# Complete genome sequence reveals evolutionary dynamics of an emerging and variant pathovar of *Xanthomonas euvesicatoria*

**Authors:** Kanika Bansal[1], Sanjeet Kumar[1,†] and Prabhu B. Patil[*]

[1]Bacterial Genomics and Evolution Laboratory, CSIR-Institute of Microbial Technology, Chandigarh, India.

[†]Present address: Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena, Germany.

***Author for Correspondence:** Dr. Prabhu B. Patil, Bacterial Genomics and Evolution Laboratory, CSIR-Institute of Microbial Technology, Chandigarh, India. pbpatil@imtech.res.in

Downloaded from https://academic.oup.com/gbe/advance-article-abstract/doi/10.1093/gbe/evy238/5140196 by MPI Science of Human History user on 06 November 2018

© The Author(s) 2018. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

## Abstract

*Xanthomonas,* a complex group of pathogens infecting more than 400 plants which is expanding to new hosts causing serious diseases. Genome based studies are transforming our understanding on diversity and relationship of host specific members, known as pathovars. In present study, we report complete genome sequence of a novel pathovar *Xanthomonas axonopodis* pv. commiphorae (Xcom) from India. It causes gumming disease of *Commiphora wightii*, a medicinally important plant. Genome-based phylogenetic and taxonomic investigations revealed that the pathovar belongs to *X. euvesicatoria* and not *X. axonopodis* as reported earlier. Interestingly, it is a novel host and novel geographic origin for a *X. euvesicatoria* pathovar. A core-genome based phylogenetic analysis resolved the pathovar complex of this species on the basis of their hosts. Interestingly, this pathovar harbours a unique 35 kb plasmid encoding type III effectors and toxin-antitoxin gene that is absent in other *X. euvesicatoria* pathovars and infects tomato, pepper, rose, onion, philodendron, alfalfa and citrus plants. The pathovar contains two TAL genes, one on plasmid and another on genomic region with an additional pseudo TAL gene flanked by IS elements in the plasmid. Further, Xcom has acquired a novel set of lipopolysaccharide biosynthesis genes after its divergence from the closely related pathovar that infects rose and supports the role of horizontal gene transfer in hypervariation at this locus in the species. Complete genome sequence of this variant pathovar has provided novel insights into evolution of an emerging pathovar in *Xanthomonas* and will be valuable resource in pathogenomics of *X. euvesicatoria*.

**Keywords:** *Xanthomonas euvesicatoria,* genomics, evolution, pathovar, TAL, complete genome.

Downloaded from https://academic.oup.com/gbe/advance-article-abstract/doi/10.1093/gbe/evy238/5140196 by MPI Science of Human History user on 06 November 2018

## Introduction

*Xanthomonas* is a highly successful pathogen known to infect more than 400 plants in a host specific manner (Leyns, et al. 1984). *Xanthomonas* species and the constituent pathovars have undergone numerous re-classifications because of challenges associated with the conventional methods of bacterial classification (Dye, et al. 1980; Hauben, et al. 1997; Rademaker, et al. 2005; Van den Mooter and Swings 1990; Vauterin, et al. 1995; Young, et al. 1978). Advent of DNA based typing and grouping methods provides accurate understanding on species diversity of *Xanthomonas* (Constantin, et al. 2015; Rodriguez, et al. 2012). In the present scenario, around 150 host specific members, known as pathovars, are grouped into 33 valid species (List of Prokaryotic names with Standing in Nomenclature (LPSN) http://www.bacterio.net/xanthomonas.html) (Parte 2018), which indicates intra-species diversification as a major phenomenon. Some species like *Xanthomonas citri, X. axonopodis* and *X. euvesicatoria* constitute numerous pathovars (Bansal, et al. 2017; Barak, et al. 2016; Constantin, et al. 2015; Midha and Patil 2014). Interestingly, majority of the pathovars are first reported from India which is one of the major centre of diversity and major cultivator of host plants. Surprisingly, nineteen out of twenty-two pathovars were found to belong to *X. citri* and not several species as reported earlier (Bansal, et al. 2017; Parkinson, et al. 2009). These studies indicate both ecological and evolutionary investigations are important in systematic understanding *Xanthomonas* group of pathogens.

With the advent of high-throughput genomics, there is an opportunity to look into diversity and evolution of pathovars and individual isolate at unprecedented details. Now, it is possible to accurately establish species status of new pathovar or an isolate using modern genome based taxonomic criteria (Lee, et al. 2016; Meier-Kolthoff, et al. 2013). At the same time, availability of large number of phylogenomic markers and core gene content allows to resolve the relationship of a pathovar that forms complex of a particular species. Apart from providing

Downloaded from https://academic.oup.com/gbe/advance-article-abstract/doi/10.1093/gbe/evy238/5140196 by MPI Science of Human History user on 06 November 2018

robust phylogeny, comparative studies are leading to the identification of genomic determinants of host specificities and role of horizontal gene transfer in variable gene-content. With the emergence of long-read technology, it is now possible to study the highly repetitive transcription activator like (Barbolla, et al.) proteins that are secreted by type III secretion system to modulate host genes for pathogenicity and role of plasmids in diversification of bacterial isolates (Boch and Bonas 2010). This aspect is particularly relevant in *Xanthomonas* species, which displays high infra-species diversity (Bansal, et al. 2017; Constantin, et al. 2015; Parkinson, et al. 2009).

Recently, a new pathovar of *Xanthomonas* was reported from India and found to cause a serious gummosis in *Commiphora wightii* (Samanta, et al. 2013). This plant is cultivated for medicinally important oleo-gum-resin and is native to South Asia. Biochemical and initial phylogenetic analysis has classified it as *X. axonopodis* pv. commiphorae (Xcom). In the present study, we report complete genome sequence of this novel pathovar and surprisingly, its genomotaxonomic revealed that Xcom belongs to *X. euvesicatoria* and not to *X. axonopodis*. *X. euvesicatoria* is a major pathogen of tomato and pepper, and all other pathovars of this species were reported outside India, mainly in the United States (Albuquerque, et al. 2012; Barak, et al. 2016; Potnis, et al. 2011). In this context, Xcom represents expanded host range and geographic origin of a *X. euvesicatoria* pathovar and has provided us an opportunity to look for the genome dynamics in the emergence of a novel pathovar.

## Results

### Complete genome sequence of Xcom

Complete genome sequencing of Xcom was performed using Oxford Nanopore MinION and Illumina Miseq platform (methods). Xcom has a circular chromosomal DNA of 4.8 Mb and one plasmid of 35 kb with NCBI accession numbers: CP031059 and CP031060 respectively. The genome coverage obtained is 124x with 4,203 numbers of CDS and 53 tRNAs. GC content

Downloaded from https://academic.oup.com/gbe/advance-article-abstract/doi/10.1093/gbe/evy238/5140196 by MPI Science of Human History user on 06 November 2018

for the genomic and plasmid regions are 65.2% and 62.5% respectively. Further, its completeness and contamination for the genome was found to be 99.64% and 0% respectively.

Since we have a complete genome sequence, highly repetitive *tal* genes could be easily fetched. TAL effectors are trans kingdom remote controls of gene expression, which once inside the plant cell nucleus, induces the transcription of eukaryotic genes in a sequence specific manner (Boch and Bonas 2010). Interestingly, with complete genome, *tal* gene profile of the Xcom was found to be unique in having two *tal* genes, one on the genomic (Xcom_11830) and one on the plasmid (Xcom_21330) region, in addition to a pseudo *tal* gene (Xcom_21360) in the plasmid region. Both the copies of *tal* genes in plasmid are surrounded by transposable elements. Interestingly, the chromosomal *tal* gene is also flanked by a transposable element. Apart from *tal* genes, the plasmid also encodes for non-TAL effectors and type II toxin-antitoxin genes. The unique plasmid also harbors a cluster of nine hypothetical genes and an acetyltransferase gene.

**Taxonogenomic and genealogical investigation**

Taxonogenomic analysis (using OrthoANI and dDDH) including type strains of phylogenetic groups of *X. axonopodis* depicted that Xcom belongs to *X. euvesicatoria* (orthoANI value of 98.72% and dDDH value of 88.9%) much above the cut-off for species delineation of 96% and 70% respectively and not to *X. axonopodis* (orthoANI value of 93.14% and dDDH value of 49.6%) (Figure 1). Hence, now onwards, we will refer the pathovar as *X. euvesicatoria* pv. commiphorae LMG26789 (Xcom).

*X. euvesicatoria* is a complex species consisting of pathovars infecting asterid, rosid and monocot hosts (supplementary table 1). Hence, we carried out genealogical investigation to understand the evolution of diverse range of pathovars (figure 2). Interestingly, strains from asterids hosts, which are native to America, formed a distinct clonal group (CG-I); and strains

Downloaded from https://academic.oup.com/gbe/advance-article-abstract/doi/10.1093/gbe/evy238/5140196 by MPI Science of Human History user on 06 November 2018

infecting rosids, including Xcom and monocots hosts formed CG-II, while *X. alfalfae* were not included in CGs. Interestingly, host of CG-II strains are mainly native to Asia (except *Philodendron*, known to be originated in South America).

**Variations in lipopolysaccharide (LPS) biosynthetic gene clusters**

LPS locus in genus *Xanthomonas* is hypervariable at species and pathovar level (Patil, et al. 2007). *X. euvesicatoria* pathovar display hypervariability in LPS biosynthetic gene cluster (figure 3). *X. perforans* and *X. euvesicatoria*, both infecting asterids from CG I are having diverse LPS gene clusters. In contrast, *X. perforans* and *X. axonopodis* pv. dieffenbachiae that infect asterids and rosids respectively, have same LPS cassette. At the same time, *X. euvesicatoria* and *X. alfalfa* strain GEV-Rose-07 from CG II are having identical LPS. Moreover, Xcom shares half of its cassette with these pathovars, while the other half of Xcom LPS cassette is showing homology with *X. albelineans.*

## Discussion

Earlier, genome based investigation has revealed that five known pathovars (*X. axonopodis* pv. dieffenbachiae strain LMG12749, *X. axonopodis* pv. allii CFBP6369, *X. alfalfae* strain GEV-Rose-07, *X. axonopodis* pv. citrumelo F1 and *X. perforans*) of *Xanthomonas* belonged to species *X. euvesicatoria* (Barak, et al. 2016). These were misclassified into three different species i.e. *X. axonopodis*, *X. alfalfa*, *X. perforans,* and some were reported as *Xanthomonas* sp. Originally, Xcom was also classified as *X. axonopodis* pathovar based on traditional approaches and limited sequence information (Samanta, et al. 2013). However, similar to the case of other pathovars, genome based taxonomy and phylogenetic analysis, clearly established Xcom as a new pathovar of *X. euvesicatoria*. Hence, genome based studies are critical in understanding the complex group of pathogens and particularly pathovar rich species like *X. euvesicatoria, X. axonopodis* and *X. citri*.

Downloaded from https://academic.oup.com/gbe/advance-article-abstract/doi/10.1093/gbe/evy238/5140196 by MPI Science of Human History user on 06 November 2018

In an earlier genome-based study, we have reported that nineteen pathovars, previously classified into three different species, belonged to one species i.e., *X. citri* (Bansal, et al. 2017). All these pathovars were first reported from India, in the last century and believed to have spread to other parts of the world (Parkinson, et al. 2009). Many *X. citri* pathovars still do not have reports from India. In contrast, all the other pathovars of *X. euvesicatoria* were first reported outside India and Xcom will be the first pathovar from *X. euvesicatoria* known from India. *X. euvesicatoria* and *X. citri* are evolutionary diverse species in the genus. Hence, it is surprising that a new pathovar that infects a host that is native to India belongs to *X. euvesicatoria*. This provided us an opportunity to carry out a complete genome based investigation on diversity and evolution of *X. euvesicatoria* pathovars by including an ecologically variant pathovar.

Initial core-genome phylogeny revealed that Xcom is closely related to pathovar that infects rose. Interestingly, the rose pathovar was first reported in 2004 from Texas and Florida regions of North America (Huang, et al. 2013). Present analysis has also suggested that the two recently reported rose and Xcom pathovars are from a rapidly diversifying lineage and might have originated from a common ancestor that has evolved into pathovars infecting other rosids like *Commiphora*, rose, alfalfa, citrus, etc. Apart from phylogeny, it is also important to look into the dynamic variations that plays role in the evolution of Xcom. We were successful in using short-reads from Illumina platform and long-reads from nanopore sequencer in obtaining complete genome of Xcom pathovar. Complete genome sequence is particularly relevant for new and emerging pathogens like Xcom, as we can also study plasmids, repetitive transcriptional activator like (Barbolla, et al.) proteins and hyper-variable/dynamic genomic regions known in *Xanthomonas*.

Complete genome sequence allowed us to identify a unique plasmid in Xcom that harbors one complete and one pseudo TAL genes surrounded by IS elements. TAL proteins are secreted by

Downloaded from https://academic.oup.com/gbe/advance-article-abstract/doi/10.1093/gbe/evy238/5140196 by MPI Science of Human History user on 06 November 2018

*Xanthomonas* pathovar using type III secretion system to regulate host genes. However, due to modular structure of *tal* genes, they are nearly impossible to get assembled using short sequencing reads. Among the available complete genomes of *X. euvesicatoria*, only *X. perforans* LH3 (Richard, et al. 2017) was found to be have a TAL gene on the plasmid. Interestingly, none of the other complete genomes such as *X. euvesicatoria* 85-10, *X. euvesicatoria* LMG930 and *X. perforans* 91-118 were found to have TAL genes. To the best of our knowledge, this is the first report of genome sequence of an *X. euvesicatoria* pathovar encoding a *tal* gene on the chromosome. At the same time, presence of a complete and a pseudogene of TAL protein surrounded by IS elements suggests on-going dynamics mediated through a unique plasmid.

In *X. euvesicatoria* pathovars, LPS gene clusters are hyper-variable (Potnis, et al. 2011). LPS is both a major virulence factor and a pathogen associated molecular pattern (PAMP) (Patil, et al. 2007). Acquisition of a novel 17.3 kb LPS gene cluster is associated with a host-specific variation in *X. perforans* that infects only tomato compared to *X. euvesicatoria* that infects both tomato and pepper. LPS variation also seems to play a major role in the evolution of Xcom, as half of the LPS gene cluster is novel while other half is similar to that of closely related pathovar that infects rose. This suggests that the acquisition of a new set of LPS genes has happened after divergence of rose and Xcom pathovars from the common ancestor.

Overall, our genome based investigation allowed us to clarify the taxonomic status of Xcom, and its relationship with the other pathovars of *X. euvesicatoria*. The complete genome sequence and findings will be valuable in developing unique markers for identification and surveillance of this novel pathogen. At the same time, identification of a unique plasmid with type III effectors and, both TAL and non-TAL along with toxin-antitoxin genes suggests its evolution as a major pathogen through dynamic variation.

Downloaded from https://academic.oup.com/gbe/advance-article-abstract/doi/10.1093/gbe/evy238/5140196 by MPI Science of Human History user on 06 November 2018

## Methods

### Procurement of Bacterial Culture and revival

The bacterial culture of then defined *X. euvesicatoria (axonopodis)* pv. commiphorae LMG26789 (Xcom) was procured from the Belgium Coordinated Collections of Microorganisms/ LMG (BCCM/LMG) and was revived according to the prescribed instructions.

### Genomic DNA isolation and Illumina Sequencing

ZR Fungal/Bacterial DNA Mini-Prep Kit (Zymo Research, Irvine, CA, USA) was used to obtain high quality bacterial DNA. DNA quantification was performed using Nanodrop 1000 (Thermo Fisher Scientific) and Qubit 2.0 Fluorometric (Invitrogen; Thermo Fisher Scientific). Paired end Illumina sequencing libraries were prepared with Illumina Nextera XT sample preparation kit (Illumina, Inc., San Diego, CA, USA) in accordance to the manufacturer's recommendations and was sequenced on in house Illumina Miseq (Illumina) platform with 2x250 bp paired end sequencing kit.

### Bacterial genomic DNA isolation and ONT MinION Sequencing

Complete bacterial DNA was obtained using the DNeasy Blood & Tissue Kit (Qiagen; 69504) (Qiagen DNeasy, Valencia, CA, USA) and quantified using Nanodrop 1000 and Qubit 2.0. Initial DNA concentration of 4 μg was taken for the shearing of genomic using Covaris g-TUBE (Covaris, Brighton, UK). Purification of the genomic DNA was again performed with AMPure beads to ensure the concentration remains > 1μg (in 45μl). DNA end repair was performed using NEBNext Ultra II End repair/dA-Tailing Module (NEB, Ipswich, USA) and again Ampure XP beads (Beckman Coulter) were used to performed the cleanup. The ONT library was prepared using ONT 1D ligation sequencing KIT (SQK-LSK108) with native barcoding kit (EXP-NBD103). The barcode ligation was performed with New England

Downloaded from https://academic.oup.com/gbe/advance-article-abstract/doi/10.1093/gbe/evy238/5140196 by MPI Science of Human History user on 06 November 2018

Biolabs/ligase master mix module and purified using 0.45X AMPure beads. 50 µl of the pooled DNA sample was used for adapter ligation using T4 DNA ligase.

Prior to MinION Sequencing on Oxford Nanopore MinION Mk1B using MinKNOW (http://community.nanoporetech.com) (Oxford Nanopore Technologies) software (v.1.13.1), quality checks of the flow cell (FLO-MIN106, version R9.4) (Oxford Nanopore Technologies) was performed. The library was combined with the reagents supplied with the Oxford Nanopore and loaded onto the flow cell in accordance to the manufacturer's recommendations with 48 hours sequencing procedure and the data generated was live base called. The raw fastq reads were filtered based on the minimum length of 2 kb using Filtlong v0.2.0 (https://github.com/rrwick/Filtlong).

**Genome assembly & annotation using ONT & Illumina reads**

Unicycler v0.4.4 (Wick, et al. 2017) bold mode was used for genome assembly using ONT long-reads resulting in complete circular chromosomal and plasmid. The assembled genomes were then error corrected with multiple rounds of pilon v1.22 (Walker, et al. 2014) with Illumina short reads. The assembled genome was then checked for the completeness and contamination using CheckM v1.0.11 (Parks, et al. 2015) and submitted to NCBI WGS portal and annotated using NCBI Prokaryotic Genome Annotation Pipeline (https://www.ncbi.nlm.nih.gov/genome/annotation_prok/).

**Genealogical analysis**

ClonalFrameML (Didelot and Wilson 2015) analysis was used to generate phylogenetic tree using genome alignment obtained using MAUVE v2.3.1 (Darling, et al. 2004) and maximum likelihood tree was obtained using PhyML v3.1 (Guindon, et al. 2010). CFML analysis was carried out using MAUVE alignment and PhyML tree, performing 100 simulations.

**LPS cassette and *tal* genes analysis**

Downloaded from https://academic.oup.com/gbe/advance-article-abstract/doi/10.1093/gbe/evy238/5140196 by MPI Science of Human History user on 06 November 2018

Full length LPS cassettes were retrieved from genomes using *etfA* and *metB* conserved genes and annotated using PROKKA 1.11 (Seemann 2014). Easyfig 2.2.2 was used to generate figure (Sullivan, et al. 2011). Further, TAL genes analysis on complete genome was carried out using AnnoTALE (Grau, et al. 2016) software.

## Author Contributions

KB and SK have performed Whole Genome Sequencing and submission to NCBI and analysis. KB have drafted manuscript with inputs from PBP and SK. PBP conceived and participated in its design with inputs from KB. All the authors have read the manuscripts and approved the manuscript.

## Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Acknowledgements

KB is supported by fellowship of University Grant Commission (UGC). SK is supported by the fellowship from CSIR OLP0133. This work is supported by a project entitled “High throughput and integrative genomic approaches to understand adaptation of probiotic and pathogenic bacterium” OLP148 of Council of Scientific and Industrial Research (CSIR) to PBP.

Downloaded from https://academic.oup.com/gbe/advance-article-abstract/doi/10.1093/gbe/evy238/5140196 by MPI Science of Human History user on 06 November 2018

## References

Albuquerque P, et al. 2012. Evolutionary and experimental assessment of novel markers for detection of Xanthomonas euvesicatoria in plant samples. PLoS One 7: e37836. doi: 10.1371/journal.pone.0037836

PONE-D-11-21030 [pii]

Bansal K, Midha S, Kumar S, Patil PB 2017. Ecological and evolutionary insights into Xanthomonas citri pathovar diversity. Applied and environmental microbiology 83: e02993-02916.

Barak JD, et al. 2016. Whole-genome sequences of Xanthomonas euvesicatoria strains clarify taxonomy and reveal a stepwise erosion of type 3 effectors. Frontiers in plant science 7: 1805.

Barbolla R, et al. 2004. Class 1 integrons increase trimethoprim-sulfamethoxazole MICs against epidemiologically unrelated Stenotrophomonas maltophilia isolates. Antimicrobial agents and chemotherapy 48: 666-669.

Boch J, Bonas U 2010. Xanthomonas AvrBs3 family-type III effectors: discovery and function. Annual review of phytopathology 48: 419-436.

Constantin E, et al. 2015. Genetic characterization of strains named as Xanthomonas axonopodis pv. dieffenbachiae leads to a taxonomic revision of the X. axonopodis species complex. Plant Pathology.

Darling AC, Mau B, Blattner FR, Perna NT 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14: 1394-1403.

Didelot X, Wilson DJ 2015. ClonalFrameML: Efficient inference of recombination in whole bacterial genomes. PLoS Comput Biol 11: e1004041-e1004041.

Dye D, et al. 1980. International standards for naming pathovars of phytopathogenic bacteria and a list of pathovar names and pathotype strains. Review of plant pathology 59: 153-168.

Grau J, et al. 2016. AnnoTALE: bioinformatics tools for identification, annotation, and nomenclature of TALEs from *Xanthomonas* genomic sequences. Sci Rep 6.

Guindon S, et al. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59: 307-321.

Hauben L, Vauterin L, Swings J, Moore E 1997. Comparison of 16S ribosomal DNA sequences of all Xanthomonas species. International Journal of Systematic and Evolutionary Microbiology 47: 328-335.

Downloaded from https://academic.oup.com/gbe/advance-article-abstract/doi/10.1093/gbe/evy238/5140196 by MPI Science of Human History user on 06 November 2018

Huang C-H, et al. 2013. A novel Xanthomonas sp. causes bacterial spot of rose (Rosa spp.). Plant disease 97: 1301-1307.

Lee I, Kim YO, Park S-C, Chun J 2016. OrthoANI: an improved algorithm and software for calculating average nucleotide identity. International Journal of Systematic and Evolutionary Microbiology 66: 1100-1103.

Leyns F, De Cleene M, Swings J-G, De Ley J 1984. The host range of the genusXanthomonas. The Botanical Review 50: 308-356.

Meier-Kolthoff JP, Göker M, Spröer C, Klenk H-P 2013. When should a DDH experiment be mandatory in microbial taxonomy? Archives of Microbiology 195: 413-418.

Midha S, Patil PB 2014. Genomic insights into the evolutionary origin of Xanthomonas axonopodis pv. citri and its ecological relatives. Applied and environmental microbiology 80: 6266-6279.

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

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome research 25: 1043-1055.

Parte AC 2018. LPSN-List of Prokaryotic names with Standing in Nomenclature (bacterio. net), 20 years on. International Journal of Systematic and Evolutionary Microbiology 68: 1825-1829.

Patil PB, Bogdanove AJ, Sonti RV 2007. The role of horizontal transfer in the evolution of a highly variable lipopolysaccharide biosynthesis locus in xanthomonads that infect rice, citrus and crucifers. BMC Evol Biol 7: 243. doi: 1471-2148-7-243 [pii]

10.1186/1471-2148-7-243

Potnis N, et al. 2011. Comparative genomics reveals diversity among xanthomonads infecting tomato and pepper. BMC Genomics 12: 146. doi: 1471-2164-12-146 [pii]

10.1186/1471-2164-12-146

Rademaker J, et al. 2005. A comprehensive species to strain taxonomic framework for Xanthomonas. Phytopathology 95: 1098-1111.

Richard D, et al. 2017. Complete Genome Sequences of Six Copper-Resistant Xanthomonas Strains Causing Bacterial Spot of Solaneous Plants, Belonging to X. gardneri, X. euvesicatoria, and X. vesicatoria, Using Long-Read Technology. Genome announcements 5: e01693-01616.

Downloaded from https://academic.oup.com/gbe/advance-article-abstract/doi/10.1093/gbe/evy238/5140196 by MPI Science of Human History user on 06 November 2018

Rodriguez RL, et al. 2012. Genomes-based phylogeny of the genus Xanthomonas. BMC Microbiol 12: 43. doi: 1471-2180-12-43 [pii]

10.1186/1471-2180-12-43

Samanta JN, Mandal K, Maiti S 2013. A novel pathovar of Xanthomonas axonopodis causes gumming of Guggal (Commiphora wightii). European journal of plant pathology 135: 115-125.

Seemann T 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30: 2068-2069.

Sullivan MJ, Petty NK, Beatson SA 2011. Easyfig: a genome comparison visualizer. Bioinformatics 27: 1009-1010.

Van den Mooter M, Swings J 1990. Numerical analysis of 295 phenotypic features of 266 Xanthomonas strains and related strains and an improved taxonomy of the genus. International Journal of Systematic and Evolutionary Microbiology 40: 348-369.

Vauterin L, Hoste B, Kersters K, Swings J 1995. Reclassification of Xanthomonas. Int. J. Syst. Bacteriol. 45: 472-489.

Walker BJ, et al. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9: e112963.

Wick RR, Judd LM, Gorrie CL, Holt KE 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS computational biology 13: e1005595.

Young J, Dye D, Bradbury J, Panagopoulos C, Robbs C 1978. A proposed nomenclature and classification for plant pathogenic bacteria. New Zealand Journal of Agricultural Research 21: 153-177.

Downloaded from https://academic.oup.com/gbe/advance-article-abstract/doi/10.1093/gbe/evy238/5140196 by MPI Science of Human History user on 06 November 2018

**Figure legends:**

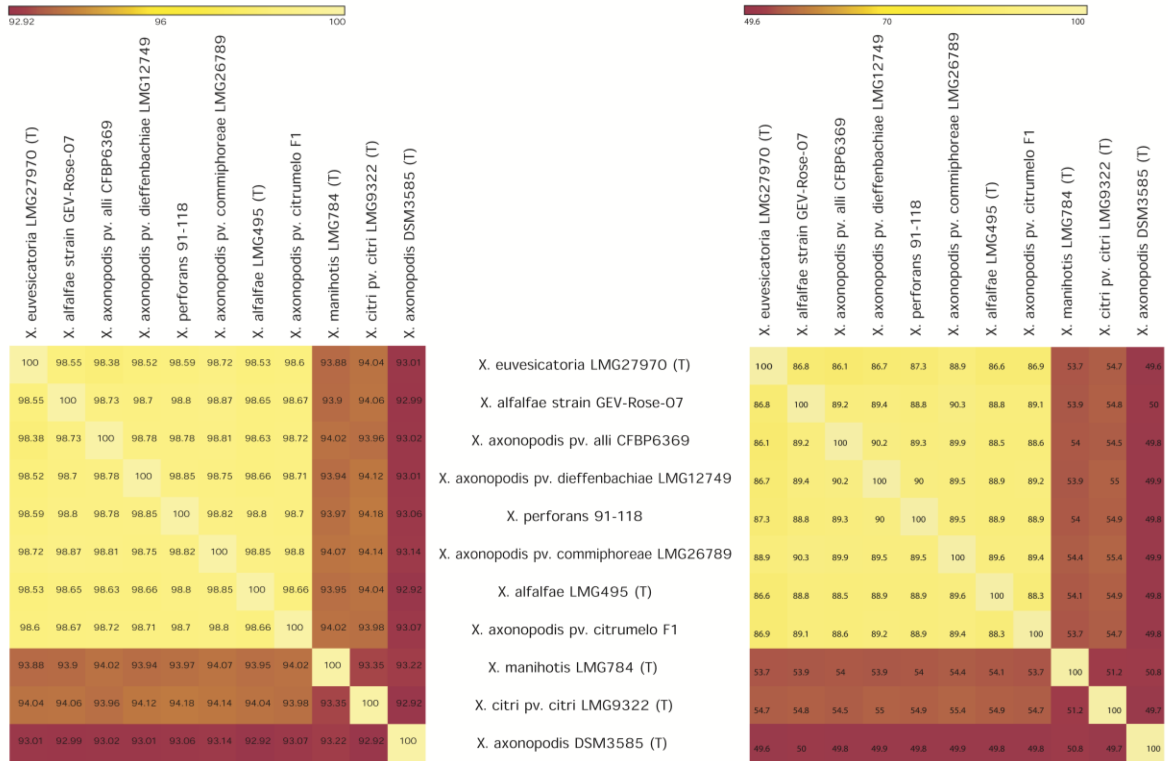
**Figure 1:** Heatmap displaying orthoANI (right hand side) and dDDH (left hand side) values showing Xcom belongs to *X. euvesicatoria* based on >96% ANI and >70% dDDH cutoff values.

**Figure 2: Genealogical analysis of recombination and mutation events of *X. euvesicatoria* pathovars as depicted by ClonalFrameML**. Strains isolated from asterids or monocots or rosids is indicated along with the tree. Here, the variations detected by comparing each clade with its most recent common ancestor are depicted in the graph. Substitutions are represented by vertical lines and recombination events by dark blue horizontal bars. Light blue vertical lines represent no substitution and white lines refer to non homoplasic substitutions. Any other color represents homoplastic substitutions with increase in homoplasy associated with increase in degree of redness from white to red. Xp: *X.* perforans; Xe: *X.* euvesicatoria; XafGEV-Rose-07: *X. alfalfa* strain GEV-Rose-07; Xdf12749: *X. axonopodis* pv. dieffenbachiae LMG12749; Xal6369: *X. axonopodis* pv. alli CFBP6369; XacF1: *X. aconopodis* pv. citrumelo F1; Xcom26789: *X. euvesicatoria* pv. commiphorae LMG26789.

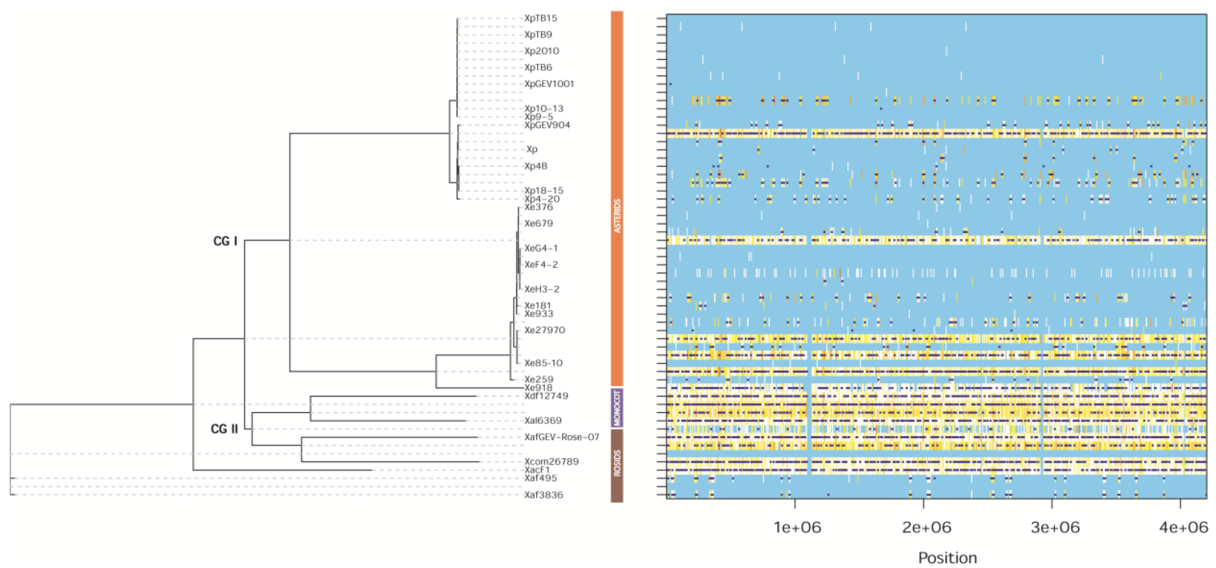
Downloaded from https://academic.oup.com/gbe/advance-article-abstract/doi/10.1093/gbe/evy238/5140196 by MPI Science of Human History user on 06 November 2018

# Figures

## Figure 1




## Figure 2




Downloaded from https://academic.oup.com/gbe/advance-article-abstract/doi/10.1093/gbe/evy238/5140196 by MPI Science of Human History user on 06 November 2018