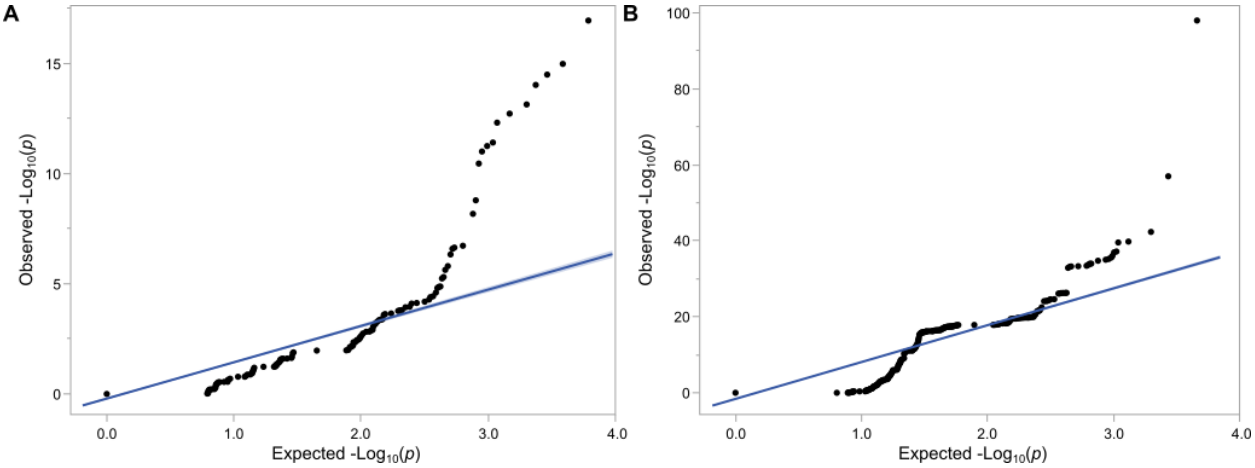
* Strain was included in the conservative dataset for association testing

Supplementary Table 3. Core genome single-nucleotide polymorphisms associated with pepper-pathogenic *X. perforans* and *X. euvesicatoria*, based on a conservative genome sample (n = 71) using Scoary (Brynildsrud et al. 2016).

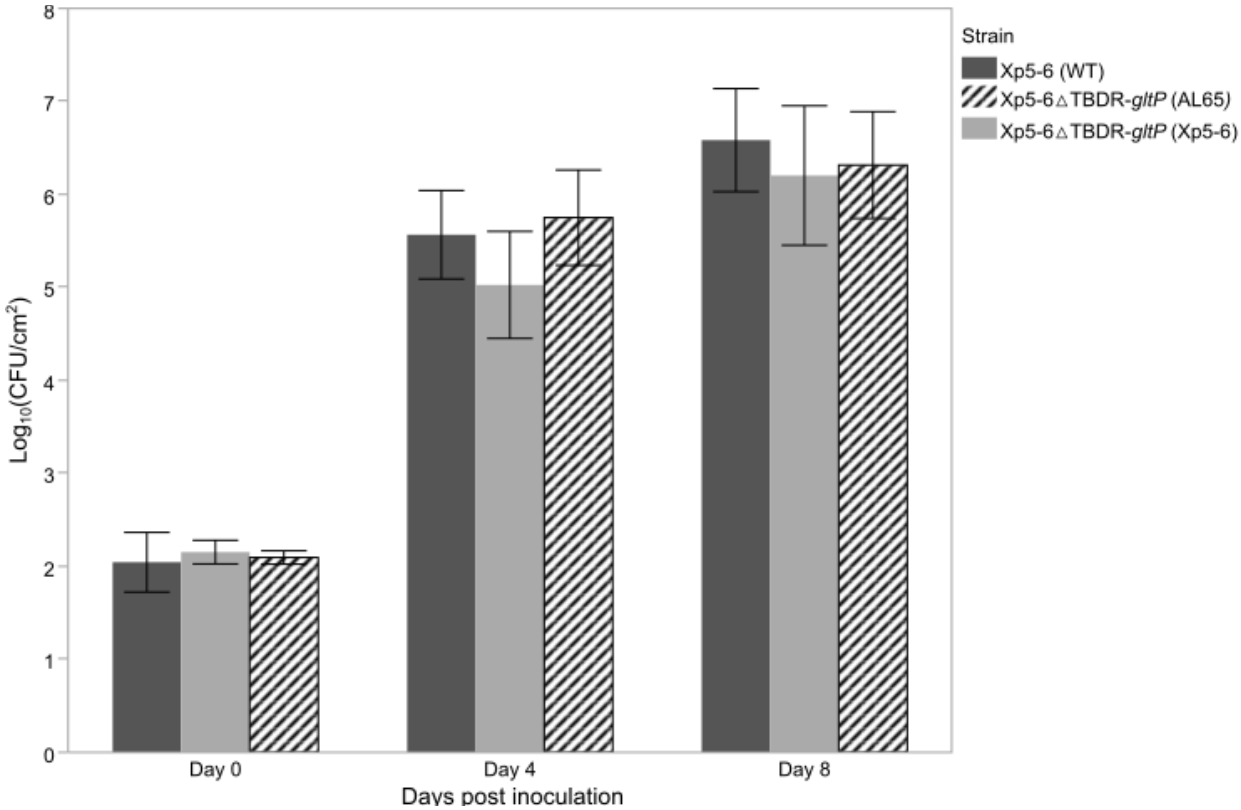
Genome position[a]	Gene/Annotation	Sensitivity	Specificity	Bonferroni (p)
4587220	*gltP*	100	100	0.00316
4587076	*gltP*	100	100	0.00316
4586977	*gltP*	100	100	0.00316
4586869	*gltP*	100	100	0.00316
4586850	*gltP*	100	100	0.00316
4586786	*gltP*	100	100	0.00316
4586618	*gltP*	100	100	0.00316
4586494	*gltP*	100	100	0.00316
4586475	*gltP*	100	100	0.00316
4586383	Intergenic region	100	100	0.00316
4585626	Membrane protein	100	100	0.00316
4584918	Membrane protein	100	100	0.00316
4584840	Membrane protein	100	100	0.00316
3738181	Type VI secretion tetratricopeptide repeat protein	100	100	0.00316
3737192	Type VI secretion tetratricopeptide repeat protein	100	100	0.00316
3737156	Type VI secretion tetratricopeptide repeat protein	100	100	0.00316
3737110	Type VI secretion tetratricopeptide repeat protein	100	100	0.00316
3737108	Type VI secretion tetratricopeptide repeat protein	100	100	0.00316
3737027	Type VI secretion Rhs protein	100	100	0.00316
2570910	Type II secretion protein XpsD	0	0	0.00316
1486583	Hypothetical protein	100	100	0.00316
1486408	Hypothetical protein	100	100	0.00316
1486291	Fe-S oxidoreductase	100	100	0.00316
1486241	Fe-S oxidoreductase	100	100	0.00316
1486180	Fe-S oxidoreductase	100	100	0.00316
1486143	Fe-S oxidoreductase	100	100	0.00316
1486050	Fe-S oxidoreductase	100	100	0.00316
1485607	Fe-S oxidoreductase	100	100	0.00316

1485597	Fe-S oxidoreductase	100	100	0.00316
1485321	Fe-S oxidoreductase	100	100	0.00316
1484739	Intergenic region	100	100	0.00316
1484609	Hpr(Ser) kinase/phosphatase	100	100	0.00316
1484489	Hpr(Ser) kinase/phosphatase	100	100	0.00316
1484317	Hpr(Ser) kinase/phosphatase	100	100	0.00316
1484300	Hpr(Ser) kinase/phosphatase	100	100	0.00316
1481444	MFS transporter	100	100	0.00316
1481321	MFS transporter	100	100	0.00316
1481192	MFS transporter	100	100	0.00316
1480775	MFS transporter	100	100	0.00316
1480592	MFS transporter	100	100	0.00316
1480580	MFS transporter	100	100	0.00316
1480298	MFS transporter	100	100	0.00316
1480295	MFS transporter	100	100	0.00316
1480147	Intergenic region	100	100	0.00316
1480129	Intergenic region	100	100	0.00316

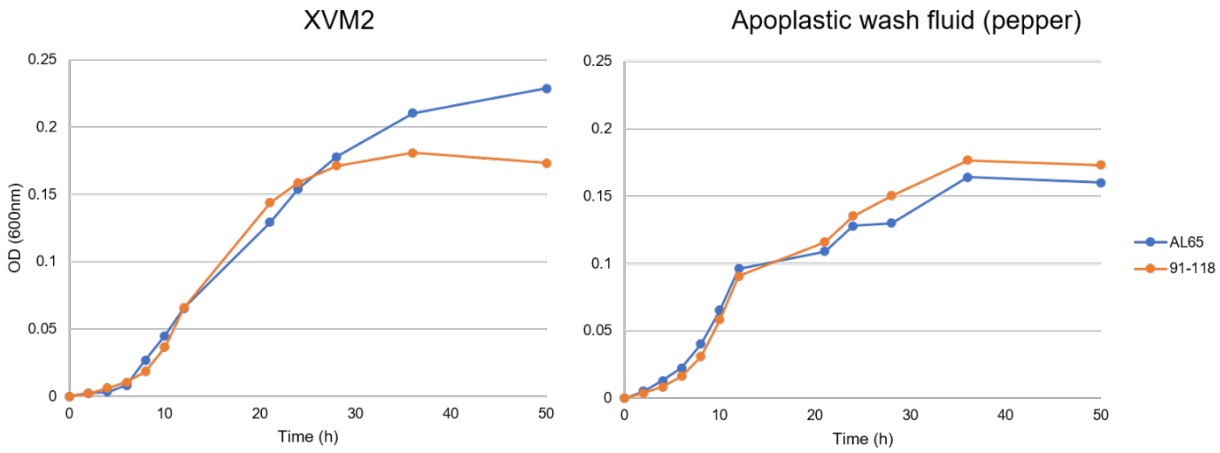
[a]Genome position in the LH3 chromosome (NZ_CP018475.1)



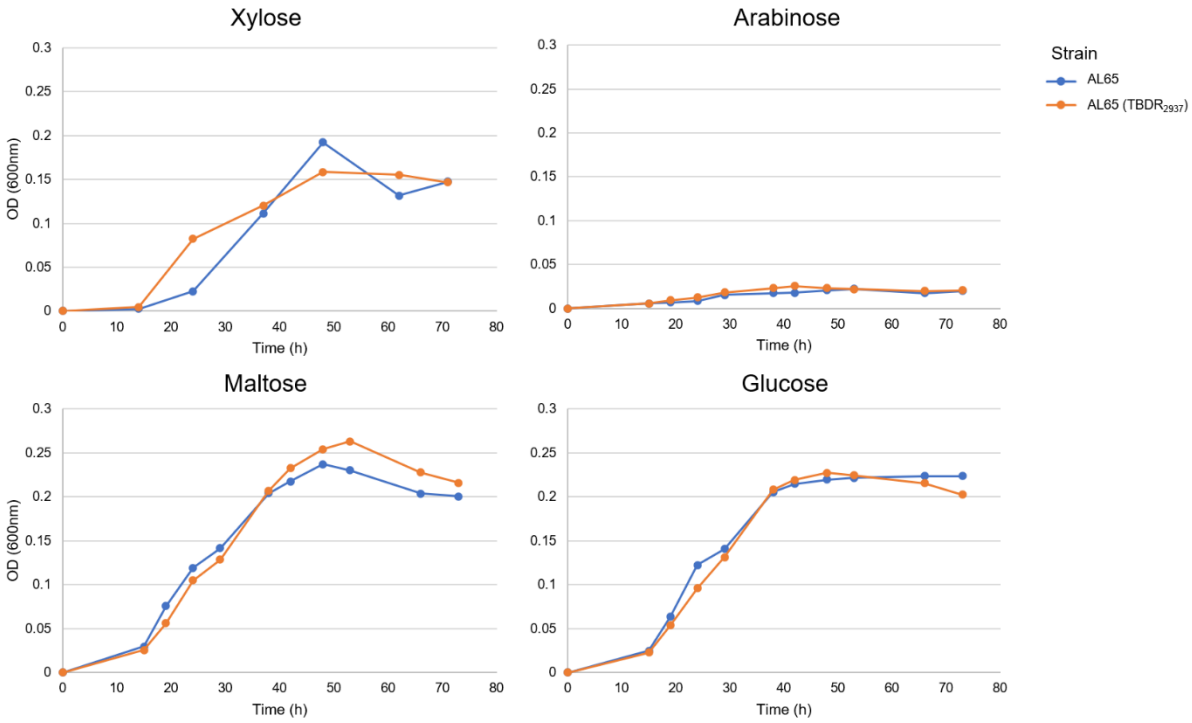
Supplementary Figure 1. Quantile-quantile plots for the -Log10 transformed P-values for the **A)** pan-genome orthologous groups and **B)** genome-wide unitigs, based on the exploratory dataset.



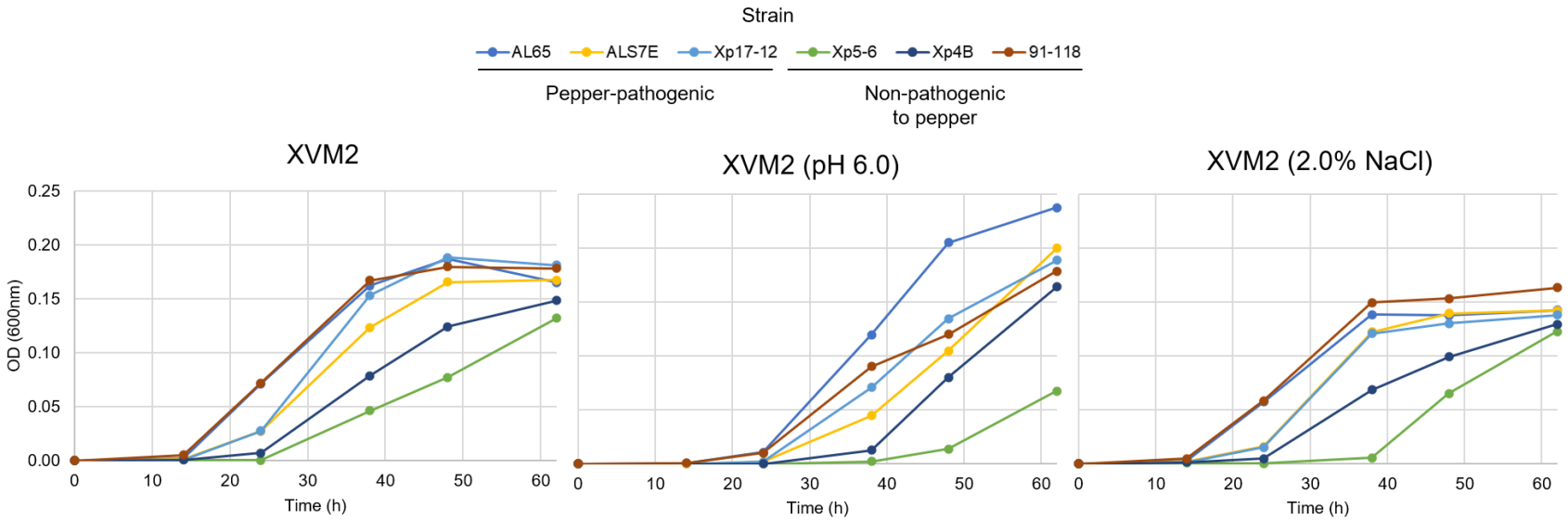
Supplementary Figure 2. Effect of a pepper-associated variants on *in-planta* growth of *X. perforans* strain Xp5-6, which was non-pathogenic to pepper. Mean Log_{10}-transformed CFU/cm^2 of leaf tissue in pepper cv. ECW of wildtype (WT) strain Xp5-6, Xp5-6 double TonB-dependent receptor (TBDR) and *gltP* mutant complemented with the *gltP* allele derived from pepper-pathogenic strain AL65, and Xp5-6 double TBDR and *gltP* mutant complemented with the wildtype *gltP* allele. Error bars show one standard error of the mean (n = 3), the experiment was repeated twice with similar results. No significant differences in growth were observed ($\alpha = 0.05$).



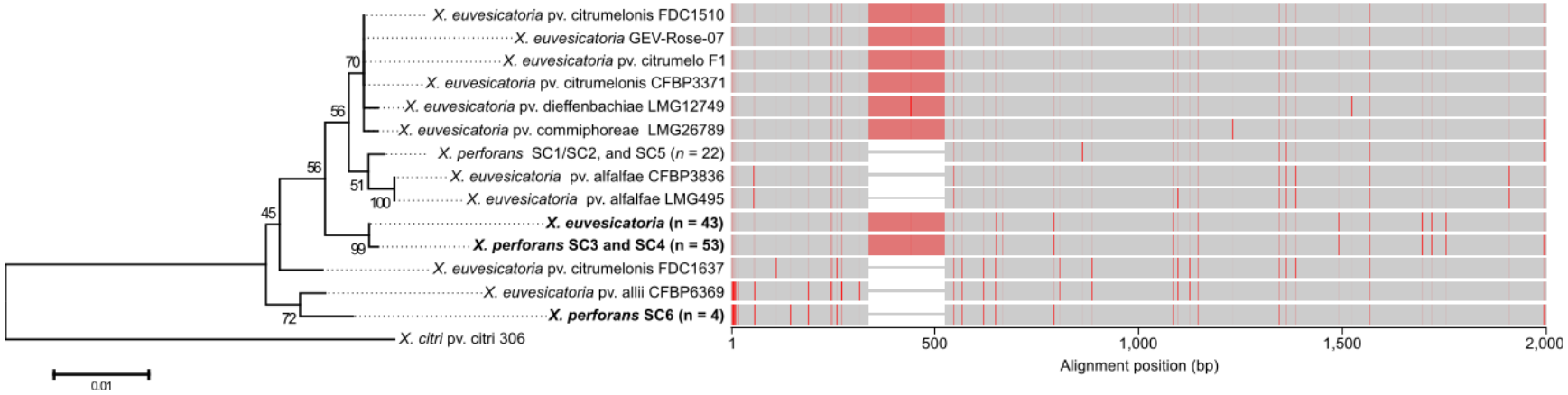
Supplementary Figure 3. Growth of pepper-pathogenic *X. perforans* strain AL65 and non-pathogenic strain 91-118 in XVM2 medium (left) and apoplastic was fluid extracted from pepper leaves (right). Representative results are shown, the experiment was conducted twice with similar results.



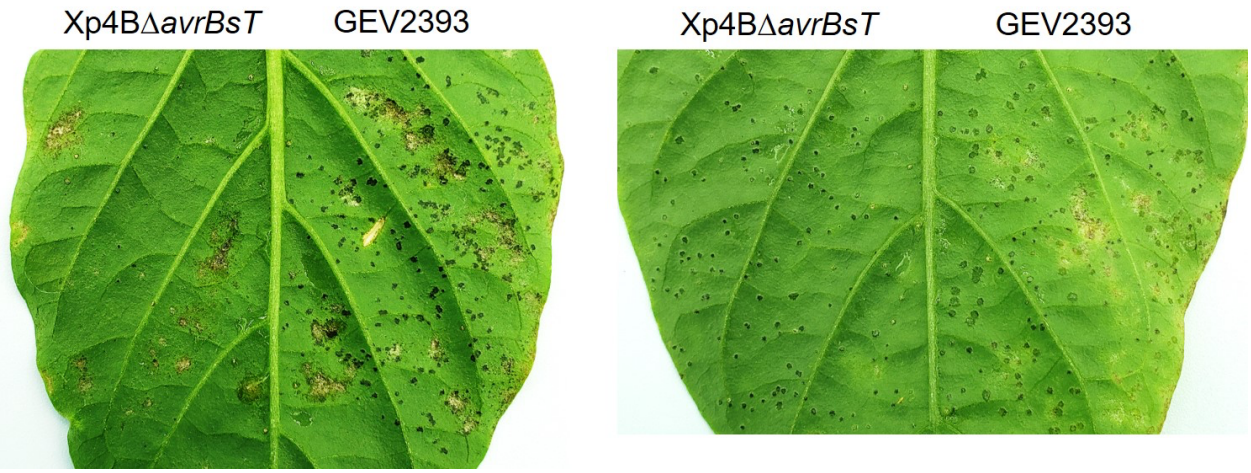
Supplementary Figure 4. Growth of pepper-pathogenic *X. perforans* strain AL65 and AL65 TonB-dependent receptor ($TBDR_{2937}$) allele exchange mutant in XVM2 medium modified with various sugars (10 mM) as the sole carbon source. Representative results are shown. These experiments were conducted at least twice with similar results.



Supplementary Figure 5. Effect of acidic and hypertonic conditions on the growth of pepper-pathogenic and non-pathogenic *X. perforans* strains in XVM2 medium. Representative results are shown. These experiments were conducted at least twice with similar results.



Supplementary Figure 6. Neighbor-joining phylogenetic tree (left) and multiple sequence alignment (right) of *gltP* and the intergenic region upstream of the coding sequence among members of the *X. euvesicatoria* species complex. Taxa with pepper-pathogenic *X. perforans* and *X. euvesicatoria* strains are shown in bold. The number of *X. euvesicatoria* and/or *X. perforans* sequence clusters (SCs) with the same *gltP* allele is shown in parentheses. The tree was rooted on *X. citri* pv. citri strain 306 and branch support, shown at the nodes, was assesed using 1,000 bootstrap replicates.



Supplementary Figure 7. Differentiation of susceptible and resistant lesions in pepper leaves incited by *X. perforans*. Laves were infiltrated with a bacterial suspension adjusted to 10^3 CFU ml^{-1} and photographed 10 days post-inoculation. Resistant lesions induced by strain Xp4BΔ*avrBsT* were characterized by the dark necrotic specks shown on the right. These are differentiated from the susceptible lesions induced by strain GEV2393 in that they are not accompanied by water-soaking and their development was variable from leaf to leaf.

References

Abrahamian, P., Timilsina, S., Minsavage, G. V., Potnis, N., Jones, J. B., Goss, E. M., et al. 2019. Molecular Epidemiology of *Xanthomonas perforans* Outbreaks in Tomato Plants from Transplant to Field as Determined by Single-Nucleotide Polymorphism Analysis. Appl. Environ. Microbiol. 85.

Amaral, A. M. do, Toledo, C. P., Baptista, J. C., and Machado, M. A. 2005. Transformation of *Xanthomonas axonopodis* pv. *citri* by electroporation. Fitopatologia Brasileira. 30:292–294.

Bansal, K., Kumar, S., and Patil, P. B. 2018. Complete Genome Sequence Reveals Evolutionary Dynamics of an Emerging and Variant Pathovar of *Xanthomonas euvesicatoria*. Genome Biol Evol. 10:3104–3109.

Brynildsrud, O., Bohlin, J., Scheffer, L., and Eldholm, V. 2016. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. Genome Biol. 17.

DeFeyter, R., Kado, C. I., and Gabriel, D. W. 1990. Small, stable shuttle vectors for use in *Xanthomonas*. Gene. 88:65–72.

Figurski, D. H., and Helinski, D. R. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proc Natl Acad Sci U S A. 76:1648–1652.

Huang, C.-H., Vallad, G. E., Adkison, H., Summers, C., Margenthaler, E., Schneider, C., et al. 2013. A Novel *Xanthomonas* sp. Causes Bacterial Spot of Rose (Rosa spp.). Plant Disease. 97:1301–1307.

Newberry, E. A., Bhandari, R., Minsavage, G. V., Timilsina, S., Jibrin, M. O., Kemble, J., et al. 2019. Independent Evolution with the Gene Flux Originating from Multiple *Xanthomonas* Species Explains Genomic Heterogeneity in *Xanthomonas perforans*. Appl. Environ. Microbiol. 85.

Potnis, N., Krasileva, K., Chow, V., Almeida, N. F., Patil, P. B., Ryan, R. P., et al. 2011. Comparative genomics reveals diversity among xanthomonads infecting tomato and pepper. BMC Genomics. 12:146.

Roach, R., Mann, R., Gambley, C. G., Shivas, R. G., and Rodoni, B. 2018. Identification of *Xanthomonas* species associated with bacterial leaf spot of tomato, capsicum and chilli crops in eastern Australia. Eur J Plant Pathol. 150:595–608.

Roden, J. A., Belt, B., Ross, J. B., Tachibana, T., Vargas, J., and Mudgett, M. B. 2004. A genetic screen to isolate type III effectors translocated into pepper cells during *Xanthomonas* infection. Proc Natl Acad Sci U S A. 101:16624–16629.

Samanta, J. N., Mandal, K., and Maiti, S. 2013. A novel pathovar of *Xanthomonas axonopodis* causes gumming of Guggal (*Commiphora wightii*). Eur J Plant Pathol. 135:115–125.

Sambrook, J., Fritsch, E., and Maniatis, T. 1989. *Molecular Cloning*. New Yrok, NY: Cold Spring Harbor Laboratory Press.

Schwartz, A. R., Potnis, N., Timilsina, S., Wilson, M., Patané, J., Martins, J. J., et al. 2015. Phylogenomics of *Xanthomonas* field strains infecting pepper and tomato reveals diversity in effector repertoires and identifies determinants of host specificity. Front. Microbiol. 6.

Yaripour, Z., Mohsen Taghavi, S., Osdaghi, E., and Lamichhane, J. R. 2018. Host range and phylogenetic analysis of *Xanthomonas alfalfae* causing bacterial leaf spot of alfalfa in Iran. Eur J Plant Pathol. 150:267–274.