Comparative genomics reveals diversity among xanthomonads infecting tomato and pepper

Permalink
https://escholarship.org/uc/item/4ht7x40d

Journal
BMC Genomics, 12(1)

ISSN
1471-2164

Authors
Potnis, Neha
Krasileva, Ksenia
Chow, Virginia
et al.

Publication Date
2011-03-11

DOI
http://dx.doi.org/10.1186/1471-2164-12-146

Supplemental Material
https://escholarship.org/uc/item/4ht7x40d#supplemental

Peer reviewed
Comparative genomics reveals diversity among xanthomonads infecting tomato and pepper

Neha Potnis1†, Ksenia Krasileva2†, Virginia Chow3, Nalvo F Almeida4, Prabhu B Patil5, Robert P Ryan6, Molly Sharlach2, Franklin Behlau1,7, J Max Dow8, MT Momol1, Frank F White8, James F Preston3, Boris A Vinatzer9, Ralf Koebnik10†, João C Setubal11†, David J Norman12, Brian J Staskawicz2†, Jeffrey B Jones1*†

Abstract

Background: Bacterial spot of tomato and pepper is caused by four Xanthomonas species and is a major plant disease in warm humid climates. The four species are distinct from each other based on physiological and molecular characteristics. The genome sequence of strain 85-10, a member of one of the species, Xanthomonas euvesicatoria (Xcv) has been previously reported. To determine the relationship of the four species at the genome level and to investigate the molecular basis of their virulence and differing host ranges, draft genomic sequences of members of the other three species were determined and compared to strain 85-10.

Results: We sequenced the genomes of X. vesicatoria (Xv) strain 1111 (ATCC 35937), X. perforans (Xp) strain 91-118 and X. gardneri (Xg) strain 101 (ATCC 19865). The genomes were compared with each other and with the previously sequenced Xcv strain 85-10. In addition, the molecular features were predicted that may be required for pathogenicity including the type III secretion apparatus, type III effectors, other secretion systems, quorum sensing systems, adhesins, extracellular polysaccharide, and lipopolysaccharide determinants. Several novel type III effectors from Xg strain 101 and Xv strain 1111 genomes were computationally identified and their translocation was validated using a reporter gene assay. A homolog to Ax21, the elicitor of XA21-mediated resistance in rice, and a functional Ax21 sulfation system were identified in Xcv. Genes encoding proteins with functions mediated by type II and type IV secretion systems have also been compared, including enzymes involved in cell wall deconstruction, as contributors to pathogenicity.

Conclusions: Comparative genomic analyses revealed considerable diversity among bacterial spot pathogens, providing new insights into differences and similarities that may explain the diverse nature of these strains. Genes specific to pepper pathogens, such as the O-antigen of the lipopolysaccharide cluster, and genes unique to individual strains, such as novel type III effectors and bacteriocin genes, have been identified providing new clues for our understanding of pathogen virulence, aggressiveness, and host preference. These analyses will aid in efforts towards breeding for broad and durable resistance in economically important tomato and pepper cultivars.

Background

Bacterial spot disease of tomato and pepper presents a serious agricultural problem worldwide, leading to significant crop losses especially in regions with warm and humid climate. The disease is characterized by necrotic lesions on leaves, sepals and fruits, reducing yield and fruit quality [1]. The disease is caused by a relatively diverse set of bacterial strains within the genus Xanthomonas; strain nomenclature and classification for the strains that infect pepper and tomato have gone through considerable taxonomic revision in recent years. Currently, the pathogens are classified into four distinct pathogen groups (A, B, C, and D) within the genus Xanthomonas. Strains belonging to groups A, B and D infect both tomato and pepper. Group C strains are pathogenic only on tomato [2,3]. These phenotypically and genotypically distinct strains have different geographic distributions. Strains of group A and B are found worldwide. C strains have been increasingly found in the U.S., Mexico, Brazil, Korea and regions bordering...
the Indian Ocean, and D group strains are found in the former Yugoslavia, Canada, Costa Rica, U.S, Brazil and regions of the Indian Ocean [4-8]. Three of the four groups except for D were originally described as a single pathovar within Xanthomonas campestris and referred to as X. campestris pv. vesicatoria. The D group consisted of a strain isolated from tomato that had been designated ‘Pseudomonas gardneri’ for many years [9] although De Ley provided evidence for placement in the genus Xanthomonas [10]. Subsequently all four groups were classified as separate species on the basis of physiological and molecular characteristics as follows: Xanthomonas euvesicatoria (group A), Xanthomonas vesicatoria (group B), Xanthomonas perforans (group C), and Xanthomonas gardneri (group D) [11].

Based on 16S rRNA analysis, X. euvesicatoria strain 85-10 (A group) and X. perforans (C group) together form a monophyletic group, whereas X. vesicatoria (B group) and X. gardneri (D group) cluster together with X. campestris pv. campestris (Xcc) Xcc strain 33913 [11]. Recently, a phylogenetic tree was constructed based on MLST (multi-locus sequence typing) data for A, B, C and D group strains and other xanthomonads [12]. The MLST approach revealed that X. euvesicatoria and X. perforans form a group along with X. citri strain 306 (Xac). X. gardneri is most closely related to X. campestris pv. campestris strains while X. vesicatoria forms a distinct clade [12]. This diversity among the four groups makes the Xanthomonas-tomato/pepper system an excellent example to study pathogen co-evolution, as distinct species have converged on a common host.

While integrated management approaches for control of bacterial spot disease are available, the development of host resistance is more economical and environmentally benign for the control of the disease [13,14]. Host resistance may also be required to replace the loss of some integrated management tools. Use of copper and streptomycin sprays over the years, for example, has led to the development of resistant strains [5]. At the same time, genetic resistance has been lost due to race shifts in pathogen populations [15-17]. Designing new and possibly durable resistance requires knowledge of pathogenicity factors possessed by the four groups.

Many candidate pathogenicity factors have been identified in strains of Xanthomonas. A number of virulence factors are employed by xanthomonads to gain entry into leaf or fruit tissue, and gain access to nutrients, while simultaneously overcoming or suppressing plant defenses. Different secretion systems and their effectors have been shown to contribute to the virulence of plant pathogens. The type III secretion system (T3SS) encoded by the hrrp (Hypersensitive Response and Pathogenicity) gene cluster [18,19] and type III secreted effectors have been widely studied for their role in hypersensitivity and pathogenicity. Effectors common between strains are believed to be responsible for conserved virulence function and avoidance of host defense. Differences in effector suites have evolved in closely related strains of plant pathogens and strain-specific effectors may help to escape recognition by host-specific defenses [20-25]. Important insights into pathogenicity mechanisms of X. euvesicatoria strain 85-10 (hereafter, Xcv) have been obtained with its genome sequence [26]. Here we report draft genome sequences of type strains of the other three bacterial spot pathogen species: X. vesicatoria strain 1111 (Xv 1111) (ATCC 35937), X. perforans strain 91-118 (Xp 91-118), and X. gardneri strain 101 (Xg 101) (ATCC 19865). We have annotated and analyzed predicted pathogenicity factors in the draft genomes. Additionally, we have investigated differentiation between xanthomonads that might explain differences in disease phenotypes and in host range.

**Results and Discussion**

**Draft genome sequences of Xv strain 1111, Xp strain 91-118 and Xg strain 101 were obtained by combining Roche-454 (pyrosequencing) and Illumina GA2 (Solexa) sequencing data**

Initially, we sequenced Xv strain 1111 (ATCC 35937) (hereafter Xv), Xp strain 91-118 (hereafter Xp) and Xg strain 101 (ATCC 19865) (hereafter Xg) by 454 pyrosequencing [27]. *De novo* assembly using Newbler assembler resulted in 4181, 2360 and 4540 contigs, respectively, for Xv, Xp and Xg, with approximately 10-fold coverage for each strain (Additional file 1: Table S1). Many pathogenicity genes, including type III effectors, existed in the form of fragments given the relatively low coverage of the 454-based assembly. More complete assemblies were obtained using Illumina sequencing [28]. *De novo* assemblies of around 100-fold coverage were constructed from the Illumina data alone or combined with pre-assembled 454 long reads using CLC Genomic Workbench [29]. Combined 454 and Illumina sequencing produced a much better assembly than either technology alone (Table 1). Therefore, combined assemblies were chosen for all subsequent analyses. The average contig size in the combined 454 and Illumina assemblies was around 18 kb for Xv and Xp, and 10 kb for Xg. The N50 (minimum number of contigs needed to cover 50% of the assembly) values were 37 and 40 for Xv and Xp, respectively, and 83 for Xg indicating that final assemblies consist of a few large contigs allowing reasonably accurate whole genome comparisons. The three strains were deduced to contain plasmids as evidenced by the presence of genes that are known to be involved in plasmid maintenance (e.g. parB/F genes). We have used adjacency to such genes to infer occurrence of certain other genes on plasmids.
Relationships of the strains to other xanthomonads using whole genome comparisons
16S rRNA analysis and MLST-based phylogenetic analysis showed the diversity among the four bacterial spot species. We carried out phylogenetic analysis based on orthologous protein-coding genes from draft genomes and reference xanthomonads (Figure 1). Whole genome comparisons were performed using the MUMi index [30] to assess pairwise distance between the draft genomes and available reference Xanthomonas genomes as shown in the phylogenetic tree and the distance matrix (Additional file 2: Fig. S2). Another program, dnadiff, based on nucmer [31] showed the extent of homologies among the shared regions of the genomes by pairwise comparisons (Additional file 3: Table S3). All of the methods yielded consistent results: we were able to ascertain that among the three newly sequenced strains in relationship to the previously sequenced strains, Xp and Xcv form the closest pair, which is in turn closest to Xac. Next, Xg is closest to Xcc, with Xv forming a clade with Xg and the Xcc species group (Figure 1, Additional files S2 and S3).

Four xanthomonads show variation in the organization of the type III secretion gene clusters
Annotation of the respective type III secretion gene clusters, or hrp genes showed that Xp has an almost identical and syntenic hrp cluster to that of Xcv (Figure 2). The most notable difference is that hpaG and hpaF encode the fusion protein XopAE in Xp, while they are present as separate genes in Xcv. Adjacent hypothetical protein XCV0410 (126 amino acid protein) is absent from Xp. Xv and Xg show greater similarity to the core hrp cluster genes of Xcc than to that of Xcv. Xv and Xg contain hrpW associated with the hrp cluster as in Xcc. Additionally, xopD in Xv and Xg is not associated with the hrp cluster as in Xcc (referred to as psv in Xcc). PsvA shows 74% and 84% sequence identity to the respective homologs from Xv and Xg. XopA (hpa1) from Xcv seems to be absent from Xv and Xg. Interestingly, we found a novel candidate effector gene (named xopZ2) upstream of hrpW in Xv and Xg (See below, Additional file 4: Fig. S4).

Table 1 General sequencing and combined (454 and solexa) de novo assembly features of draft genomes of Xv, Xp and Xg

<table>
<thead>
<tr>
<th></th>
<th>Xanthomonas vesicatoria (Xv)</th>
<th>Xanthomonas perforans (Xp)</th>
<th>Xanthomonas gardneri (Xg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of contigs</td>
<td>296</td>
<td>291</td>
<td>552</td>
</tr>
<tr>
<td>N50*</td>
<td>37</td>
<td>40</td>
<td>83</td>
</tr>
<tr>
<td>Mean contig length</td>
<td>18,686</td>
<td>18,082</td>
<td>10,014</td>
</tr>
<tr>
<td>Longest contig</td>
<td>153,834</td>
<td>133,836</td>
<td>88,536</td>
</tr>
<tr>
<td>Total length of contigs</td>
<td>5,531,090</td>
<td>5,262,127</td>
<td>5,528,125</td>
</tr>
</tbody>
</table>

*N50 - number of contigs that cover 50% of the genome assembly.

Figure 1 Maximum likelihood tree based on orthologous genes from xanthomonads and Stenotrophomonas
Concatenated amino acid sequences of the orthologous genes from four bacterial spot pathogen strains along with other sequenced xanthomonads were considered in the analysis. Stenotrophomonas maltophilia (Sm) was used as an outgroup. The evolutionary history was inferred using the Maximum likelihood method. The tree is drawn to scale, with branch lengths corresponding to the evolutionary distances. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site.
Finally, the \textit{hrp}-associated effector \textit{xopF1} is conserved and intact in all four tomato and pepper pathogens.

A \textbf{reporter gene assay confirms translocation of novel type III effectors}

We identified and annotated T3SS effectors from the three newly sequenced xanthomonads (See Methods). Several candidate effectors, which had not yet been experimentally confirmed in xanthomonads, and candidate effectors with plausible translocation motifs were identified (Tables 2, 3, and 4). Corroborative evidence for T3SS-mediated translocation of the candidate effectors was assessed by constructing fusion genes with the C-terminal end of \textit{AvrBs2} coding sequence (\textit{avrBs2}$_{62-574}$aa) in a race 6 strain of \textit{X. euvesicatoria}. Translocation was measured in pepper cv. ECW 20R, containing the resistance gene \textit{Bs2} (Additional file 4: Fig. S4). Genes \textit{xopAO}, \textit{xopG}, \textit{xopAM}, and \textit{XGA}_0724 (belonging to the \textit{avrBs1} class of effectors), of which homologs were previously found in \textit{Pseudomonas} species, were demonstrated to direct \textit{AvrBs2}-specific hypersensitive reactions in ECW 20R (Tables 3, Table 4, Additional file 4: Fig. S4). Another candidate effector gene \textit{xopZ2}, associated with the \textit{hrp} clusters in \textit{Xv} and \textit{Xg} (Figure 2), was also functional in the \textit{AvrBs2}-based assay. Thus, we identified five effectors (\textit{xopAO}, \textit{xopG}, \textit{xopAM}, \textit{xopZ2}, \textit{XGA}_0724) that have not been previously recognized in \textit{Xanthomonas} and showed their functionality.

\textbf{Core effectors among four xanthomonads give insight into infection strategies of the pathogen}

Comparing the draft genome sequences of the three xanthomonads with that of \textit{Xcv} allowed us to identify the core effectors conserved in all four strains as well as strain-specific effectors (Tables 2, 3, and 4).
### Table 3 Type III effectors specific to each species

<table>
<thead>
<tr>
<th>Effector class</th>
<th>Effector Locus tags</th>
<th>Effector homolog</th>
<th>Pfam domains/biochemical motifs</th>
<th>Comments/Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>XopJ2</td>
<td>XVE_4395, XVE_3709</td>
<td>AvrBsT</td>
<td>CSS-family cysteine protease or Ser/Thr acetyltransferase</td>
<td>[40]</td>
</tr>
<tr>
<td>XopAG</td>
<td>XVE_2415</td>
<td>AvrGf1</td>
<td>-</td>
<td>[39]</td>
</tr>
<tr>
<td>XopAI</td>
<td>XVE_4756</td>
<td>XAC3230</td>
<td>-</td>
<td>[25]</td>
</tr>
<tr>
<td>AvrB8s1</td>
<td>XGA_0724</td>
<td>AvrA (84% identity)</td>
<td>-</td>
<td>This study</td>
</tr>
<tr>
<td>AvrHah1</td>
<td>XGA_3187</td>
<td>AvrBs3</td>
<td>Transcriptional activator, nuclear localization</td>
<td>AvrBs3 present in few euvesicatoria strains [41].</td>
</tr>
<tr>
<td>XopAO</td>
<td>XGA_1250</td>
<td>AvrRpm1 (61% identity)</td>
<td>-</td>
<td>This study</td>
</tr>
<tr>
<td>XopAQ</td>
<td>XGA_2091</td>
<td>Rip6/rip11</td>
<td>No known domains</td>
<td>[46]</td>
</tr>
<tr>
<td>XopAS</td>
<td>XGA_0764/0765</td>
<td>HopA51</td>
<td>No known domains</td>
<td>This study</td>
</tr>
<tr>
<td>XopC2</td>
<td>XPE_3585</td>
<td>Rsp1239</td>
<td>Halochlorophyllase-like hydrolase</td>
<td>[24]</td>
</tr>
<tr>
<td>XopJ4</td>
<td>XPE_1427</td>
<td>AvrXv4</td>
<td>SUMO protease (experimental); YopJ-like serine threonine acetyl transferase domain (predicted)</td>
<td>[38, 105]</td>
</tr>
<tr>
<td>XopAF</td>
<td>XPE_4185</td>
<td>AvrXv3</td>
<td>Transcriptional activator domain</td>
<td>[37]</td>
</tr>
<tr>
<td>XopAE</td>
<td>XPE_2919</td>
<td>HpaF/G</td>
<td>LRR protein</td>
<td>[24]</td>
</tr>
<tr>
<td>AvrBs1</td>
<td>XCV0052</td>
<td>AvrBs1</td>
<td>-</td>
<td>[26]</td>
</tr>
<tr>
<td>XopC1</td>
<td>XCV2435</td>
<td>XopC</td>
<td>Phosphoribosyl transferase domain and haloacid dehalogenase-like hydrolase</td>
<td>[105]</td>
</tr>
<tr>
<td>XopJ1</td>
<td>XCV2156</td>
<td>XopJ</td>
<td>CSS-family cysteine protease or Ser/Thr acetyltransferase</td>
<td>[105]</td>
</tr>
<tr>
<td>XopJ3</td>
<td>XCV0471</td>
<td>AvrXv7</td>
<td>CSS-family cysteine protease or Ser/Thr acetyltransferase</td>
<td>[26]</td>
</tr>
<tr>
<td>XopO</td>
<td>XCV1055</td>
<td>Unknown</td>
<td>-</td>
<td>[26]</td>
</tr>
<tr>
<td>XopAA</td>
<td>XCV3785</td>
<td>ECF</td>
<td>Early chlorosis factor, proteasome/cyclosome repeat</td>
<td>[26]</td>
</tr>
<tr>
<td>XopAI</td>
<td>XCV4428</td>
<td>AvrRxo1</td>
<td>-</td>
<td>[26]</td>
</tr>
</tbody>
</table>
At least 11 effector genes form a core set of common effectors for xanthomonads infecting tomato and pepper (Table 2). Of these 11, eight effector genes (avrBs2, xopK, xopL, xopN, xopQ, xopR, xopX and xopZ) were found to be conserved in all sequenced xanthomonads including the three draft genomes presented here with the exceptions of X. albilineans and X. campestris pv. armoraciae. These genes might be necessary for maintaining pathogenicity of these xanthomonads in a wide range of host plants. XopN has been reported to suppress PAMP (pathogen-associated molecular pattern)-triggered immunity by interacting with tomato TARK1 and TFT1 [32]. XopF1 is conserved in tomato and pepper xanthomonads. Although a homolog of xopF1 is found in Xcc, the respective gene is truncated [34]. Hence, xopF1 is a potential pathogenicity determinant in tomato. A xopF1 deletion mutant of Xcv did not show any difference in virulence when compared to wild type Xcv on the susceptible cultivar of pepper ECW, suggesting XopF1 is not the lone factor for pathogenicity of Xcv on pepper [33]. Another effector gene, xopD, is associated with the hrp gene cluster in Xcv and Xp. However, xopD appears to have translocated to another location in the genome in case of Xg, Xv and Xcc strains. XopD is annotated as "Psv virulence protein" in Xcc genome [34] and has been shown to be a chimeric protein sharing a C terminus with XopD from Xcv [35]. Although xopD homologs from Xv and Xg are syntenic with the psv gene in Xcc, Xv and Xg have intact full-length copies of xopD as in Xcv, indicating that the xopD could be another effector exclusive to the tomato pathogens and a possible pathogenicity determinant in tomato. XopD has been shown to enhance pathogen survival in tomato leaves by delaying symptom development [36]. Two tandem copies of xopX are found in Xg. However, one gene in Xg appears to be inactive due to a frameshift mutation. In Xp, the two copies of xopX are found in different locations in the genome with neighboring genes, including chaperone gene groEL, which is also duplicated. Orthologs of xopZ are also found in all four xanthomonads, with 82% identity for Xcv and Xp and 35% identity for Xg and Xv. Apart from low sequence identity in Xv and Xg, gene-specific rearrangements appear to have occurred within each ortholog. We propose that the overall low amino acid relatedness (pairwise sequence identities below 50%) of this effector
in Xv and Xg warrants assigning the proteins to a new family within the xopZ class, named xopZ2, while the orthologs from Xcv and Xp belong to family of xopZ1 as originally described in Xoo and as supported by pairwise sequence identities of at least 60% (see above, Figure 2, Table 4).

**Effectors unique to Xp might be responsible for restricting growth on pepper**

Xp is pathogenic only on tomato. The avirulence gene, avrXv3, present in Xp, was previously shown to elicit an hypersensitive response (HR) in pepper cv. ECW [37]. An avrXv3 knockout mutant of Xp is not virulent in pepper cv. ECW indicating that other factors are associated with host specificity. Comparing effector repertoires of the pepper pathogens Xg, Xcv, and Xv with Xp may provide clues to the factors that are responsible for reduced virulence (Table 4). Besides avrXv3, the only effectors present in Xp and absent or inactive in Xg, Xv and Xcv are xopC2, xopAE and xop4 (avrXv4) (Table 3). The gene avrXv4 is absent from other sequenced xanthomonads and shows gene-for-gene interaction with the Xv4 resistance gene from the wild tomato relative Solanum pennellii but does not contribute to restricted growth of Xp on pepper [38]. The effector xopC2 is a homolog of the effector rspi239 from Ralstonia solanacearum GMI1000 and xopAE encodes a LRR protein with homology to the R. solanacearum effector PopC. Both genes, xopC2 and xopAE, are truncated in Xcv. Therefore, these two effectors may trigger immunity in pepper. Interestingly, Xp contains a paralog of xopP. The two copies are found next to each other in the genome and share 75% identity at the amino acid level. The second copy is next to the candidate effector xopC2, which is unique to Xp among tomato and pepper pathogens. Effectors xopC2 and xopP may both act to restrict growth in pepper. Moreover, there are at least two effectors, xopE2 and xopG, present in the pepper pathogens Xcv, Xv and Xg but absent from Xp. These effectors may be essential pathogenicity factors in pepper.

**Species-specific effectors**

Xv possesses two unique effector genes, xopAG (avrGf1) and xopAI (Table 3). A phylogenetic analysis of xopAG from Xv is closely related to xopAG from X. citri A*, which has been shown to be responsible for causing an HR on grapefruit [39]. XopAI is a chimeric protein, which contains a conserved myristoylation motif at its N terminus, like XopJ1. This effector class also includes the homolog XAC3230 from Xac as well as XAUB_26830 and XAUC_23780 from X. fuscans subsp. aurantiifolii strains B and C, respectively [25]. The presence of transposons and phage elements in close proximity helps to explain the evolution of this novel effector in Xac by terminal reassortment [35]. Xv also contains effector gene avrBsT, which is responsible for the hypersensitive response on pepper. Loss of the plasmid containing avrBsT in Xcv strain 75-3 allows the strain to cause disease on pepper [40].

Xg contains at least two effectors, avrHah1 (an avrBs3-like effector gene) and xopB as does Xcv, and share sequence identity of 82% and 86% respectively to the corresponding effectors of Xcv. However, AvrHah1 appears to specify a different phenotype when compared to avrBs3 from Xcv. AvrHah1 was shown to be responsible for increased watering on pepper ECW-50R and 60R, whereas Xcv strains carrying avrBs3 show a phenotype that consists of small raised fleck lesions on pepper [41]. Another effector gene, xopB, has a PIP box at the 5’ end in Xcv, whereas the homolog in Xg does not contain a PIP box. Neighboring genes to xopB in the respective strains are completely different between genomes, suggesting lack of synteny between the two species in this region (Table 4). XopB from Xg is 92% identical at the amino acid level to the homolog in Xcv. Deletion mutants of xopB from Xcv did not show any difference in virulence, indicating it does not contribute significantly to virulence [42]. However, xopB may contribute to virulence in Xg. We also identified eight effector genes that are unique to Xcv (Table 3). With the exception of xopAA (early chlorosis factor), all of these genes belong to regions of low GC content compared to average genome GC content (64.75%): avrBsI (42%), xopC1 (48%), xopI (xopI) (57%), xopI3 (avrXv3) (52%), xopO (52%), xopAI (avrRxo1) (51%).

**Few effectors are shared among phylogenetically related group strains**

Although Xp and Xcv, and Xv and Xg form distinct phylogenetic groups (Figure 1), relatively few effectors are shared between these species. For Xp and Xcv, they share at least six effectors - xopE1, xopF2, xopP, xopV, xopAK, xopAP, which are absent from the other two genomes (Table 4). Xv and Xg appear to be most closely related to strains of X. campestris pv. campestris, and this relationship is reflected in the suite of effector genes. In fact, Xg and Xv share four effector genes with Xcc, namely, xopAM, avrXccA1, hrpW and xop22, with the caveat that hrpW and avrXccA1 may not function as intracellular effectors (Table 4). Furthermore, the genomic regions containing these genes are syntenic in Xg, Xv and Xcc.

**X. gardneri shows evidence of effector acquisition by horizontal gene transfer**

Effector homologs of avrA, hopAS1 and avrRpm1 from P. syringae pv. tomato T1 and P. syringae pv. syringae...
B728a are found in Xg with 79%, 41% and 61% identity at the amino acid level, respectively (Table 3, Additional file 4: Fig. S4). Other X. gardneri strains also contain these effectors based on PCR screening (data not shown). These three effectors, XGA_0724 (belonging to avrBs1 class), XGA_0764/XGA_0765 (xopAS) and XGA_1250 (xopAO), are unique to X. gardneri. The C terminal region of XGA_0724 shows 53% identity to avrBs1 from Xcv. Hence according to the Xanthomonas effector nomenclature [24], XGA_0724 from Xg was placed under the class avrBs1. XGA_0764/XGA_0765 and XGA_1250 have not yet been reported to be found in xanthomonads and were assigned to new classes xopAS and xopAO. X. gardneri strains have been found to be associated with tomato and have a lower optimum temperature for disease development similar to that of pathovars of Pseudomonas syringae [43]. A high score by Alien_hunter analysis [44], along with very low GC content (45% for XGA_0724 and 48% for XGA_01250, 59% for XGA_0764/XGA_0765) and the proximity of mobile genetic elements provides evidence for horizontal gene transfer (Additional file 5: Table S5). Effector xopAS appears to be separated into two ORFs XGA_0764 and XGA_0765 by internal stop codon. The functionality of effector xopAS needs to be confirmed by in planta reporter gene assay. AvrA of P. syringae pv. tomato PT23 was shown to contribute to virulence on tomato plants [45]. Acquisition of XGA_0724 by Xg might have conferred increased virulence on tomato. AvrRpm1 from P. syringae pv. syringae possesses a myristoylation motif, which is absent from homologs in Xg. This modification in Xg might have been acquired to escape host recognition. Another candidate effector gene, xopAQ, in Xg is found 68 bps downstream of a perfect PI P box. The gene shows 65% identity at the amino acid level to rip6/11, a novel effector from R. solanacearum RS1000 [46].

All four xanthomonads contain Ax21 coding gene but only Xcv contains a functional sulfation gene

The ax21 (activator of AX21-mediated immunity) gene is conserved among Xanthomonas species and is predicted to encode a type I-secreted protein that may serve as a quorum sensing signaling molecule [47]. A 17-amino acid sulfated peptide from the N-terminal region of Xanthomonas oryzae pv. oryzae (Xoo) Ax21 (axYS22) was shown to bind and activate the AX21 receptor kinase from rice, demonstrating that Ax21 is a conserved PAMP that can activate plant immune signaling [48]. The ax21 gene is present in Xcv (93% identity with Xoo PXO99 protein), Xp (94%), Xv (91%), and Xg (88%). The axYS22 peptide is 100% conserved in Xcv, Xp and Xv, while in Xg there is a change from leucine to isoleucine at residue 20; this is unlikely to alter the activity of the peptide, since changing this residue to alanine had no effect on recognition by AX21 [48].

Recognition of axYS22 by the AX21 receptor requires sulfation of tyrosine 22, which requires the putative sulfotransferase RaxST. In contrast to ax21, the raxST gene is more variable in these genomes, which is consistent with a report of sequence differences in this gene among Xoo strains [49]. Furthermore, in Xp, there is a single-nucleotide insertion at position 65, causing a frameshift mutation. The Xv and Xg genomes do not contain raxST; therefore, the ax21 gene products may be nonfunctional in these strains. These findings have implications for the further study of the role of Ax21 in quorum sensing and virulence, as well as for the usefulness of the AX21 receptor to confer resistance to xanthomonads in crop plants.

Two type II secretion systems are conserved in all four Xanthomonas genomes

Most cell-wall degrading enzymes, such as cellulases, polygalacturonases, xylanases, and proteases, are secreted by a type II secretion system (T2SS). The Xps T2SS, present in all xanthomonads, has been studied for its contribution to virulence in Xcc and Xoo [50,51]. Another T2SS cluster, known as the Xcs system, is found only in certain species of Xanthomonas, e.g. Xcc, Xac, and Xcv. The Xps system secretes xylanases and proteases and is under control of hrpG and hrpX [52], indicating differential regulation. Both Xps and Xcs systems are present in all three draft genomes.

Xanthomonads possess diverse repertoires of cell-wall degrading enzymes, which are present in diverse genomic arrangement patterns

Each species of Xanthomonas has its own collection of genes encoding endoxylanases, endoglucanases, and pectate lyases which contribute to cell wall deconstruction during pathogenesis. We have compared these repertoires from the three draft genomes and other xanthomonads as detailed in Table 5. The genes are designated for different families of glycosyl hydrolases (GH) and polysaccharide lyases (PL) that include the enzymes that cleave glycosidic bonds in the structural polysaccharides of plant cell walls.

Genes encoding secreted endoxylanases regulated by the xps genes have been described for their contributions to virulence, including XCV0965 [52] encoding GH30 endoxylanase. The GH30 family catalyses the cleavage of methylglucuronoxylans in the cell walls of monocots and dicots at a β-1,4-xyllosidic bond penultimate to one linking the xylose residue that is substituted by an α-1,2-linked 4-O-methylglucuronate residue [53,54]. Such an enzyme secreted by Erwinia chrysanthemi generates oligosaccharides that are not
assimilated for growth, suggesting a function in which it contributes to cell wall deconstruction for access to pectates for growth substrate [53]. It is interesting to note the orthologous genes encoding GH30 enzymes are absent in Xg and Xv, with a truncated xyn30 gene in Xac. On the basis of sequence homology, xyn30 genes may also contribute to virulence in Xoo, Xcc and Xp.

The more common GH10 endoxylanases, which occur in several bacterial and fungal phyla, have been implicated in the virulence of plant pathogenic bacteria and fungi [55,56]. In Xoo, deletion of the gene encoding a GH10 xyn10B resulted in diminished virulence [57]. All sequenced Xanthomonas genomes contain either two or three copies of xyn10 genes, all of which are within a gene cluster that may comprise a single operon (Figure 3). The GH10 endoxylanases are the best studied of all of the xylanases, and structure/function relationships may be inferred on the basis of gene sequence. The action of these enzymes on glucuronoxylans generates xylotriose, xylobiose, and small amounts of xylose that generally serve as substrates for growth. Also generated is methylglucuronoxylotriose, that is formed to the extent that xylose residues in the β-1,4 xylan backbone are substituted with α-1,2-linked 4-O-methylglucuronate residues [58].

An adjacent gene cluster in an opposite orientation contains agu67 gene encoding a GH67 α-glucuronidase that serves to catalyze the removal of 4-O-methylglucuronate from the reducing terminus of methylglucuronoxylotriose. This activity provides a synergistic function to the overall xylanolytic process to generate xylotriose, which is converted to xylose by xylanases and xylosidases for complete metabolism [59]. The coregulation of operons encoding XynB and Agu67 enzymes occurs as a logical condition to coordinate expression of genes that encode these and additional enzymes that collectively process glucuronoxylans and glucuronoarabinoxylans for complete metabolism. The accessory enzymes and transporters necessary for the function of these enzymes are embedded within these operons in Gram positive bacteria [60-62] and share similarities noted here with Xanthomonas spp.. These include the genes encoding two glycohydrolases, a β-xylosidase and an α-L-arabinofuranosidase. Also included in this cluster are genes encoding enzymes for intracellular metabolism of

Table 5 Repertoire of cell wall degrading enzymes in xanthomonads

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Family</th>
<th>Enzymatic function</th>
<th>Xp</th>
<th>Xac</th>
<th>Xcv</th>
<th>Xv</th>
<th>Xg</th>
<th>Xcc strain 33913</th>
<th>Xoo strain KACC</th>
</tr>
</thead>
<tbody>
<tr>
<td>xyn10A</td>
<td>GH10</td>
<td>Endo-β-1,4-xylanase EC.3.2.1.8</td>
<td>2014</td>
<td>4254</td>
<td>4360</td>
<td>2337</td>
<td>1172</td>
<td>4118</td>
<td>4429</td>
</tr>
<tr>
<td>xyn10B</td>
<td></td>
<td></td>
<td>2016</td>
<td>4252</td>
<td>4358</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4428</td>
</tr>
<tr>
<td>xyn10C</td>
<td></td>
<td></td>
<td>2020</td>
<td>4249</td>
<td>4355</td>
<td>2333</td>
<td>0341</td>
<td>4115</td>
<td>-</td>
</tr>
<tr>
<td>aguA</td>
<td>GH67</td>
<td>α-glucuronidase EC.3.2.1.139</td>
<td>4318</td>
<td>4227</td>
<td>4333</td>
<td>4712</td>
<td>2473</td>
<td>4102</td>
<td>4419</td>
</tr>
<tr>
<td>xyn51A</td>
<td>GHS1</td>
<td>β-D-Arabinofuranosidase EC.3.2.1.55</td>
<td>0180</td>
<td>1286</td>
<td>1335</td>
<td>1029/1030</td>
<td>2303</td>
<td>1191</td>
<td>1317</td>
</tr>
<tr>
<td>xyn5A</td>
<td>GH5</td>
<td>Endo-β-1,4-xylanase EC.3.2.1.8</td>
<td>4682</td>
<td>0933/34 partial</td>
<td>0965</td>
<td>-</td>
<td>0857</td>
<td>3618</td>
<td></td>
</tr>
</tbody>
</table>

Glucanases

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Family</th>
<th>Enzymatic function</th>
<th>Xp</th>
<th>Xac</th>
<th>Xcv</th>
<th>Xv</th>
<th>Xg</th>
<th>Xcc strain 33913</th>
<th>Xoo strain KACC</th>
</tr>
</thead>
<tbody>
<tr>
<td>cel8A</td>
<td>GH8</td>
<td>Endo-1,4 β-D glucanase</td>
<td>1965</td>
<td>3516</td>
<td>3641</td>
<td>0432</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>cel9A</td>
<td>GH9</td>
<td></td>
<td>2345</td>
<td>2522</td>
<td>2704</td>
<td>1327</td>
<td>0588</td>
<td>2387</td>
<td>-</td>
</tr>
</tbody>
</table>

Pectate lyases

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Family</th>
<th>Enzymatic function</th>
<th>Xp</th>
<th>Xac</th>
<th>Xcv</th>
<th>Xv</th>
<th>Xg</th>
<th>Xcc strain 33913</th>
<th>Xoo strain KACC</th>
</tr>
</thead>
<tbody>
<tr>
<td>pel1A</td>
<td>PL1</td>
<td>Pectate lyase EC.4.2.2.2</td>
<td>3841</td>
<td>3562</td>
<td>3687</td>
<td>1933</td>
<td>4024</td>
<td>0645</td>
<td>0821</td>
</tr>
<tr>
<td>pel1B</td>
<td></td>
<td></td>
<td>1563</td>
<td>2986</td>
<td>3132</td>
<td>3512</td>
<td>0893</td>
<td>2815</td>
<td>-</td>
</tr>
<tr>
<td>pel1C</td>
<td></td>
<td></td>
<td>-</td>
<td>2373</td>
<td>2569</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>pel3A</td>
<td>PL3</td>
<td>Pectate lyase EC.4.2.2.2</td>
<td>-</td>
<td>2922</td>
<td>-</td>
<td>3222</td>
<td>2761</td>
<td>1219</td>
<td>-</td>
</tr>
<tr>
<td>pel4A</td>
<td>PL4</td>
<td>Rhamnogalacturonan lyase EC.4.2.2.2</td>
<td>1975</td>
<td>3505</td>
<td>3632</td>
<td>2592</td>
<td>4531</td>
<td>3377/78/79</td>
<td>1078</td>
</tr>
<tr>
<td>pel9A</td>
<td>PL9</td>
<td>Pectate lyase EC.4.2.2.2</td>
<td>-</td>
<td>-</td>
<td>2278</td>
<td>1927</td>
<td>1853</td>
<td>-</td>
<td>2265</td>
</tr>
<tr>
<td>pel10A</td>
<td>PL10</td>
<td>Pectate lyase EC.4.2.2.2</td>
<td>-</td>
<td>-</td>
<td>4069</td>
<td>5124</td>
<td>0122</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Different cell wall degrading enzymes, such as xylanase, pectate lyase, glucanases, were compared for their repertoires among already sequenced xanthomonads including our draft genomes. Genes identified by their locus tags in the respective genomes.
glucuronate and xylose, including glucuronate isomerase; xylose isomerase; D-mannonate dehydratase; and D-mannonate oxidoreductase. Genes encoding mannitol dehydrogenase and the hexuronic transporter, as well as the TonB-dependent receptor and LacI transcriptional regulator, flank these two operons.

The arrangement and content of xylanolytic enzymes differentiate \textit{Xanthomonas} species into three groups (Figure 3). Here, we propose a common nomenclature for xylanases, the genes for which have been annotated in the sequenced genomes. Members of the first group are \textit{Xac}, \textit{Xcv} and \textit{Xp} in which all three genes encoding GH10 endoxylanases (\textit{xyn10A}, \textit{xyn10B} and \textit{xyn10C}) are present, and with additional genes further downstream in this cluster. Members of the second group are \textit{Xcc}, \textit{Xv} and \textit{Xg} in which genes encoding two of the three endoxylanases are present (\textit{xyn10A} and \textit{xyn10C}) and where one or more of the downstream genes are absent. \textit{Xoo} strains represent a third group in which a different set of two endoxylanase encoding genes are present (\textit{xyn10A} and \textit{xyn10B}) and where the \(\beta\)-galactosidase and gluconolactonase genes flanking \textit{xyn10C} are absent. It is noteworthy that the organization of genes in the cluster encoding the \(\alpha\)-glucuronidase is conserved across \textit{Xanthomonas} species.

\textbf{Genes involved in several Type IV secretion systems are present in genomes and plasmids}

Like \textit{Xcv}, the tomato pathogens, \textit{Xg}, \textit{Xv} and \textit{Xp}, also appear to contain more than one copy of a type IV secretion system (T4SS) cluster (Figure 4A, B). Two T4SS clusters (Vir and Dot/Icm type) are present in \textit{Xcv}, and genes belonging to both of these systems are found on plasmids [26]. The Dot/Icm type system is absent from \textit{Xv}, \textit{Xp} and \textit{Xg}.

In \textit{Xv} and \textit{Xp}, genes for one T4SS are on a plasmid and the second one on the chromosome while in \textit{Xg}, two T4SS gene clusters are on a plasmid and one is on the chromosome. The two T4SS clusters on plasmids of \textit{Xg} do not show any similarity to the genes for T4SS in \textit{Xac}, \textit{Xcv}, \textit{Xcc} and \textit{Xoo}. Of the two T4SS clusters in \textit{Xg},
one is also found in \( Xv \) and \( Xp \). This cluster appears to be exclusive to these three tomato pathogens (Figure 4A). The genes belonging to this cluster show low (30-45%) identity to the T4SS clusters from \textit{Ralstonia}, \textit{Burkholderia}, \textit{Bradyrhizobium}, and \textit{Stenotrophomonas maltophila}. The other cluster from \( Xg \), which is absent from \( Xv \) and \( Xp \), shows very high identity (98%) and synteny to the T4SS cluster of \textit{Burkholderia multivorans} and around 89% identity to a T4SS cluster of \textit{Acidovorax avenae} subsp. \textit{citrulli} (Figure 4B).

Apart from the plasmid borne T4SS genes, \( Xcv \) also contains a portion of a type IV system cluster on the chromosome and consists of \textit{VirB6}, \textit{VirB8}, \textit{VirB9}, \textit{VirD4} genes. This chromosomal cluster is flanked by a transposon element (IS1477) that might indicate its horizontal gene transfer. \( Xp \), \( Xg \) and \( Xv \) genomes contain a complete chromosomal T4SS cluster showing high identity to the T4SS chromosomal clusters from \( Xcc \) (Figure 4C).

**Type V secreted adhesins function in synergism during pathogenesis**

Different adhesins have been shown to function at different stages of the infection process starting with attachment, entry, later survival inside host tissue and colonization by promoting virulence [63,64]. FhaB hemagglutinin, important for leaf attachment, survival inside plant tissue and biofilm formation, is present in all four tomato pathogens. In \( Xcv \), \( fhaB \) is divided into two separate open reading frames, XCV1860 and XCV1861, with the two-partner secretion domains being present in all four tomato pathogens. In \( Xoo \), \( xadA \) and \( xadB \) promote virulence by enhancing colonization of the leaf surface and leaf entry through hydathode [64]. As in \( Xcv \) and \( Xac \), \( Xp \) encodes

---

**Figure 4** Type IV secretion system. A) Schematic representation of type IV secretion system cluster common to \( Xp \), \( Xv \) and \( Xg \) (Plasmid borne); B) Type IV cluster unique to \( Xg \) (plasmid borne); C) Chromosomal type IV cluster organization in \( Xcv \), \( Xv \), \( Xp \) and \( Xg \).
two copies of xadA, while Xv and Xg possess a single ortholog of xadA as does Xcc. YapH and the type IV pilus protein PilQ were shown to be involved in virulence in Xoo during later stages of growth and migration in xylem vessels. In Xcv, Xc, and XooKACC, two copies of yapH are present. There are two pilQ orthologs in Xcv and only one in other sequenced xanthomonads. Next to the fhaB and fhaC adhesin genes, hms operon is present in the genomes of xanthomonads, the homologs of which are pga operon genes in E. coli involved in biofilm formation [65].

Type VI secretion system is present in Xcv, Xv and Xp
Type VI secretion system (T6SS) has been shown recently to contribute to host pathogen interactions during pathogenesis in Vibrio cholerae, Burkholderia pseudomallei and Pseudomonas aeruginosa. Hcp (Haemolysin-coregulated protein) and Vgr (valine-glycine repeats) proteins are exported by the T6SS [66]. T6SS clusters can be assigned to three different types in xanthomonads (Table 6). Xcv and Xp possess two types of T6SSs (type 1 and 3); whereas Xv contains only a single type of T6SS, type 3. As in Xcc, there is no T6SS cluster in Xg (Table 6, Additional file 6: Table S6).

LPS locus displays remarkable variation in sequence and number of coding genes and shows host specific variation
The lipopolysaccharide (LPS) biosynthesis cluster has been studied in detail in Xcc [67], which comprises three regions; region 1 from wxcA to wxcE involved in biosynthesis of water soluble LPS antigen; region 2 (gnd, rmd) coding for LPS core genes; and region 3 from wxcK to wxcO coding for enzymes for modification of nucleotide sugars and sugar translocation systems. This LPS biosynthesis locus is positioned between highly conserved housekeeping genes, namely cystathionine gamma lyase (metB) and electron transport flavoprotein (etfA), as reported in other xanthomonads [68]. Comparison of this cluster from draft genomes to the already sequenced xanthomonads revealed high variability in the number of genes and their sequences. Xv and Xg have an identical type of LPS gene cluster of 17.7 kb encoding 14 open reading frames (Figure 5A) which is similar in organization and sequence identity to the LPS locus from Xcc strains. Interestingly, Xg and Xv also contain two glycosyl transferases involved in synthesis of xylosylated polyrhamnan as seen in Xcc [69], in contrast to glycosyl transferases (wbdA1, wbdA2) involved in synthesis of polymannan in Xcv [26]. This suggests that basic structure of O-antigen in Xg and Xv is similar to Xcc. The three tomato/pepper pathogens Xcv, Xv and Xg have retained an ancestral type of LPS gene cluster (Figure 5A and 5B). On the other hand, Xp has acquired a novel LPS gene cluster during the course of evolution and is completely different in sequence and number of genes that are encoded. In Xp, this LPS locus is 17.3 kb long and encodes 12 ORFs, all of which are absent in the corresponding genomic region of Xcv, Xv or Xg. Also the first five ORFs flanking the metB side of the LPS locus in Xp (Figure 5A, ORFs colored in red) showed very low or no identity to region 1 of the LPS

<p>| Table 6 Type VI secretion clusters in different xanthomonads |
|-----------------------------------------------|---------------|---------------|-----------|</p>
<table>
<thead>
<tr>
<th>Strain</th>
<th>T6SS #1</th>
<th>T6SS #2</th>
<th>T6SS #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phosphorylation-type regulators:</td>
<td>Kinase/Phosphatase/Forkhead</td>
<td>-</td>
<td>Kinase/Phosphatase/Forkhead</td>
</tr>
<tr>
<td>AraC-type regulators:</td>
<td>-</td>
<td>-</td>
<td>AraC</td>
</tr>
<tr>
<td>Xvv NCPPB702</td>
<td>YES</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Xcm NCPPB4381</td>
<td>YES</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Xaub</td>
<td>/</td>
<td>/</td>
<td>YES</td>
</tr>
<tr>
<td>Xauc</td>
<td>/</td>
<td>/</td>
<td>YES</td>
</tr>
<tr>
<td>Xac</td>
<td>/</td>
<td>/</td>
<td>XAC4116 - XAC4148</td>
</tr>
<tr>
<td>Xv</td>
<td>YES</td>
<td>/</td>
<td>YES</td>
</tr>
<tr>
<td>Xp</td>
<td>YES</td>
<td>/</td>
<td>YES</td>
</tr>
<tr>
<td>Xcv</td>
<td>XCV2120 - XCV2143</td>
<td>/</td>
<td>XCV206 - XCV244</td>
</tr>
<tr>
<td>Xoo KACC10331</td>
<td>XOO3034 - XOO3052</td>
<td>XOO3466 - XOO3517</td>
<td></td>
</tr>
<tr>
<td>Xoo MAFF 311018</td>
<td>XOO2886 - XOO2906</td>
<td>XOO3286 - XOO3319</td>
<td></td>
</tr>
<tr>
<td>Xoo PX099A</td>
<td>XOO0245 - XOO0270</td>
<td>XOO2029 - XOO2060</td>
<td></td>
</tr>
<tr>
<td>Xoc BLS256</td>
<td>XOC2523 - XOC2545</td>
<td>XOC1309 - XOC1370</td>
<td></td>
</tr>
<tr>
<td>Xg</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Xcc ATCC33913</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Xcc 8004</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Xca 756C</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Paxb</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
</tbody>
</table>

Potnis et al. BMC Genomics 2011, 12:146
http://www.biomedcentral.com/1471-2164/12/146
locus in the other xanthomonads. However, these ORFs still belong to the same Pfam families [70] that are usually present in this region, for example, ABC transporters and glycosyl transferases. The second half of the LPS cluster flanking \textit{etfA} side encodes six ORFs, which are homologs of the LPS cluster genes from \textit{Xac}, \textit{Xcm} and \textit{Xoo}. Phylogenetic insight based on conserved \textit{metB} and \textit{etfA} genes that flank the variable LPS locus suggest that the ancestor of all the \textit{Xanthomonas} pathogens of pepper and tomato studied in this paper had the same LPS gene cluster, however putative horizontal gene transfer events at this locus have led to the acquisition of a novel LPS gene cluster in \textit{Xp} (Figure 5B). Alien\textunderscore hunter analysis also supports this acquisition with a high score showing this region to belong to an anomalous region (Additional file 5: Table S5). This event might have played a major role in changing the specificity of \textit{Xp} towards tomato and its dominance over its relative(s) as reported previously [71], similar to variant epidemic strain of \textit{Vibrio cholerae}, reported to be a major reason for its emergence and cholera outbreak during the 1990’s in the Indian subcontinent [72]. Identity in terms of sequences and gene organization among pepper pathogens and absence of those genes from \textit{X. perforans} and a novel LPS cluster in the tomato pathogen \textit{X. perforans} suggest a role of this cluster in host specific variation.

Analysis of DSF cell-cell signaling system

RpfC/RpfG are two-component signaling factors and are involved in DSF (diffusible signal factor) cell-cell signaling [73-76], known to co-ordinate virulence and biofilm gene expression. The genomes of \textit{Xv}, \textit{Xp}, and \textit{Xg} carry an \textit{rpf} (regulation of pathogenicity factors) gene cluster (Table 7) that is found in all xanthomonads and which encodes components governing the synthesis and perception of the signal molecule DSF [74,75]. The Rpf of the DSF system regulates the synthesis of virulence factors and biofilm formation and is required for the full virulence of \textit{Xcc}, \textit{Xac}, \textit{Xoc}, and \textit{Xoo} [77-81]. RpfF is responsible for the synthesis of DSF, whereas, RpfC and RpfG are implicated in DSF perception and signal transduction [73-76]. RpfC is a complex sensor kinase, whereas RpfG is a response regulator with a CheY-like receiver domain that is attached to an HD-GYP domain. HD-GYP domains act in degradation of the second messenger cyclic di-GMP [82]. In addition to genes encoding these products, \textit{Xg} and \textit{Xp} have \textit{rpfH}, which encodes a membrane protein related to the sensory input domain RpfC but whose function is unknown. \textit{Xv} contains \textit{rpfH} but with an internal stop codon, whereas functional \textit{rpfH} is present in \textit{Xv} and \textit{Xcc}, and totally absent in \textit{Xac} and \textit{Xoo}.

Cyclic di-GMP signaling

Cyclic di-GMP is a second messenger known to regulate a range of functions in diverse bacteria, including the virulence of animal and plant pathogens [83-85].

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Xcc8004</th>
<th>Xoo</th>
<th>Xcv</th>
<th>Xv</th>
<th>Xp</th>
<th>Xg</th>
</tr>
</thead>
<tbody>
<tr>
<td>rpfB</td>
<td>XC_2331</td>
<td>XOO2868</td>
<td>XCV1921</td>
<td>2934</td>
<td>0530</td>
<td>2948</td>
</tr>
<tr>
<td>rpfC</td>
<td>XC_2332</td>
<td>XOO2869</td>
<td>XCV1920</td>
<td>2932</td>
<td>0528</td>
<td>2950</td>
</tr>
<tr>
<td>rpfG</td>
<td>XC_2333</td>
<td>XOO2870</td>
<td>XCV1919</td>
<td>2930</td>
<td>0526</td>
<td>2952</td>
</tr>
<tr>
<td>rpfH</td>
<td>XC_2334</td>
<td>Absent</td>
<td>XCV1918</td>
<td>2928/2926*</td>
<td>0524</td>
<td>2954</td>
</tr>
<tr>
<td>rpfI</td>
<td>XC_2335</td>
<td>XOO2871</td>
<td>XCV1917</td>
<td>2924</td>
<td>0522</td>
<td>2956</td>
</tr>
</tbody>
</table>

Table 7 A comparison of \textit{rpf} cluster from \textit{rpfB} to \textit{rpfG} found across a range of \textit{Xanthomonas} genomes

Figure 5 The Structure and phylogeny of the LPS cluster. A) Schematic comparison of LPS gene clusters described in the present study. Genes conserved in different strains are given identical color. Genes specific to individual strains are given unique color. “Hpo pro” indicates an ORF encoding for a hypothetical protein. The red color-coded genes in \textit{Xp} genes are absent in any of the sequenced xanthomonads. B) Phylogenetic tree based on conserved \textit{metB} and \textit{etfA} genes that flank the variable LPS locus. Strains abbreviations are as in the main text. Arrow indicates the horizontal gene transfer event in the lineage that gave rise to \textit{Xp}.

URL: http://www.biomedcentral.com/1471-2164/12/146
The cellular level of cyclic di-GMP is controlled by a balance between synthesis by GGDEF domain diguanylate cyclases and degradation by HD-GYP or EAL domain phosphodiesterases. GGDEF, EAL and HD-GYP domains are largely found in combination with other signaling domains, suggesting that their activities in cyclic di-GMP turnover can be modulated by environmental cues. A number of proteins involved in cyclic di-GMP signaling have been implicated in virulence of Xcc [86,87]. The genome of Xcv encodes 3 proteins with an HD-GYP domain and 33 proteins with GGDEF and/or EAL domains. As in other Xanthomonas spp., the HD-GYP domain proteins are completely conserved in Xcv, Xv, Xg and Xp. There is also almost complete conservation of GGDEF/EAL domain proteins between Xcv and three draft genomes, although Xv has no ortholog of XCV1982 (Additional file 7: Table S7). In addition, the EAL domain protein (XCVd0150) encoded on a plasmid in Xcv is absent in the other strains.

Copper resistance (cop) genes are present in Xv and copper homeostasis (coh) genes are present in all strains

Among the Xcv, Xv, Xp and Xg strains sequenced, Xv is the only one resistant to copper and the only strain harboring a set of plasmid borne genes, namely copL, copA, copB, copM, copG, copC, copD, and copF that are also present in copper resistant strains of Xac (unpublished data/Behlau, F. personal communication) and S. maltophilia [88]. Genes copA and copB have been previously annotated as copper resistance related genes for many different xanthomonad genomes including Xoo, Xac, Xcv, Xac and Xcc. Homologs of these genes are also present in Xv, Xg and Xp and are located on the chromosome. Additionally, upstream of copA on the chromosome of all strains, there is an ORF that shares homology with plasmid copL. In contrast to what has been published, chromosomal copA and copB are not responsible for copper resistance but likely for copper homeostasis and/or tolerance. While strains harboring the plasmid-borne cop genes, like in Xv, are resistant to copper and can grow on MGY agar (mannitol-glutamate yeast agar) amended with up to 400 mg L\(^{-1}\) of copper sulfate pentahydrate, strains that have only the chromosomal cop genes as for Xcv, Xp and Xg, are sensitive to copper and can only grow on media amended up to 75 mg L\(^{-1}\) of copper. Nucleotide sequence of plasmid cop genes in Xv are 98% similar to the ones found in Xac and Stenotrophomonas, whereas chromosomal copLAB from Xv is 83% identical to homolog ORFs in Xcv, Xg and Xp. When copL, copA and copB genes from Xv located on the plasmid are compared to the homologs on the chromosome of the same strain, the identity of nucleotide sequences is 27, 73, and 65%, respectively. To avoid further confusion or misinterpretation, we suggest that the nomenclature of the chromosomal copL, copA and copB genes in xanthomonads should be changed to cohL, cohA and cohB, respectively, referring to copper homeostasis genes. New nomenclature has been adopted in the annotation of the draft genomes.

**Genes unique to X. perforans as compared to pepper pathogens give clues to its predominance over Xcv in the field and host specificity**

Thirteen gene clusters were found to be specific to the tomato pathogen Xp when compared to the other three strains (Additional file 8: Table S8). A part of the clusters are syntenic to the genomic regions specific to the three pepper pathogens, suggesting the replacement of these genomic regions from pepper pathogens in correspondence to these region in Xp. These replaced regions in Xp might provide potential candidates for host range determinants. Most notable among these regions was the LPS cluster genes (See above). Other such regions include the avirulence genes avrXv3 and avrXv4, a TIR-like domain containing protein, oxidoreductases, and bacteriocin-like proteins that were not found in any other sequenced xanthomonads. Importance of bacteriocin-like genes in Xp has already been studied for its predominance in the field over T1 strains [89,90]. Alien_hunter analysis showed that the bacteriocin BCN-A region belongs to an anomalous region indicating possible horizontal gene transfer of this region (Additional file 5: Table S5).

**Pepper pathogenicity/aggressiveness factors increased in planta growth of Xp**

Comparison of proteomes of Xv, Xg, Xcv against Xp showed 68 genes exclusive to pepper pathogens which might be candidate virulence factors on pepper (Additional file 9: Table S9). These include 16 genes with known function, 35 coding for mobile genetic elements, and 17 genes with unknown function/hypothetical proteins. Out of the 16 genes with known function, copG was confirmed to be a type III effector using the avrBs2 reporter gene assay and 6 genes belong to the LPS biosynthesis gene cluster. These 16 genes were searched against already sequenced genomes of Xac, Xcc and Xoo. The wxcO gene, which codes for O-antigen, has been identified to be a virulence factor in the X. fuscans bean pathosystem by subtractive hybridization [91]. Three genes, XCV1298, XCV1839 and wxcO, were initially selected for the verification of their contribution to virulence in pepper. Individual genes along with their promoter regions were cloned into pLAFR3 and conjugated individually and in combination into X. perforans
ME24 (91-118ΔavrXv3), which no longer elicits an HR in pepper. However, in planta growth of ME24 is more similar to that of an avirulent strain than the virulent pepper strain TED3 race 6. ME24 transconjugants carrying wxcO and XCV1839 in combination showed increased in planta growth and also comparatively increased number of lesions on pepper cv. ECW when compared to ME24 revealing that these two genes play in fact a role in pepper pathogenicity (Figure 6).

**Genes specific to Xg as compared to other tomato/pepper pathogens may explain its aggressive nature on tomato and pepper**

Comparison of genes from Xg against Xcv, Xp and Xv genes showed the presence of 625 genes specific to Xg (Additional file 10: Table S10). These include four type III effectors (avrBs1 member, xopAO, avrHah1, xopAQ), twenty-one genes belonging to the unique type IV secretion system cluster and associated genes. These genes can be speculated to contribute to the aggressive nature of Xg strains on tomato and pepper. Xg also contains a unique beta xylosidase not present in any other xanthomonads. Type II secreted beta xylosidase has been studied for its role in plant cell wall digestion. Moreover, Xg contains XGA_3730 coding for a hemolysin-type calcium-binding repeat containing protein, a homolog of which is found in Xylella strains with 55% sequence identity. In Xylella, this gene is annotated as a member of a family of pore forming toxins/RTX toxins. Its homolog is also found in other plant pathogens (i.e. P. syringae pv. syringae B728a and R. solanacearum GMI1000). This protein has been described as a type I effector in X. fastidiosa strain temecula (PD1506) [92]. RTX toxin family members, especially of the hemolysin type, have been shown to be virulence factors in a variety of cell types in eukaryotes [93,94]. Finally, a gene XGA_0603 coding for lantionine synthetase (lantibiotic biosynthesis) is found among these Xg specific genes, a homolog of which is found in Xvm NCPPB702. LanL enzymes in pathogenic bacteria contribute to virulence by modifying the host signaling pathways, in most cases by inactivating MAPKs [95].

**Genes common to all tomato pathogens but absent from other sequenced xanthomonads**

In order to see what defines the tomato pathogens, we compared the four sequenced genomes (Xv, Xp, Xg and Xcv) to other sequenced xanthomonads. We found seven genes that were conserved in all four tomato pathogens and absent from most of other sequenced xanthomonads with the exception of Xcm, Xvv, Xaub and Xauc, which possess homologs for six out of these seven genes (Table 8). Only the hypothetical protein XCV2641 seems to be specific to the four tomato pathogens. This gene shows only 35% sequence identity to a gene from Xvv and Xcm. A homolog of the hypothetical protein, XCV4416 was found in Xau, but is absent from all other sequenced xanthomonads. Genes homologous in Xcm and Xvv include two transposase genes both belonging to the transposase 17 superfamily (XCV0615, XCV0623), XCV0041 (putative penicillin amidase fragment), XCV0111 (lignostilbene- alpha, beta dioxigenase), XCV0112 (uncharacterized protein conserved in bacteria) (Table 8). Interestingly, XCV0111 encodes a protein known to be involved in phenylpropanoid degradation. Phenylpropanoids are well known plant secondary metabolites induced during defense response upon pathogen attack [96]. It appears that the four tomato pathogens along with Xvv and Xcm have acquired this function to disarm the basal plant defense.

**The evolution of pathogenicity clusters corresponds to the MLST-based phylogeny**

The correlation between tree topology using MLST and phylogeny based on the sequences of pathogenicity clusters and the avrBs2 effector gene, which is found in all xanthomonads, was tested. Based on MLST, Xp and Xcv group together along with Xac while Xg is more closely related to Xcc. Xv forms a different clade and is more closely related to the Xcc group. As can be seen in Figure 7, phylogeny based on MLST is congruent with phylogeny based on the pathogenicity clusters (gun, hrp cluster) and based on the avrBs2 effector, suggesting that overall these clusters were vertically inherited from the most recent common ancestor of these strains.

**Conclusions**

The interaction of Xanthomonas strains with tomato and pepper represents a model system for studying plant-pathogen co-evolution because of the diversity present
among the strains causing bacterial spot. Although the four *Xanthomonas* species infect the same host, tomato, and cause very similar disease, they are genetically diverse pathogens. The comparative genomic analysis has provided insights into the evolution of these strains. Whole genome comparisons revealed that *Xg* and *Xv* are more closely related to *Xcc* than *Xcv* and *Xp*. A few pathogenicity clusters, such as *hrp*, *xcs* and *xps* of *Xg* and *Xv*, were similar in terms of genetic organization and sequence identity to *Xcc* (Figure 8). However, a few pathogenicity clusters of the four strains belonging to four phylogenetic groups showed different evolutionary origins. While the

### Table 8 Genes present in all four tomato and pepper pathogens but absent from all other sequenced xanthomonads

<table>
<thead>
<tr>
<th>Locus tag for Xcv85-10</th>
<th>Possible function</th>
<th>Homolog present in any other genera</th>
<th>GC content</th>
</tr>
</thead>
<tbody>
<tr>
<td>XCV0623</td>
<td>Transposase 17 superfamily Hypo protein -COG belonging to transposase, inactive derivatives</td>
<td>In <em>Stenotrophomonas</em>, <em>Acidovorax</em> Xanthomonas campestris pv. musacearum NCPPB4381</td>
<td>0.59</td>
</tr>
<tr>
<td>XCV2641</td>
<td>Hypothetical protein</td>
<td><em>X. c. musacearum</em> and <em>X. c. vasculorum</em> (identity 37, 31% respectively)</td>
<td>0.65</td>
</tr>
<tr>
<td>XCV4416</td>
<td>Hypothetical protein</td>
<td><em>Pectobacterium carotovorum</em> <em>X. fuscans</em> pv. <em>aurantifoli</em></td>
<td>0.57</td>
</tr>
<tr>
<td>XCV0615</td>
<td>Transposase 17 superfamily Hypothetical protein COG1943 (transposase, inactivated derivatives)</td>
<td><em>Acidovorax</em>, <em>X. c. musacearum</em> and <em>X. c. vasculorum</em></td>
<td>0.62</td>
</tr>
<tr>
<td>XCV0112</td>
<td>COG4704 uncharacterized protein conserved in bacteria</td>
<td><em>Stenotrophomonas</em>, <em>X. c. musacearum</em> and <em>X. c. vasculorum</em></td>
<td>0.65</td>
</tr>
<tr>
<td>XCV0111</td>
<td>putative lignostilbene-alpha,beta-dioxygenase-phenylpropanoid compound degradation</td>
<td><em>Stenotrophomonas</em>, <em>Ralstonia</em>, <em>X. c. musacearum</em> and <em>X. c. vasculorum</em></td>
<td>0.66</td>
</tr>
<tr>
<td>XCV0041</td>
<td>putative penicillin amidase (fragment)</td>
<td><em>Ralstonia</em>, <em>X. c. musacearum</em> and <em>X. c. vasculorum</em></td>
<td>0.64</td>
</tr>
</tbody>
</table>

---

**Figure 7** Correlation between phylogenies based on Multi-Locus Sequence Typing (MLST) core genome and pathogenicity clusters. Concatenated amino acid sequences of the six genes *fusA*, *gapA*, *gltA*, *gyrB*, *lacF*, *lepA* from four bacterial spot pathogen strains along with other sequenced xanthomonads are considered in the analysis. The evolutionary history was inferred using the Neighbor-Joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. Phylogenetic analyses were conducted in MEGA4.
pepper pathogens Xcv, Xv and Xg possess similar LPS biosynthesis cluster, part of the LPS cluster from Xp is similar to the one from Xac (Figure 8). Xv contains few effectors, including xopAG (avrGf1) and xopAL the latter of which was previously found to be unique to citrus pathogens Xac, Xaub and Xauc [25]. Xg has a number of effectors homologous to P. syringae type III effectors suggesting probable horizontal transfer of these effectors. Xg contains a unique T4SS along with the one that is exclusive to Xp, Xv and Xg. Xp has two T6SSs, as found in Xcv. Xv has only one T6SS which is similar to that of Xac. Xg has no T6SS as seen for Xcc (Figure 8). While Xg and Xv show close relationship to Xcc based on whole genome comparisons, few pathogenicity clusters mentioned above seem to be conserved among tomato/pepper xanthomonads.

Type III effectors have been investigated for their contribution to pathogenicity and host-range specificity. In addition to homologs of the known effectors, we identified novel effectors in the draft genomes. By comparing effector repertoires of tomato pathogens, two possible candidate pathogenicity determinants, xopF1 and xopD, were identified, of which xopD is responsible for delaying symptom development, and in turn, is important for pathogen survival. Unique genes present in Xg include the novel effectors xopAO, xopAQ, xopAS and an avrBs1 member as well as a few other virulence factors, which have been characterized in other plant pathogens and which could explain the aggressive nature of Xg on pepper. Each species contains at least three unique type III effectors, which could explain host preferences among the strains and their aggressiveness on tomato/pepper. Comparison of the LPS clusters between the four species revealed significant variation. Xp has acquired a novel LPS cluster during evolution, which might be responsible for its predominance and its limited host range. As seen from the in planta growth assay of Xp ΔavrXv3 mutant carrying the LPS O-antigen from Xcv, the LPS cluster from pepper pathogens can be a contributor to the increased in planta growth of Xp ΔavrXv3 mutant on pepper, but is not the absolute virulence determinant. Use of the XA21 receptor similar to the

---

**Figure 8** A diagrammatic representation of relationship among bacterial spot xanthomonads, Xac and Xcc with respect to presence or absence of pathogenicity clusters. Similar color shade indicates high identity and similar cluster organization. Lower sequence identities compared to the reference are indicated by faded gray shades. Reference strain is indicated by asterisk next to the symbol. The absence of certain part of cluster is indicated by white. In the case of LPS cluster, Xv and Xcc contain novel cluster regions in the C terminal region which is indicated by a different color. Xac and Xcv contain a plasmid borne type IV cluster. Although it differs from type IVA present in other bacterial spot xanthomonads, Xac and Xcv cluster is mentioned here under type IVA with different colors. A blank space indicates complete absence of gene cluster in that particular species. A more detailed representation of individual clusters can be found in figures 2 through 5.
Xoo-rice system in Xcv - tomato/pepper could be one of the ways to confer resistance to xanthomonads due to presence of a similar AX21 peptide and a functional rax system in Xcv. Common and unique genes encoding enzymes involved in cell wall deconstruction are candidates for further study to define host preference and virulence.

In conclusion, comparison of draft genomes obtained by next generation sequencing has allowed an in-depth study of diverse groups of bacterial spot pathogens at the genomic level. This analysis will serve as a basis to infer evolution of new virulent strains and overcoming existing host resistance. The knowledge of potential virulence or pathogenicity factors is expected to aid in devising effective control strategies and breeding for durable resistance in tomato and pepper cultivars.

Methods
Genome sequencing
Xv, Xp and Xg were sequenced by 454-pyrosequencing [27] at core DNA sequencing facility, ICBR, University of Florida. Xanthomonas isolates were grown overnight in nutrient broth. Genomic DNA was isolated using CTAB-NaCl extraction method [97] and resuspended in TE buffer (10 mM Tris pH 8, 1 mM EDTA pH 8). Libraries of fragmented genomics DNA were sequenced on 454-Genome Sequencer, FLX instrument at Interdisciplinary Center for Biotechnology Research (ICBR) at UF. De novo assemblies were constructed using 454 Newbler Assembler [27]. The three draft genomes were obtained with around 10× coverage. For Illumina sequencing, the Xanthomonas strains were purified from single-colony and grown overnight in liquid cultures. Genomic DNA was isolated by phenol extraction and precipitated twice with isopropanol, and finally dissolved in TE buffer. DNA was then purified by cesium chloride density gradient centrifugation and precipitated with 95% ethanol, then dissolved in TE buffer. DNA was then purified by illumina reads, wherein, 454 based contigs were used as long reads to fill in gaps generated during combined de novo assembly. These combined assemblies of each genome were uploaded on IMG-JGI (Joint Genome Institute, Walnut Creek, California) server for gene calling. The gene prediction was carried out using GeneMark. Pfam, InterPro, COGs assignments were carried out for identified genes. Pathogenicity clusters described in the paper were manually annotated.

Whole genome comparisons
We aligned draft genomes against reference Xanthomonas genomes using nucmer [31] of MUMmer program (version 3.20) and dnadiff was used to calculate percentage of aligned sequences. We have also compared genomes using the MUM index [30] to measure distances between two genomes. The maximal unique exact matches index (MUMi) distance calculation was performed using the Mummer program (version 3.20). Mummer was run on concatenated contigs or replications (achieved by inserting a string of 20 symbols ‘N’ between contig or replicon sequences) of each genome. The distance calculations performed using the MUMi script are based on the number of maximal unique matches of a given minimal length shared by two genomes being compared. MUMi values vary from 0 for identical genomes to 1 for very distant genomes [30].

Phylogenetic analysis
MLST sequences (fusA, gapA, gltA, gyrB, lacF, lepA) for all the genomes were obtained in concatenated form from PAMDB website http://pamdb.org. Genes and their corresponding amino acid sequences spanning gum, hrp cluster were downloaded from NCBI genbank sequences of sequenced genomes. Amino acid sequences of proteins of these clusters for Xcv and Xcc were used as query to search for homology against draft genomes of Xp, Xv and Xg. The amino acid sequences were then concatenated for each pathogenicity cluster and then aligned using CLUSTALW ignoring gaps. Neighbour-joining trees were constructed with bootstrap value for 1000 replicates using MEGA4 [98]. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). There were a total of 2723 positions in the final dataset.

Phylogeny reconstruction
Species tree. We used a supermatrix approach as in previous work [25]. Protein sequences of six Xanthomonas genomes (ingroups) and the S. maltophilia R551-3 genome (outgroup) were clustered in 5,096 families using OrthoMCL [99]. We then selected families with one and only one representative from each of the ingroup genomes and at most one outgroup protein, resulting in
2,282 families. Their sequences were aligned using MUSCLE [100] and the resulting alignments were concatenated. Non-informative columns were removed using Gblocks [101], resulting in 792,079 positions. RAxML [102] with the PROTGAMMAWAGF model was used to build the final tree.

Prediction of effector repertoires, cloning of candidate effectors and confirmation using avrBs2 reporter gene assay
A database was created collecting all the known plant and animal pathogen effectors. Using all these known effectors as query, tblastn analysis was performed against all contigs of the draft genomes of Xv, Xg and Xp with e-value of 10^{-5} [103]. Pfam domains were searched for possible domains found in known effectors in predicted set of ORFs of draft genome sequences. Candidate effectors were classified according to the nomenclature and classification scheme for effectors in xanthomonads recently [24]. Candidate effectors showing < 45% identity at amino acid level to the known effectors were confirmed for their translocation using avrBs2 reporter gene assay.

N-terminal 100 amino acid region along with upstream 500 bps sequence of candidate genes were PCR amplified using primers with BglII restriction sites at the 5’ ends. Following digestion with BglII, PCR amplicons were ligated with BglII-digested pBS (BglII::avrBs2_{62-574::HA}) (courtesy of Dr. Mary Beth Mudgett, Stanford university), and later transformed into E. coli DH5a. In-frame fusions were confirmed by DNA sequencing using F20 and R24 primers. BamHI–KpnI fragments containing the candidate gene fused to avrBs2 was then cloned into pUFR034. Resulting plasmids were then introduced into pepper race 6 (cv. TED3) by tri-parental mating. The transconjugants with the three individual genes and combined ones along with virulent pepper race 6 strain were infiltrated in pepper cv. ECW and leaves were sampled at every 48 hours after inoculation. The samples were plated on nutrient agar, incubated at 27°C and CFU/ml counts were enumerated. Experiment was carried out in triplicate and repeated three times.

Cloning of pepper specificity genes in Xp
The three genes mentioned above were cloned individually and in combination in pLAFR3 vector and conjugated in Xp 91-118 ΔavrXv63 mutant PM1. The PM1 transconjugants with the three individual genes and combined ones along with virulent pepper race 6 strain were infiltrated at 10^5 CFU/ml concentration in pepper cv. ECW and leaves were sampled at every 48 hours after inoculation. The samples were plated on nutrient agar, incubated at 27°C and CFU/ml counts were enumerated. Experiment was carried out in triplicate and repeated three times.

Database submission
The draft genome sequences of Xanthomonas vesicatoria ATCC 35937 (Xv) have been deposited at DDBJ/EMBL/GenBank under accession number AEQV00000000. The draft genome sequences of Xanthomonas perforans 91-118 (Xp) have been deposited at DDBJ/EMBL/GenBank under accession number AEQW00000000. The draft genome sequences of Xanthomonas gardneri ATCC 19865 (Xg) have been deposited at DDBJ/EMBL/GenBank under accession number AEQX00000000. The version described in this paper is the first version, AEQV01000000, AEQW01000000, AEQX01000000. All three draft genomes will be released upon manuscript acceptance.

Additional material

Additional file 1: Table S1: General features of the sequencing data
Additional file 2: Figure S2: 1a Phylogenetic tree based on MUMi indices; 1b) Distance matrix based on MUMi indices. MUMi program was used to calculate pairwise distances between draft genomes and reference Xanthomonas genomes.
Additional file 3: Table S3: Whole genome comparisons using MUMmer mndiff program. % coverage of the aligned contigs and % identities of the respective contigs against reference genomes has been shown for each draft genome.
Additional file 4: Figure S4: AvrBs2-based HR assay confirms translocation of novel effectors. Hypersensitive response reaction indicating presence of translocation signal was recorded 24 hrs after inoculation on pepper cv. ECW20R with candidate effectors xopZ2 (a), avrB17 (b), xopG (d), xopAM (e), xopA0 (f) conjugated in race 6 strain along with control race 6 strain (c). All the strains showed water-soaking on pepper cv. ECW after 48 hrs after inoculation.
Additional file 5: Table S5: Evidence of the horizontal gene transfer using Alien Hunter analysis.
Additional file 6: Table S6: Genes/contigs representing T6SS in draft genomes as compared to Xcv
Additional file 7: Table S7: Domain architecture and distribution of proteins with HD-GYP, GGDEF and/or EAL domains encoded by genomes of different Xanthomonas strains.
Additional file 8: Table S8: Genes unique to pepper pathogens but absent from Xp
Additional file 9: Table S9: Genes common to all pepper pathogens
Additional file 10: Table S10: Genes unique to Xg. Genes of special interest are highlighted in red and yellow.

List of abbreviations
Xv: Xanthomonas vesicatoria strain 85-10; formerly, Xanthomonas campestris pv. vesicatoria strain; Xg: Xanthomonas vesicatoria strain 1111 (ATCC 35937); Xp: Xanthomonas perforans strain 91-118, Xg: Xanthomonas gardneri strain 101 (ATCC 19865); Xoo: Xanthomonas oryzae pv. Oryzae; Xcc: Xanthomonas campestris pv. Campestris; Xcv: Xanthomonas campestris pv. Musacearum NCPPB4381; Xvv: Xanthomonas vesicatoria pv. vesicatoria NCPPB702; Xvu: Xanthomonas citrulli subsp. citri strain 306, formerly, Xanthomonas avanopodis pv. citrulli strain 306; Xab: Xanthomonas fuscans subsp. auranofiloi B strain; Xauc: Xanthomonas fuscans subsp. auranvioloi C strain; Xac: Xanthomonas oryzae pv. Oryzaicola; Xxc: Xanthomonas campestris pv. musacearum NCPPB4381; Xv: Xanthomonas vesicatoria subsp. citri.
Acknowledgements

This research was supported by USDA-NIFA special grant T-STAR (J. B. Jones, J. F. Preston and M. T. Monro, USDA 2006-34135-17969). We thank ILG-JGI for providing genome annotation pipeline and interface for the comparative genomic analyses. Special thanks to Kostantinos Movromatis in helping preparation of Genbank files.

Authors’ contributions

JBI conceived the project. JBI and BJS oversaw genomic sequencing. JBI provided the strains. KK did genome assembly. NP, RPR, RK, VC, JFP, PBP, JMD, MS, TM did analyses of different genomic validations of the pathogenicity clusters. NFA created ortholog families, and JBI and BJS helped create an effector database. NP carried out effector analysis and confirmed annotated pathogenicity clusters in the genomes. DJN and BJS helped with preparation of Genbank files.

Received: 1 December 2010 Accepted: 11 March 2011

Published: 11 March 2011

References


BMC Genomics 2011, 12:146

http://www.biomedcentral.com/1471-2164/12/146

Page 22 of 23


