ACCESS MICROBIOLOGY
an open research platform

SHORT COMMUNICATION
Harrison *et al.*, *Access Microbiology* 2023;5:000532.v3
DOI 10.1099/acmi.0.000532.v3

MICROBIOLOGY SOCIETY
OPEN DATA OPEN ACCESS

Draft genome sequences for ten strains of *Xanthomonas* species that have phylogenomic importance

Jamie Harrison[1], Rana Muhammad Fraz Hussain[2,3], Shannon F. Greer[2,3], Vardis Ntoukakis[2], Andrew Aspin[4], Joana G. Vicente[3,4], Murray Grant[2] and David J. Studholme[1,*]

Abstract

Here we report draft-quality genome sequences for pathotype strains of eight plant-pathogenic bacterial pathovars: *Xanthomonas campestris* pv. *asclepiadis*, *X. campestris* pv. *cannae*, *X. campestris* pv. *esculenti*, *X. campestris* pv. *nigromaculans*, *X. campestris* pv. *parthenii*, *X. campestris* pv. *phormiicola*, *X. campestris* pv. *zinniae* and *X. dyei* pv. *eucalypti* (= *X. campestris* pv. *eucalypti*). We also sequenced the type strain of species *X. melonis* and the unclassified *Xanthomonas* strain NCPPB 1067. These data will be useful for phylogenomic and taxonomic studies, filling some important gaps in sequence coverage of *Xanthomonas* phylogenetic diversity. We include representatives of previously under-sequenced pathovars and species-level clades. Furthermore, these genome sequences may be useful in elucidating the molecular basis for important phenotypes, such as biosynthesis of coronatine-related toxins and degradation of fungal toxin cercosporin.

DATA SUMMARY

Protocols used in this study are deposited at protocols.io:

- Culturing *Xanthomonas* strains: https://dx.doi.org/10.17504/protocols.io.ewov1nr92gr2/v1.
- Preparing genomic DNA: https://dx.doi.org/10.17504/protocols.io.5jyl89428v2w/v1.
- Genome sequence assembly from Illumina reads: https://dx.doi.org/10.17504/protocols.io.kxygxzrqzv8j/v1.
- Phylogenomic analysis: https://dx.doi.org/10.17504/protocols.io.261geny57g47/v1.

The sequence data described in this article are available under BioProject accession PRJNA774146 via https://www.ncbi.nlm.nih.gov/bioproject/PRJNA774146/. Accession numbers of individual genome assemblies are listed in Table 1. Configuration files and tree files are available from GitHub at https://github.com/davidjstudholme/phylogenomics-Xanthomonas-1.

INTRODUCTION

The genus *Xanthomonas* is comprised of Gram-negative bacteria that are usually associated with plants and it includes causative agents of economically important disease of crops such as brassicas, rice, cassava, banana and tomatoes [1]. The taxonomy of *Xanthomonas* spp. has a long and sometimes confusing history. The List of Prokaryotic names with Standing in Nomenclature (accessed 3 April 2023) lists 35 validly named species for the genus, plus several synonyms and others that have not been validly published [2]. Many *Xanthomonas* species are further divided into intra-specific groups called pathovars, which are defined primarily by their host range and in some cases also by biochemical or physiological differences [3].

Received 23 November 2022; Accepted 25 June 2023; Published 14 July 2023
Author affiliations: [1]Biosciences, University of Exeter, Exeter, EX4 4QD, UK; [2]Gibbet Hill Campus, School of Life Sciences, University of Warwick, Coventry, CV4 7AL, UK; [3]Wellesbourne Campus, School of Life Sciences, University of Warwick, Coventry, CV35 9EF, UK; [4]Fera Science Ltd., York Biotech Campus, Sand Hutton, York, YO41 1LZ, UK.
***Correspondence:** David J. Studholme, d.j.studholme@exeter.ac.uk
Keywords: genome sequencing; phylogenomics; *Xanthomonas*; taxonomy.
Abbreviations: ANI, average nucleotide identity; CFBP, Collection Francaise des Bacteries Phytopathogenes; dDDH, digital DNA-DNA hybridization; gyrB, DNA gyrase subunit B; HMW, high molecular weight; HSP, high-scoring pair; LMG, Laboratorium voor Microbiologie; MLSA, multi-locus sequence analysis; NCBI, National Center for Biotechnology Information; NCPPB, National Collection of Plant-Pathogenic Bacteria; PGAP, Prokaryotic Genome Annotation Pipeline; PhaME, phylogenetics and molecular evolution analysis tool; pv., pathovar; slc, species-level clade; SPAdes, St. Petersburg genome assembler; WHRI, Warwick Horticulture Research International.
Two supplementary tables are available with the online version of this article.

Recent efforts have attempted to reconcile the taxonomy of *Xanthomonas* species and pathovars with their phylogeny, often informed by sequence data from one or several genes or from genome-wide sequence data; this has led to the proposal of new species [4–8] and the transfer of pathovars from one species to another [9–19]. Analysis of partial DNA sequences of the *gyrB* gene suggested that many pathovars of the species *X. campestris* are phylogenetically not closely related to the type strain of this species; they are much closer to other named species or so-called species-level clades (slc) [20].

Although partial *gyrB* gene sequences are available for many of these taxonomically incongruent strains, genome-wide sequence data would facilitate phylogenomic studies at a higher resolution. Here, we present draft-quality genome sequences for strains that nominally belong to *X. campestris*, but whose inclusion in that species is incongruent with phylogeny [6, 20–23]; evolutionarily, they fall within species '*X. cannabis*', *X. dyei*, *X. hortorum*, *X. melonis*, *X. sacchari*, Slc 4 and Slc 6. Furthermore, we sequenced strain NCPPB 4013 [24], representing a putative new species, and the type strain of *X. melonis*, thereby filling some gaps in the resources for phylogenomics of *Xanthomonas*.

METHODS

Bacterial strains were obtained from the National Collection of Plant Pathogenic Bacteria (https://www.fera.co.uk/ncppb). We cultured the bacteria in King's B liquid medium at 28 °C as described in protocols.io at https://dx.doi.org/10.17504/protocols.io.ewov1nr92gr2/v1. Genomic DNA was extracted from frozen bacterial pellets using the Qiagen 67 563 MagAttract HMW DNA Kit, following protocol described at https://dx.doi.org/10.17504/protocols.io.5jyl89428v2w/v1. Whole-genome shotgun sequencing was performed on the Illumina NovaSeq 6000 platform to generate pairs of 151-nucleotide sequence reads. We subjected the resulting sequence reads to quality control with fastp [25], *de novo* assembly with SPAdes version 3.15.1 [26] and polishing with Pilon [27] as documented in protocols.io at https://www.protocols.io/view/de-novo-assembly-of-xanthomonas-genomes-from-illum-kxygxzrqzv8j/v1. We generated a phylogenomic tree from genome sequence assemblies using PhaME [28] and FastTree [29] as described in the protocol at https://protocols.io/view/phylogenomic-analysis-of-xanthomonas-ces2tege. Average nucleotide identities (ANI) were calculated using FastANI [30] as described in the protocol at https://www.protocols.io/view/fastani-analysis-protocol-261ge36bol47/v1. For calculating dDDH values, we used the Type Strain Genome Server [31], using formula two (sum of all identities found in high-scoring pairs [HSP] divided by total HSP length) [32]. Genome assemblies were submitted to GenBank [33] via the NCBI Submission Portal [34], after which they were annotated via the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) version 5.3 [35]. Assembly quality was assessed using CheckM version 1.2.2 [36].

RESULTS AND DISCUSSION

We assembled draft genome sequences *de novo* from short sequence reads for ten strains of *Xanthomonas* species and deposited these in GenBank under the accession numbers listed in Table 1. These strains include the type strain for *X. melonis*, pathotype strains for eight pathovars of *X. campestris* and one strain (from radish) not assigned to any named species. For nine of the ten newly sequenced genomes, partial *gyrB* sequences are available under GenBank accessions EU007531.1, EU285211.1, EU285210.1, EU285202.1, EU285197.1, EU285181.1, EU285180.1, EU285177.1 and EU285064.1. In all cases, BLASTN searches confirmed 100% nucleotide sequence identity over the full length of the partial *gyrB* gene versus the corresponding genome sequence assembly. CheckM estimated contamination levels at below 5% for all assemblies except for that of *X. campestris* pv. *nigromaculans* (Tables 1 and S1), available in the online version of this article. Pairwise ANI values are tabulated in Table S2.

'X. cannabis'

The species '*X. cannabis*' was proposed to include *X. campestris* pv. *cannabis*, and *X. sp.* strain Nyagatare isolated from beans [37, 38], though according to The List of Prokaryotic names with Standing in Nomenclature [2] (accessed 29 March 2023) this name is not yet validly published. Originally, *X. cannabis* was proposed in 1955 when Okabe and Goto transferred *Pseudomonas cannabis* [39] into the genus *Xanthomonas* [40]; however, no type strain was deposited. In 2014, Netsu and colleagues [41] speculated that a Japanese isolate identified as *X. campestris* pv. *cannabis* might represent the same pathogen [39] previously described as '*P. cannabis*' and '*X. cannabis*'.

According to Parkinson [20], *X. campestris* pv. *cannabis* falls within Slc 1, along with pathotype strains of *X. campestris* pv. *zinniae* (NCPPB 2439) and *X. campestris* pv. *esculenti* (NCPPB 2190), originally '*X. rubrisorghi*' [20]. This inclusion was based on partial *gyrB* gene sequences; genomes had not been previously sequenced. Our sequencing and phylogenomic analysis (Fig. 1, Table 2) confirm the close relationship between these pathovars and *X. campestris* pv. *cannabis*. Furthermore, they share 96.58% and 97.67% ANI (70.6% and 79.4% dDDH) with the proposed strain NCPPB 2877, consistent with their belonging to the same species.

A description of leaf spot on *Zinnia* plants in Zimbabwe [42] proposed the causative pathogen to be a *forma specialis* of *X. nigromaculans*, where *X. nigromaculans* was the pathogen responsible for leafspot on burdock [22] now known as *X. campestris* pv. *nigromaculans* and phylogenetically falling within *X. hortorum*. Subsequently, this *Zinnia* pathogen was renamed *X. campestris* pv. *zinniae* [3]. Previous analyses based on the *gyrB* locus [20] and our phylogenomic tree (Fig. 1) both place the *Zinnia* pathogen

as completely distinct from *X. campestris* pv. *nigromaculans* and *X. hortorum*. Rather, it falls within the clade corresponding to '*X. cannabis*' and Slc 1.

In addition to its significance as a pathogen in Africa [42], Asia [43, 44] and Europe [45], *X. campestris* pv. *zinniae* is of interest for its ability to degrade the toxin cercosporin that is produced by phytopathogenic fungi of the genus *Cercosporus* [46–49]. The search for genes encoding the degradation pathway identified an oxidoreductase and a putative transcriptional regulator but also highlighted that additional factors are required for cercosporin degradation [49]. Availability of this draft genome sequence may enable identification of additional genes that could facilitate the engineering of *Cercosporus*-resistant crop plants [49].

X. dyei

Strain NCPPB 2337 causes dieback on *Eucalyptus citriodora* in Australia and was originally described as a new species: *X. eucalypti* [50]. In the major taxonomic revision that saw many species reclassified as pathovars of *X. campestris*, this pathogen was renamed as *X. campestris* pv. *eucalypti* [3], with NCPPB 2337 designated as the pathotype strain. In Parkinson and colleagues' *gyrB*-based phylogenetic analysis [20], it was placed within Slc 2 along with unclassified xanthomonads isolated from *Lobelia* spp. and with *X. campestris* pv. *laureliae* isolated from *Laurelia novae-zelandiae*. Following multi-locus sequence analysis (MLSA), this slc was assigned the species name *X. dyei*, so that NCPPB 2337 became the pathotype strain of *X. dyei* pv. *eucalypti* [12]. Consistent with this, our phylogenomic analysis places *X. dyei* pv. *eucalypti* close to the type strain of *X. dyei* (Fig. 1), which was isolated from *Metrosideros excelsa* [12]. Values of ANI and dDDH were 96.86% and 93.9%, respectively.

X. melonis

We sequenced the genome of NCPPB 3434 (=LMG 8670=CFBP 4644), type strain for species *X. melonis* [16], isolated from *Cucumis melo* in Brazil. This augments the previously available genome sequence for *X. melonis* CFBP 4644 (GenBank: GCA_002940015.1), and several recently sequenced isolates from Trinidad [51]. Strain CFBP 4644 is identical to NCPPB 3434 but it was not flagged as being type material in the metadata of BioSample SAMN05560295. Consistent with previous phylogenomic analyses of *X. melonis* CFBP 4644 [52, 53], the phylogenetic position of NCPPB 3434 fell within a clade shared with *X. cucurbitae*, *X. codiaei* and *X. floridensis* (Fig. 1).

Xanthomonas strain NCPPB 1067 was isolated from *Raphanus sativus* (radish) in Vanuatu. In the previous *gyrB*-based phylogenetic analysis [20], it was the sole representative of Slc 3. Our phylogenomic analysis places it close to the type strain of *X. melonis* (Fig. 1), with which it shares 98.57% ANI and 86.80% dDDH, consistent with its belonging to this species and with Parkinson's Slc corresponding to the species *X. melonis*.

Table 1. Bacterial strains used for genome sequencing. Contamination levels were assessed using CheckM [36]. Additional assembly metrics are provided in the Table S1

Taxonomic name	NCPPB no.	Host of isolation	Phylogenetic position	Genome assembly GenBank accession	Coverage	Contamination (%)
X. campestris pv. *esculenti* (Rangaswami & Easwaran 1962) Dye 1978	2190 (PT)	*Hibiscus esculentus*	'*X. cannabis*' (Slc 1)	GCA_020784125.1	141 x	0.25
X. campestris pv. *zinniae* [42] Dye 1978	2439 (PT)	*Zinnia elegans*	'*X. cannabis*' (Slc 1)	GCA_020783815.1	134 x	1.12
X. dyei pv. *eucalypti* [12, 50] (= *X. campestris* pv. *eucalypti*)	2337 (PT)	*Eucalyptus maculata var. citriodora*	*X. dyei*	GCA_020783675.1	63 x	0.58
X. melonis [16]	3434 (T)	*Cucumis melo*	*X. melonis*	GCA_020783655.1	169 x	0.70
*X*nthomonas sp.	1067	*Raphanus sativus*	*X. melonis*	GCA_020783795.1	32 x	0.37
X. campestris pv. *parthenii* [54]	3888 (PT)	*Parthenium hysterophorus*	Slc 4	GCA_020783765.1	192 x	1.08
X. campestris pv. *phormiicola* [55] Dye 1978	2983 (PT)	*Phormium tenax*	Slc 6	GCA_020783715.1	49 x	4.484
X. campestris pv. *nigromaculans* [22] Dye 1978	1935 (PT)	*Arctium lappa*	*X. hortorum*	GCA_020783725.1	57 x	12.92
X. campestris pv. *cannae* [65]	4345 (PT)	*Canna×generalis*	*X. sacchari*	GCA_020783755.1	122 x	4.24
X. campestris pv. *asclepiadis* [24]	4013 (PT)	*Asclepias syriaca*	Close to *X. euroxanthea*	GCA_020784175.1	180 x	0.89

PT, Pathotype strain; T, Type strain.

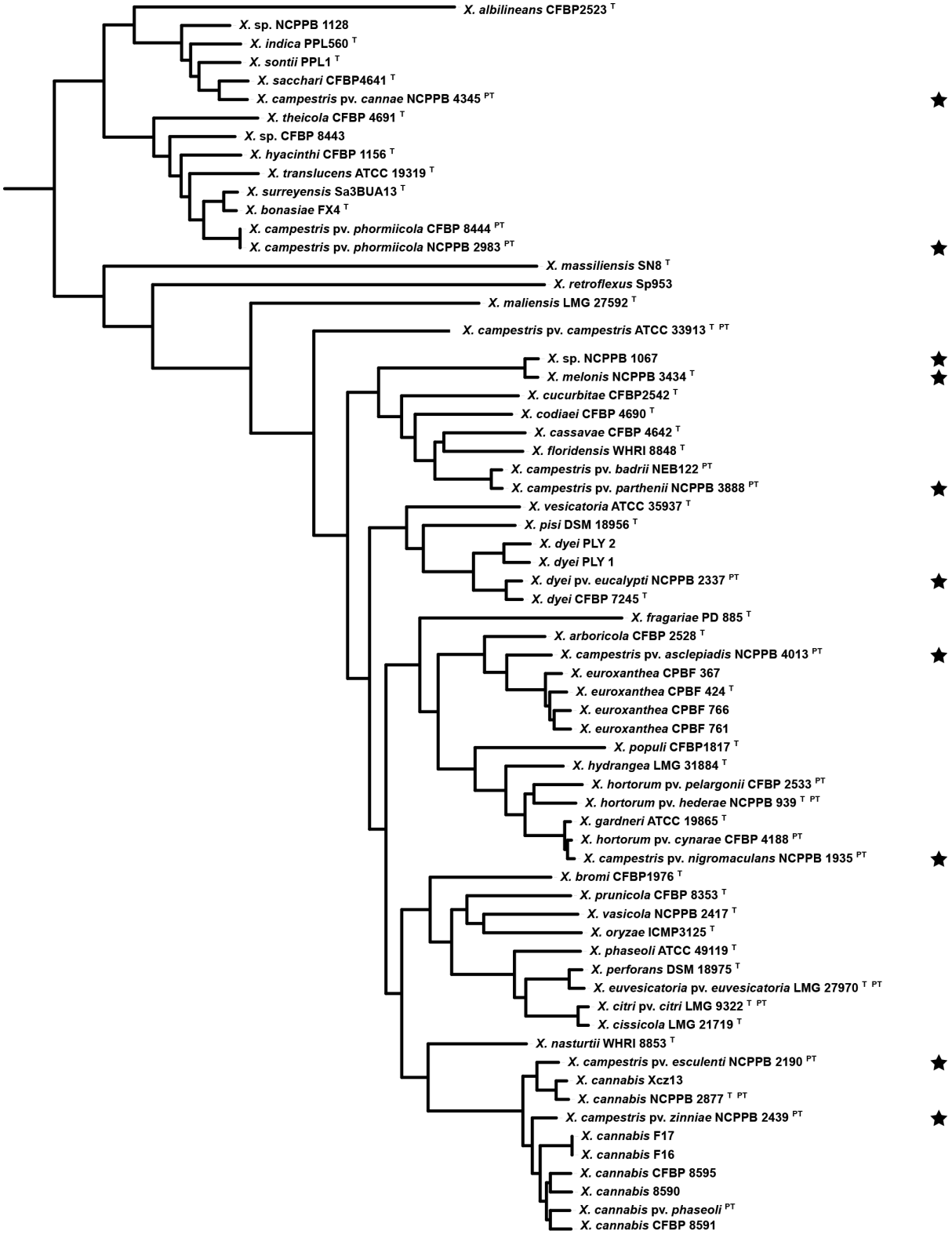

Fig. 1. Phylogenetic tree, based on core-genome sequences, for the newly sequenced strains, type strains and representative strains of *Xanthomonas* spp., generated using PhaME [28] and FastTree [29]. The tree was graphically rendered using the Interactive Tree of Life [67]. Configuration and tree files are available from https://github.com/davidjstudholme/phylogenomics-Xanthomonas-1. Accession numbers and references for the genome assemblies are listed in Table 2. Newly sequenced genomes are indicated with a black star.

Table 2. Genome assemblies used for phylogenomic analysis

GenBank accession	Bacterial strain	References
GCA_002939705.1	*X. albilineans* CFBP2523 ^{T}	–
GCA_001013475.1	*X. arboricola* pv. *juglandis* CFBP 2528 $^{T\ PT}$	[68]
GCA_017163705.1	*X. bonasiae* FX4 ^{T}	[4]
GCA_002939755.1	*X. bromi* CFBP1976 ^{T}	[69]
GCA_020784175.1	*X. campestris* pv. *asclepiadis* NCPPB 4013 ^{PT}	This study
GCA_012848175.1	*X. campestris* pv. *badrii* NEB122 ^{PT}	–
GCA_000007145.1	*X. campestris* pv. *campestris* ATCC 33913 $^{T\ PT}$	[70]
GCA_020783755.1	*X. campestris* pv. *cannae* NCPPB 4345 ^{PT}	This study
GCA_020784125.1	*X. campestris* pv. *esculenti* NCPPB 2190 ^{PT}	This study
GCA_020783725.1	*X. campestris* pv. *nigromaculans* NCPPB 1935 ^{PT}	This study
GCA_938743425.1	*X. campestris* pv. *nigromaculans* NCPPB 1935 ^{PT}	[21]
GCA_020783765.1	*X. campestris* pv. *parthenii* NCPPB 3888 ^{PT}	This study
GCA_025666215.1	*X. campestris* pv. *phormiicola* CFBP 8444 ^{PT}	[56]
GCA_020783715.1	*X. campestris* pv. *phormiicola* NCPPB 2983 ^{PT}	This study
GCA_020783815.1	*X. campestris* pv. *zinniae* NCPPB 2439 ^{PT}	This study
GCA_011761725.1	*X. cannabis* 8590	–
GCA_014195785.1	*X. cannabis* CFBP 8591	–
GCA_014195835.1	*X. cannabis* CFBP 8595	–
GCA_014199075.1	*X. cannabis* F16	–
GCA_014199515.1	*X. cannabis* F17	–
GCA_000802365.1	*X. cannabis* NCPPB 2877 $^{T\ PT}$	[38]
GCA_000764855.1	*X. cannabis* pv. *phaseoli* Nyagatare ^{PT}	[37]
GCA_020880755.1	*X. cannabis* Xcz13	–
GCA_000454545.1	*X. cassavae* CFBP 4642 ^{T}	[71]
GCA_002019225.1	*X. cissicola* LMG 21719 ^{T}	[18]
GCA_002018575.1	*X. citri* pv. *citri* LMG 9322 $^{T\ PT}$	[72]
GCA_002939785.1	*X. codiaei* CFBP 4690	–
GCA_002939885.1	*X. cucurbitae* CFBP2542 ^{T}	–
GCA_002939865.1	*X. dyei* CFBP 7245 ^{T}	–
GCA_003363885.1	*X. dyei* PLY_1	[73]
GCA_003363875.1	*X. dyei* PLY_2	[73]
GCA_020783675.1	*X. dyei* pv. *eucalypti* NCPPB 2337 ^{PT}	This study
GCA_903989455.1	*X. euroxanthea* CPBF 367	[74]
GCA_900476395.1	*X. euroxanthea* CPBF 424 ^{T}	[75, 76]
GCA_905367725.1	*X. euroxanthea* CPBF 761	[76]
GCA_905367735.1	*X. euroxanthea* CPBF 766	[76]
GCA_001401555.1	*X. euvesicatoria* pv. *euvesicatoria* LMG 27970 $^{T\ PT}$	–
GCA_001642575.1	*X. floridensis* WHRI 8848 ^{T}	[8]
GCA_900380235.1	*X. fragariae* PD 885 ^{T}	[77]
GCA_000192065.2	*X. gardneri* ATCC 19865 ^{T}	[78]

Continued

Table 2. Continued

GenBank accession	Bacterial strain	References
GCA_021352995.1	*X. hortorum* pv. *cynarae* CFBP 4188 ^{PT}	[79]
GCA_003064105.1	*X. hortorum* pv. *hederae* NCPPB 939 $^{T\ PT}$	[80]
GCA_021353095.1	*X. hortorum* pv. *pelargonii* CFBP 2533 ^{PT}	–
GCA_009769165.1	*X. hyacinthi* CFBP 1156 ^{T}	[81]
GCA_905142475.1	*X. hydrangea* LMG 31884 ^{T}	[6]
GCA_022669045.1	*X. indica* PPL560 ^{T}	[5]
GCA_009192945.1	*X. maliensis* LMG 27592 ^{T}	[82]
GCA_900018785.1	*X. massiliensis* SN8 ^{T}	[83]
GCA_020783655.1	*X. melonis* NCPPB 3434 ^{T}	This study
GCA_001660815.1	*X. nasturtii* WHRI 8853 ^{T}	[8]
GCA_004136375.1	*X. oryzae* ICMP3125 ^{T}	[84]
GCA_013112235.1	*X. perforans* DSM 18975 ^{T}	–
GCA_022749655.1	*X. phaseoli* ATCC 49119 ^{T}	–
GCA_001010415.1	*X. pisi* DSM 18956 ^{T}	[85]
GCA_002940065.1	*X. populi* CFBP1817 ^{T}	–
GCA_002846205.1	*X. prunicola* CFBP 8353 ^{T}	[86]
GCA_900143175.1	*X. retroflexus* Sp953	[87]
GCA_002940085.1	*X. sacchari* CFBP4641 ^{T}	[7]
GCA_008119715.1	*X. sontii* PPL1 ^{T}	[7]
GCA_020783795.1	*Xanthomonas* sp. NCPPB 1067	This study
GCA_014836395.1	*X. surreyensis* Sa3BUA13 ^{T}	[88]
GCA_014236795.1	*X. theicola* CFBP 4691 ^{T}	[89]
GCA_020880735.1	*X. translucens* pv. *translucens* ATCC 19319 $^{T\ PT}$	[90]
GCA_000772705.2	*X. vasicola* NCPPB 2417 ^{T}	[19]
GCA_001908725.1	*X. vesicatoria* ATCC 35937 ^{T}	[91]
GCA_025666195.1	*Xanthomonas* sp. CFBP 8443	[56]
GCA_025666235.1	*Xanthomonas* sp. CFBP 8445	[56]
GCA_001043115.1	*Xanthomonas* sp. NCPPB 1128	[92]

PT, Pathotype strain; T, Type strain.

X. campestris pv. *parthenii*: a potential new species (Slc 4)

Xanthomonas campestris pv. *parthenii* NCPPB 3880 causes leaf blight on *Parthenium*, an obnoxious exotic weed introduced into India in 1955, and might have potential as a biological control agent [54]. The pathotype strain NCPPB 3880 fell within Slc 4 along with *X. campestris* pv. *badrii* according to previous *gyrB* analysis [20]. In our phylogenomic analysis (Fig. 1), *X. campestris* pv. *parthenii* and *X. campestris* pv. *badrii* are closely related to each other, sharing 98.88% ANI and 89.6% dDDH. They appear as a distinct species neighbouring *X. floridensis* (93.66–93.68% ANI and 52.2% dDDH) and *X. cassavae* (93.31–93.35% ANI and 50.3–50.7% dDDH).

X. campestris pv. *phormiicola*: a potential new species (Slc 6)

In 1933, Takimoto [55] described *Bacterium phormicola* as the cause of bacterial streak on New Zealand flax (*Phormium tenax*). This pathogen was subsequently classified as *X. campestris* pv. *phormiicola* [3], though Parkinson and colleagues' later examination of *gyrB* sequences [20] placed it as the sole member of Slc 6, outside of *X. campestris*. Our genome-based phylogenetic analysis of NCPPB 2983 confirmed this (Fig. 1) and placed it as a distinct species close to *X. hyacinthi* and *X. translucens* [56]. It shares 95% ANI with *X. bonasiae* FX4 (Table S2), which is on the borderline of the commonly used criterion for species delimitation.

Based on dDDH values, the Type Strain Genome Server identifies that this genome does not fall within any named species and potentially represents a new species dDDH between NCPPB 2983 and *X. bonasiae* FX4 is 56.1. The recent study by Peduzzi and colleagues came to a similar conclusion [56].

Xanthomonas campestris pv. *phormiicola* is unusual among xanthomonads in that it produces the phytotoxin coronatine [57] and/or related derivatives of coronofacic acid [49], a trait more usually associated with *Pseudomonas* spp. Tamura and colleagues [57] hypothesized that phytotoxin biosynthesis might be encoded on a plasmid but failed to isolate plasmid DNA. The availability of genome sequence opens the possibility of identifying the genetic basis for this trait. After our submission of the first version of this manuscript, a high-quality genome sequence assembly was published for *Xanthomonas campestris* pv. *phormiicola* and candidate biosynthesis genes were identified [56].

X. campestris pv. *asclepiadis*: a potential new species

Bacterial blight of milkweed (*Asclepias spp*.) is attributed to *X. campestris* pv. *asclepiadis* [24]. No sequence data was available for this pathogen in GenBank and therefore its phylogenetic position within the genus was unclear. Our analysis places the pathotype strain NCPPB 4013 very distant from the type strain of *X. campestris*, and closer to *X. euroxanthea* (Fig. 1), a species isolated from walnut trees (*Juglans regia*) [58]. The ANI between NCPPB 4013 and the *X. euroxanthea* type strain is 95.97%, close to the threshold of 95–96% that is commonly used to delineate species boundaries [30, 59–64]. The Type Strain Genome Server, calculating a dDDH value of 79.% with the *X. euroxanthea* type strain, reports that NCPPB 4013 does not belong to *X. euroxanthea* and may represent a new species.

X. campestris pv. *cannae* (*X. sacchari*)

The pathogen responsible for a bacterial leaf spot and leaf blight on canna (*Canna* x *generalis*) in India was described as *X. campestris* pv. *cannae* [65], though phylogenetic analysis of its *gyrB* gene sequence places the pathotype strain NCPPB 4345 in the *X. sacchari* clade [20]. Our phylogenomic analysis (Fig. 1) confirms a close relationship between NCPPB 4345 and the type strain of *X. sacchari*. However, the current species description for *X. sacchari* [16] states that 'The strains of this species are isolated from diseased sugarcane' and therefore would exclude this pathogen, which was isolated from canna rather than sugarcane. Currently, *X. sacchari* is not subdivided into pathovars, since there is no variation in the host range. This species is distinguished from other xanthomonads by its metabolic activity on a long list of carbon substrates [16], only a subset of which have been tested on *X. campestris* pv. *cannae*. Transfer of this pathovar into *X. sacchari* would require revision of its species description to encompass isolation hosts beyond sugarcane.

X. campestris pv. *nigromaculans*

The causative agent of black spot on burdock (*Arctium lappa*) was originally described in 1927 under the name *Bacterium nigromaculans* [22] and subsequently renamed as *X. campestris* pv. *nigromaculans* [3]. Several authors have proposed, based on sequence-based phylogenetic analysis, that *X. campestris* pv. *nigromaculans* is more closely related to the species *X. hortorum* [6, 9, 10, 20, 21] than to the type strain of *X. campestris*. This proposal was initially based on just a single genetic locus, *gyrB* [20]. Later, Dia and colleagues arrived at the same conclusion on the basis of both MLSA and phylogenomic analysis of the core genome, but they did not make their *X. campestris* pv. *nigromaculans* genome sequence publicly available at that time [6], motivating our sequencing of its pathotype strain NCPPB 1935. Subsequently, Dia and colleagues publicly deposited their sequence data for this same strain [21]. Our phylogenomic analysis (Fig. 1) is consistent with the previous studies, placing this strain clearly into Dia's sub-cluster A [6] within the *X. hortorum* clade, close to *X. hortorum* pv. *gardneri* and *X. hortorum* pv. *cynarae*.

Dia and colleagues [6] cite a report by Dehgan-Niri and Rahimian [66] when mentioning *X. campestris* pv. *nigromaculans*. That report identifies a leaf-spot pathogen of burdock (same host as pv. *nigromaculans*) to be '*X. gardneri*', i.e. *X. hortorum* pv. *gardneri*, based on partial sequence of its *gyrB* gene [66]. Although Dehgan-Niri and Rahimian did not mention *X. campestris* pv. *nigromaculans*, their *gyrB* sequence shares 100% nucleotide sequence identity with *X. campestris* pv. *nigromaculans* NCPPB 1935, as well as several *X. hortorum* strains; therefore, their burdock isolates are probably indeed pv. *nigromaculans*.

In summary, our genome sequencing and phylogenetic reconstruction supports previous suggestions that *X. campestris* pv. *nigromaculans* falls within a clade that corresponds to the species *X. hortorum* [6, 9, 10, 20, 21] and should be transferred to that species. We also note that our genome assembly for *X. campestris* pv. *nigromaculans* has a rather high level of contamination of 12.92% whereas assembly GCA_938743425.1 [21] is of higher quality, with a contamination level of 0.81% (Tables 1 and S1).

Conclusion

Here we present draft-quality genome sequences for ten plant-pathogenic bacterial strains from the National Collection of Plant Pathogenic Bacteria (NCPPB). The data will be useful for phylogenomic studies, filling some important gaps in sequence coverage of the *Xanthomonas* phylogenetic diversity. Furthermore, these genome sequences may be useful in elucidating

the molecular basis for important phenotypes such as biosynthesis of coronatine-related toxins and degradation of fungal toxin cercosporin.

Funding information
This work was funded by a grant from BBSRC, NERC, Defra, and the Scottish Government, under the Strategic Priorities Fund Plant Bacterial Diseases programme under the project 'Xanthomonas plants diseases: mitigating existing, emerging and future threats to UK agriculture' (BB/T010916/1, BB/T010908/1 and BB/T010924/1). This project utilised DNA sequencing equipment funded by the Wellcome Trust (Multi-User Equipment Grant award number 218247/Z/19/Z). We thank Shea Bailey from the National Collection of Plant Pathogenic Bacteria (NCPPB) for assistance in supplying the strains.

Acknowledgements
We are grateful for the excellent technical assistance of the University of Exeter's sequencing facility in performing library preparation and DNA sequencing and delivery of the sequencing data. We also acknowledge the use of the University of Exeter's Isca high-performance computing facility.

Author contributions
J.H. performed formal analysis, data curation and investigation, including computational methods development, sequence assembly and quality control. R.M.F.H. and S.G. performed formal analysis and investigation, including development and implementation of laboratory methods for bacterial culture and extraction of high-quality genomic DNA. A.A. was responsible for provision of resources, i.e. bacterial strains from the NCPPB. J.V. resources and supervision at Fera. M.G. and V.N. were responsible for supervision of the Warwick team. M.G. developed methodology and performed investigation (DNA extraction). D.S. was responsible for writing the original draft, formal analysis (phylogenomics) supervision at Exeter and data curation. The original conceptualization of this study arose from D.S. and J.V. All authors were involved in discussions about the work and all contributed to writing (review and editing) the final version of the manuscript.

Conflicts of interest
The author(s) declare that there are no conflicts of interest.

References

1. Timilsina S, Potnis N, Newberry EA, Liyanapathiranage P, Iruegas-Bocardo F, *et al*. *Xanthomonas* diversity, virulence and plant-pathogen interactions. *Nat Rev Microbiol* 2020;18:415–427.
2. Parte AC, Sardà Carbasse J, Meier-Kolthoff JP, Reimer LC, Göker M. List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. *Int J Syst Evol Microbiol* 2020;70:5607–5612.
3. Young JM, Dye DW, Bradbury JF, Panagopoulos CG, Robbs CF. A proposed nomenclature and classification for plant pathogenic bacteria. *New Zeal J Agric Res* 1978;21:153–177.
4. Mafakheri H, Taghavi SM, Zarei S, Portier P, Dimkić I, *et al*. *Xanthomonas bonasiae* sp. nov. and *Xanthomonas youngii* sp. nov., isolated from crown gall tissues. *Int J Syst Evol Microbiol* 2022;72:005418.
5. Rana R, Madhavan VN, Saroha T, Bansal K, Kaur A, *et al*. *Xanthomonas indica* sp. nov., a novel member of non-pathogenic *Xanthomonas* community from healthy rice seeds. *Curr Microbiol* 2022;79:304.
6. Dia NC, Van Vaerenbergh J, Van Malderghem C, Blom J, Smits THM, *et al*. *Xanthomonas hydrangeae* sp. nov., a novel plant pathogen isolated from *Hydrangea arborescens*. *Int J Syst Evol Microbiol* 2021;71:005163.
7. Bansal K, Kaur A, Midha S, Kumar S, Korpole S, *et al*. *Xanthomonas sontii* sp. nov., a non-pathogenic bacterium isolated from healthy basmati rice (Oryza sativa) seeds from India. *Antonie van Leeuwenhoek* 2021;114:1935–1947.
8. Vicente JG, Rothwell S, Holub EB, Studholme DJ. Pathogenic, phenotypic and molecular characterisation of *Xanthomonas nasturtii* sp. nov. and *Xanthomonas floridensis* sp. nov., new species of *Xanthomonas* associated with watercress production in Florida. *Int J Syst Evol Microbiol* 2017;67:3645–3654.
9. Morinière L, Burlet A, Rosenthal ER, Nesme X, Portier P, *et al*. Clarifying the taxonomy of the causal agent of bacterial leaf spot of lettuce through a polyphasic approach reveals that *Xanthomonas cynarae* Trébaol et al. 2000 emend. Timilsina et al. 2019 is a later heterotypic synonym of *Xanthomonas hortorum* Vauterin e. *Syst Appl Microbiol* 2020;43:126087.
10. Jones JB, Lacy GH, Bouzar H, Stall RE, Schaad NW. Reclassification of the *Xanthomonads* associated with bacterial spot disease of tomato and pepper. *Syst Appl Microbiol* 2004;27:755–762.
11. Hu FP, Young JM, Stead DE, Got M, Goto M. Transfer of *Pseudomonas Cissicola* (Takimoto 1939) Burkholder 1948 to the genus *Xanthomonas*. *Int J Syst Bacteriol* 1997;47:228–230.
12. Young JM, Wilkie JP, Park DC, Watson DRW. New Zealand strains of plant pathogenic bacteria classified by multi-locus sequence analysis; proposal of *Xanthomonas dyei* sp. nov. *Plant Pathol* 2010;59:270–281.
13. Schaad NW, Postnikova E, Lacy GH, Sechler A, Agarkova I, *et al*. Reclassification of *Xanthomonas campestris* pv. citri (ex Hasse 1915) Dye340 1978 forms A, B/C/D, and E as *X. smithii* subsp. *citri* (ex Hasse) sp. nov. nom. rev. comb. nov.,341 *X. fuscans* subsp. *aurantifolii* (ex Gabriel 1989) sp. nov. nom. rev. comb. nov., and X. *Syst Appl Microbiol* 2005;28:494–518.
14. da Gama MAS, Mariano R de LR, da Silva Júnior WJ, de Farias ARG, Barbosa MAG, *et al*. Taxonomic repositioning of *Xanthomonas campestris* pv. *viticola* (Nayudu 1972) Dye 1978 as *Xanthomonas citri* pv. *viticola* (Nayudu 1972) Dye 1978 comb. nov. and emendation of the description of *Xanthomonas citri* pv. *anacardii* to include pigmented isolates pathogenic to cashew plant. *Phytopathology* 2018;108:1143–1153.
15. Lang JM, DuCharme E, Ibarra Caballero J, Luna E, Hartman T, *et al*. Detection and characterization of *Xanthomonas vasicola* pv. *vasculorum* (Cobb 1894) comb. nov. causing bacterial leaf streak of corn in the United States. *Phytopathology* 2017;107:1312–1321.
16. Vauterin L, Hoste B, Kersters K, Swings J. Reclassification of *Xanthomonas*. *Int J Syst Bacteriol* 1995;45:472–489.
17. Bull CT, De Boer SH, Denny TP, Firrao G, Saux M-L, *et al*. Comprehensive list of names of plant pathogenic bacteria, 1980-2007. *J Plant Pathol* 2010;92:551–592.
18. Bansal K, Kumar S, Patil PB. Phylo-taxonogenomics supports revision of taxonomic status of twenty *Xanthomonas pathovars* to *Xanthomonas citri*. *Phytopathology* 2021.
19. Studholme DJ, Wicker E, Abrare SM, Aspin A, Bogdanove A, *et al*. Transfer of *Xanthomonas campestris* pv. *arecae* and *X. campestris* pv. *musacearum* to *X. vasicola* (*Vauterin*) as *X. vasicola* pv. *arecae* comb. nov. and *X. vasicola* pv. *musacearum* comb. nov. and description of *X. vasicola* pv. *vasculorum* pv. nov. *Phytopathology* 2020;110:1153–1160.
20. Parkinson N, Cowie C, Heeney J, Stead D. Phylogenetic structure of *Xanthomonas* determined by comparison of gyrB sequences. *Int J Syst Evol Microbiol* 2009;59:264–274.

21. Dia NC, Morinière L, Cottyn B, Bernal E, Jacobs JM, *et al*. *Xanthomonas hortorum* - beyond gardens: current taxonomy, genomics, and virulence repertoires. *Mol Plant Pathol* 2022;23:597–621.
22. Takimoto S. Bacterial black spot of burdock. *J Plant Prot Tokyo* 1927;14:519–523.
23. Dye DW, Bradbury JF, Goto M, Hayward AC, Lelliott RA, *et al*. International standards for naming pathovars of phytopathogenic bacteria and a list of pathovar names and pathotype strains. *Rev Plant Pathol* 1980;59:153–168.
24. Flynn P, Vidaver AK. *Xanthomonas campestris* pv. *asclepiadis*, pv. nov., causative agent of bacterial blight of Milkweed (*Asclepias* spp.). *Plant Dis* 1995;79:1176.
25. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018;34:i884–i890.
26. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, *et al*. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19:455–477.
27. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, *et al*. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 2014;9:e112963.
28. Shakya M, Ahmed SA, Davenport KW, Flynn MC, Lo C-C, *et al*. Standardized phylogenetic and molecular evolutionary analysis applied to species across the microbial tree of life. *Sci Rep* 2020;10:1723.
29. Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* 2010;5:e9490.
30. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018;9:5114.
31. Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 2019;10:2182.
32. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:60.
33. Sayers EW, Cavanaugh M, Clark K, Ostell J, Pruitt KD, *et al*. GenBank. *Nucleic Acids Res* 2019;47:D94–D99.
34. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, *et al*. Database resources of the national center for biotechnology information. *Nucleic Acids Res* 2022;50:D20–D26.
35. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, *et al*. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 2016;44:6614–6624.
36. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 2015;25:1043–1055.
37. Aritua V, Musoni A, Kabeja A, Butare L, Mukamuhirwa F, *et al*. The draft genome sequence of *Xanthomonas* species strain Nyagatare, isolated from diseased bean in Rwanda. *FEMS Microbiol Lett* 2015;362:1–4.
38. Jacobs JM, Pesce C, Lefeuvre P, Koebnik R. Comparative genomics of a cannabis pathogen reveals insight into the evolution of pathogenicity in *Xanthomonas*. *Front Plant Sci* 2015;6:431.
39. Watanabe T. *Sen-i Sakumotsu Byo Gaku*. Tokyo: Asakura Publishing; 1947.
40. Okabe N, Goto M. Bacterial plant diseases in Japan: I. a list of bacterial diseases and their pathogens. *Rep Fac Agric Shizuoka Univ* 1955;5:63–71.
41. Netsu O, Kijima T, Takikawa Y. Bacterial leaf spot of hemp caused by *Xanthomonas campestris* pv. *cannabis* in Japan. *J Gen Plant Pathol* 2014;80:164–168.
42. Hopkins JC, Dowson WJ. A bacterial leaf and flower disease of Zinnia in Southern Rhodesia. *Trans Br Mycol Soc* 1949;32:252–IN5.
43. Myung I-S, Lee JY, Yoo HL, Wu JM, Shim H-S. Bacterial leaf spot of Zinnia caused by *Xanthomonas campestris* pv. *zinniae*, a new disease in Korea. *Plant Dis* 2012;96:1064.
44. Zhao YT, Lu BH, Bai QR, Gao J. First report of *Xanthomonas campestris* pv. *zinniae* causing bacterial leaf and flower spot disease of zinnia elegans in Jilin province, China. *Plant Dis* 2016;100:208.
45. Schwarczinger I, Vajna L, Süle S. First report of bacterial leaf and flower spot of Zinnia elegans caused by *Xanthomonas campestris* pv. *zinniae* in Hungary. *Plant Pathol* 2008;57:367.
46. Daub ME, Ehrenshaft M. The photoactivated cercospora toxin cercosporin: contributions to plant disease and fundamental biology. *Annu Rev Phytopathol* 2000;38:461–490.
47. Mitchell TK, Alejos-Gonzalez F, Gracz HS, Danehower DA, Daub ME, *et al*. *Xanosporic* acid, an intermediate in bacterial degradation of the fungal phototoxin cercosporin. *Phytochemistry* 2003;62:723–732.
48. Mitchell TK, Chilton WS, Daub ME. Biodegradation of the polyketide toxin cercosporin. *Appl Environ Microbiol* 2002;68:4173–4181.
49. Taylor TV, Mitchell TK, Daub ME. An oxidoreductase is involved in cercosporin degradation by the bacterium *Xanthomonas campestris* pv. zinniae. *Appl Environ Microbiol* 2006;72:6070–6078.
50. Truman R. Die-back of eucalyptus citriodora caused by *Xanthomonas eucalypti* sp.n. *Phytopathology* 1974;64:143.
51. Ramnarine SDBJ, Jayaraman J, Ramsubhag A. Comparative genomics of the black rot pathogen *Xanthomonas campestris* pv. *campestris* and non-pathogenic co-inhabitant *Xanthomonas melonis* from Trinidad reveal unique pathogenicity determinants and secretion system profiles. *PeerJ* 2022;9:e12632.
52. Merda D, Briand M, Bosis E, Rousseau C, Portier P, *et al*. Ancestral acquisitions, gene flow and multiple evolutionary Trajectories of the type three secretion system and effectors in *Xanthomonas* plant pathogens. *Mol Ecol* 2017;26:5939–5952.
53. Bansal K, Kumar S, Kaur A, Singh A, Patil PB. Deep phylo-taxono genomics reveals *Xylella* as a variant lineage of plant associated *Xanthomonas* and supports their taxonomic reunification along with *Stenotrophomonas* and *Pseudoxanthomonas*. *Genomics* 2021;113:3989–4003.
54. Chand R, Singh BD, Singh D, Singh PN. *Xanthomonas campestris* pv. *parthenii* pathovar nov. incitant of leaf blight of parthenium. *Antonie van Leeuwenhoek* 1995;68:161–164.
55. Takimoto S. The bacterial disease of New Zealand flax. *J Plant Prot Tokyo* 1933;20:774–778.
56. Peduzzi C, Sagia A, Burokiene D, Nagy IK, Fischer-Le Saux M, *et al*. Complete genome sequencing of three clade-1 *xanthomonads* reveals genetic determinants for a lateral flagellin and the biosynthesis of coronatine-like molecules in *Xanthomonas*. *Phytopathology* 2023.
57. Tamura K, Takikawa Y, Tsuyumu S, Goto M, Watanabe M. Coronatine production by *Xanthomonas campestris* pv. *phormiicola*. *Jpn J Phytopathol* 1992;58:276–281.
58. Martins L, Fernandes C, Blom J, Dia NC, Pothier JF, *et al*. *Xanthomonas euroxanthea* sp. nov., a new xanthomonad species including pathogenic and non-pathogenic strains of walnut. *Int J Syst Evol Microbiol* 2020;70:6024–6031.
59. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci*2009;106:19126–19131.
60. Figueras MJ, Beaz-Hidalgo R, Hossain MJ, Liles MR. Taxonomic affiliation of new genomes should be verified using average nucleotide identity and multilocus phylogenetic analysis. *Genome Announc* 2014;2:e00927-14.
61. Lee I, Ouk Kim Y, Park S-C, Chun J. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* 2016;66:1100–1103.
62. Meier-Kolthoff JP, Klenk H-P, Göker M. Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. *Int J Syst Evol Microbiol* 2014;64:352–356.

63. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:60.
64. Palmer M, Venter SN, Coetzee MPA, Steenkamp ET. Prokaryotic species are sui generis evolutionary units. *Syst Appl Microbiol* 2019;42:145–158.
65. Easwaramurthy R, Kaviyarasan V, Gnanamanickam SS. A bacterial disease of ornamental Cannas caused by *Xanthomonas campestris* pv *cannae* pv nov. *Curr Sci* 1984;53:708–709.
66. Dehgan-Niri M, Rahimian H. First report of *Xanthomonas gardneri* causing bacterial leaf spot on Burdock in Iran. *J Plant Pathol* 2016;98:677.
67. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 2021;49:W293–W296.
68. Garita-Cambronero J, Palacio-Bielsa A, López MM, Cubero J. Comparative genomic and phenotypic characterization of pathogenic and non-pathogenic strains of *Xanthomonas arboricola* reveals insights into the infection process of bacterial spot disease of stone fruits. *PLoS One* 2016;11:e0161977.
69. Hersemann L, Wibberg D, Blom J, Widmer F, Kölliker R. Draft genome sequence of the *Xanthomonas bromi* type strain LMG 947. *Genome Announc* 2016;4:5–6.
70. da Silva ACR, Ferro JA, Reinach FC, Farah CS, Furlan LR, *et al*. Comparison of the genomes of two *Xanthomonas pathogens* with differing host specificities. *Nature* 2002;417:459–463.
71. Bolot S, Munoz Bodnar A, Cunnac S, Ortiz E, Szurek B, *et al*. Draft genome sequence of the *Xanthomonas cassavae* type strain CFBP 4642. *Genome Announc* 2013;1:e00679-13.
72. Gordon JL, Lefeuvre P, Escalon A, Barbe V, Cruveiller S, *et al*. Comparative genomics of 43 strains of *Xanthomonas citri* pv. citri reveals the evolutionary events giving rise to pathotypes with different host ranges. *BMC Genomics* 2015;16:1098.
73. Wang M, Roux F, Bartoli C, Huard-Chauveau C, Meyer C, *et al*. Two-way mixed-effects methods for joint association analysis using both host and pathogen genomes. *Proc Natl Acad Sci U S A* 2018;115:E5440–E5449.
74. Teixeira M, Martins L, Fernandes C, Chaves C, Pinto J, *et al*. Complete genome sequences of walnut-associated *Xanthomonas euroxanthea* strains CPBF 367 and CPBF 426 obtained by Illumina/Nanopore hybrid assembly. *Microbiol Resour Announc* 2020;9:e00902-20.
75. Fernandes C, Blom J, Pothier JF, Tavares F. High-quality draft genome sequence of *Xanthomonas* sp. strain CPBF 424, a walnut-pathogenic strain with atypical features. *Microbiol Resour Announc* 2018;7:e00921-18.
76. Fernandes C, Martins L, Teixeira M, Blom J, Pothier JF, *et al*. Comparative genomics of *Xanthomonas euroxanthea* and *Xanthomonas arboricola* pv *juglandis* strains Isolated from a single walnut host tree. *Microorganisms* 2021;9.
77. Gétaz M, Krijger M, Rezzonico F, Smits THM, van der Wolf JM, *et al*. Genome-based population structure analysis of the strawberry plant pathogen *Xanthomonas fragariae* reveals two distinct groups that evolved independently before its species description. *Microb Genom* 2018;4:1–14.
78. Potnis N, Krasileva K, Chow V, Almeida NF, Patil PB, *et al*. Comparative genomics reveals diversity among *xanthomonads* infecting tomato and pepper. *BMC Genomics* 2011;12:146.
79. Timilsina S, Kara S, Jacques MA, Potnis N, MinsavageGV, *et al*. Reclassification of *Xanthomonas gardneri* (ex Šutič 1957) Jones et al. 2006 as a later heterotypic synonym of *Xanthomonas cynarae* trébaol et al. *Int J Syst Evol Microbiol* 2019;69 :343–349.
80. Michalopoulou VA, Vicente JG, Studholme DJ, Sarris PF. Draft genome sequences of pathotype strains for three pathovars belonging to three *Xanthomonas* species. *Microbiol Resour Announc* 2018;7:e00923-18.
81. Cohen SP, Luna EK, Lang JM, Ziegle J, Chang C, *et al*. High-quality genome resource of *Xanthomonas hyacinthi* generated via long-read sequencing. *Plant Dis* 2020;104:1011–1012.
82. Triplett LR, Verdier V, Campillo T, Van Malderghem C, Cleenwerck I, *et al*. Characterization of a novel clade of *Xanthomonas* isolated from rice leaves in mali and proposal of *Xanthomonas maliensis* sp. nov. *Antonie van Leeuwenhoek* 2015;107:869–881.
83. Ndongo S, Beye M, Dubourg G, Nguyen TTT, Couderc C, *et al*. Genome analysis and description of *Xanthomonas massiliensis* sp. nov., a new species isolated from human faeces.. *New Microbes New Infect* 2018;26:63–72.
84. Mücke S, Reschke M, Erkes A, Schwietzer C-A, Becker S, *et al*. Transcriptional reprogramming of rice cells by *Xanthomonas oryzae* TALEs. *Front Plant Sci* 2019;10:162.
85. Naushad S, Adeolu M, Wong S, Sohail M, Schellhorn HE, *et al*. A phylogenomic and molecular marker based taxonomic framework for the order *Xanthomonadales*: proposal to transfer the families *Algiphilaceae* and *Solimonadaceae* to the order Nevskiales ord. nov. and to create a new family within the order *Xanthomonadales*, the family *Rhodanobacteraceae* fam. nov., containing the genus *Rhodanobacter* and its closest relatives. *Antonie van Leeuwenhoek* 2015;107:467–485.
86. López MM, Lopez-Soriano P, Garita-Cambronero J, Beltrán C, Taghouti G, *et al*. *Xanthomonas prunicola* sp. nov., a novel pathogen that affects nectarine (Prunus persica var. nectarina) trees. *Int J Syst Evol Microbiol* 2018;68:1857–1866.
87. Hansen LBS, Ren D, Burmølle M, Sørensen SJ. Distinct gene expression profile of *Xanthomonas retroflexus* engaged in synergistic multispecies biofilm formation. *ISME J* 2017;11:300–303.
88. Gilroy R, Ravi A, Getino M, Pursley I, Horton DL, *et al*. Extensive microbial diversity within the chicken gut microbiome revealed by metagenomics and culture. *PeerJ* 2021;9:e10941.
89. Koebnik R, Burokiene D, Bragard C, Chang C, Saux M-L, *et al*. The complete genome sequence of *Xanthomonas theicola*, the causal agent of canker on tea plants, reveals novel secretion systems in clade-1 *Xanthomonads*. *Phytopathology* 2021;111:611–616.
90. Jaenicke S, Bunk B, Wibberg D, Spröer C, Hersemann L, *et al*. Complete genome sequence of the barley pathogen *Xanthomonas translucens* Pv. translucens DSM 579 18974T (ATCC 19319T). *Genome Announc* 2016;4:1–2.
91. Richard D, Boyer C, Vernière C, Canteros BI, Lefeuvre P, *et al*. 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 Announc* 2017;5:e00010-17.
92. Aritua V, Harrison J, Sapp M, Buruchara R, Smith J, *et al*. Genome sequencing reveals a new lineage associated with lablab bean and genetic exchange between *Xanthomonas axonopodis* pv. phaseoli and *Xanthomonas fuscans* subsp. fuscans. *Front Microbiol* 2015;6:1080.

Peer review history

VERSION 2

Editor recommendation and comments

https://doi.org/10.1099/acmi.0.000532.v2.3

Duncan Y. K. Ng; Quadram Institute Bioscience, UNITED KINGDOM

Date report received: 25 June 2023
Recommendation: Accept

Comments: I am pleased to tell you that your article has now been accepted for publication in Access Microbiology.

SciScore report

https://doi.org/10.1099/acmi.0.000532.v2.1

iThenticate report

https://doi.org/10.1099/acmi.0.000532.v2.2

Author response to reviewers to Version 1

Your reference: ACMI-D-22-00186

Dear Dr Ng,

the authors sincerely thank the editorial team and the anonymous reviewers for the constructive feedback and excellent suggestions for improvement. We have addressed all the points raised and I present a point-by-point response below.

Reviewer 1

(i) L25 and elsewhere: There seems to be some confusion in the literature over the proper spelling of the pathovar name "phormicola". In the original description by Takimoto (1933), the name "Bacterium phormicola" was used, and the spelling "phormicola" was later also used for Xanthomonas. However, in its first Comprehensive List of Names of Plant Pathogenic Bacteria, the Committee corrected several pathovar names according to Latinization rules, and added a connecting vowel "i" in compound names. Since then, the spelling used in publications has been "phormiicola". We therefore suggest to use the correct name following Latin rules, i.e. "phormiicola".

I am very grateful to the reviewer for their help in resolving this confusion and we now follow their suggestions to consistently use the spelling "*phormiicola*".

(ii) Line (L) 25/26: Is X. dyei pv. eucalypti a type or a pathotype strain?

I apologise for misleading information about this taxon in the Abstract. We now make it clear that we sequenced the pathotype strain.

(iii) L61/62: When looking at GenBank, it is not really true that there are no genome sequences available for Slc 2, 3, 4, 5 and 6. It's probably more precise to say that they have not been formally described yet.

Yes, the reviewer is correct. There are several genome sequences available for Slc2, which is now known as *X. dyei*and for Slc3, which is now known as *X. melonis*. Furthermore, we previously overlooked GenBank accession GCA_012848175.1, the genome

assembly for *Xanthomonas campestris*pv. *badrii*NEB122 [PT]; this strain represents Slc4. To the best of my knowledge, there are still no genome sequences available for Slc5. Shortly after we submitted the original manuscript, we discovered that Peduzzi and colleagues have recently sequenced CFBP 8444, which belongs to Slc6. So, we deleted this misleading sentence, and reorganised this section of the text for improved clarity. We also added these previously omitted genome assemblies in the revised version of Figure 1 and added short comments to this effect in the text.

(iv) L93: It was not Jacobs and colleagues who for the first time proposed "Xanthomonas cannabis" as a new species name. In fact, Netsu et al. (J. Gen. Plant Pathol. 80: 164-168) speculated in 2014 that their Japanese strain of Xanthomonas campestris pv. cannabis caused the same disease as observed earlier by Watanabe ("Sen-i sakumotsu byo gaku". Asakura Publishing, Tokyo, pp 112-113). In 1947, Watanabe had named the causal bacterium as "Pseudomonas cannabis", and subsequently, Okabe and Goto proposed in 1955 to reclassify it into the genus Xanthomonas as "X. cannabis" ("Bacterial plant diseases in Japan: I. A list of bacterial diseases and their pathogens". Rep. Fac. Agric. Shizuoka Univ. 5: 63-71). Based on this historical account, Jacobs and collaborators used the name suggested by Okabe and Goto, "X. cannabis".

Thank you for alerting us to this gap in my knowledge of the literature. We have now updated the text to reflect this narrative of the history of the taxonomic name "*X. cannabis*" and cited those references.

(v) L98: esculenti in italics

This is now resolved.

(vi) L101: NCPPB 2877 (with space in between)

This is now resolved.

(vii) L164: When ANI values are not conclusive, i.e. between 94% and 96%, it is advised to complement the analysis by calculating digital DNA-DNA hybridization (dDDH) values (e.g.https://www.dsmz.de/services/online-tools/genome-to-genome-distance-calculator-ggdc). Please add and discuss this information here.

Yes, at line 164 we observed that the ANI between *asclepiadis*NCPPB 4013 and *X. euroxanthea*is 95.97%, which is rather borderline in respect of species delineation. The Type (Strain) Genome Server, TYGS, implements includes the tool recommended by the reviewer. TYGS does not consider that NCPPB 4013 belongs to *X. euroxanthea*and suggests that it is a potential new species:

This is based on the following dDDH values:

(viii) L177/178: Please provide ANI and dDDH values when suggesting to incorporate "X. sontii" in the species X. sacchari.

On reflection, the case for incorporating *X. sontii*into *X. sacchari*is very weak and so we have deleted this suggestion.

(ix) L200: What is meant with "species X. hortorum lineage"?

This phrase was poorly worded. We changed it to: "clade that corresponds to the species *X. hortorum*".

(x) References 1, 5, 7, 11, 20, 23, 33, 34, 53, 55, 61, 62, 64 and 65 are not correct.

We have now fixed multiple formatting problems with the references.

(xi) For ref. 45, one may better cite: Daub ME, Ehrenshaft M (2000). The Photoactivated Cercospora Toxin Cercosporin: Contributions to Plant Disease and Fundamental Biology. Annu. Rev. Phytopathol. 38: 461-490. doi: 10.1146/annurev.phyto.38.1.461.

Thank you for the suggestion. We substituted this reference.

Reviewer 2

I do not see that it is necessary to include species level clades as headings as they are only informative relating to the article by Parkinson et al, 2009 and in my opinion are somewhat confusing.

Yes, I can appreciate that the emphasis on slc is potentially confusing. As the reviewer suggests, we have removed these from most of the headings (except where the slc does not correspond to a named species and is therefore a useful label) and have made some small changes to the text to put slightly less emphasis on these.

I believe the manuscript would benefit from including average nucleotide identities throughout the manuscript and in table format.

We have now generated a table of ANIs for all genome sequence assemblies mentioned in this manuscript. This is a very large table and so we include it as supplementary Excel spreadsheet file, in which the ANIs are coloured on a white-to-red scale.

The authors point to >96% ANI but >95% is ofter referred to as the cut-off for species.

We updated the relevant text to reflect this as follows: "The ANI between NCPPB 4013 and the *X. euroxanthea*type strain is 95.97 %, close to the threshold of 95 – 96 % that is commonly used to delineate species boundaries (30, 59–64)."

X cannabis is not a species and apparently the article by Jacobs et al. is not sufficient for species status.

It is indeed true that "*X. cannabis*" is not validly published according to List of Prokaryotic names with Standing in Nomenclature https://lpsn.dsmz.de/genus/xanthomonas. We state this in the manuscript. Also, this is one reason why we have to mention Slc 1, since this informal grouping corresponds closely to the entity described (invalidly) in the literature as "*X. cannabis*". Biologically speaking, it probably is a species.

Reviewer 3

DNA gyrase gene sequence is available for all the strains sequenced in the study. The authors need to make sure that gyrase gene sequence is identical to the corresponding sequence in the genome of that particular strai

Yes, indeed we checked this. Partial *gyrB*sequences are available in the databases for 9 out of the 10 sequenced strains. We have prepared a PowerPoint file containing screenshots of BLAST searches that conform 100% identity over the full lengths of the partial *gyrB*sequences for all 9 sequenced genomes. "We now state in the manuscript: For nine of the ten newly sequenced genomes, there are available partial *gyrB*sequences (GenBank accessions: EU007531.1, EU285211.1, EU285210.1, EU285202.1, EU285197.1, EU285181.1, EU285180.1, EU285177.1 and EU285064.1). In all cases, BLASTN searches confirmed 100% nucleotide sequence identity over the full length of the partial *gyrB*gene versus the corresponding genome sequence assembly."

Line 49- 34 valid species available.

According to List of Prokaryotic names with Standing in Nomenclature (https://lpsn.dsmz.de/genus/xanthomonas) there are 35 species with a validly published and correct name (excluding synonyms). We have updated our text with this corrected number.

Figure 1: Species names in the phylogenetic tree are not in italics

We have now italicised the species names in the Figure. For completeness and transparency, we also added accession numbers and citations for the genome assemblies in Figure 1; these are found in the new Table 2.

Table 1: Adding genome assembly statistics such as genome coverage and contamination, N50 value will make the genome assembly more reliable.

We have now compiled a table with the full results from CheckM analysis of the genome assemblies in a supplementary Excel file. We also added coverage and contamination to Table 1 and mentioned the high level of contamination in one assembly in the main text.

Line 108, Line 124, Line 151: Provide ANI data and dDDH values to strengthen the statement.

Our revised text now reads: "… analysis of NCPPB 2983 confirmed this and placed it as a distinct species close to *X. hyacinthi*and *X. translucens*(55) (Figure 1). It shares 95 % ANI with *X. bonasiae*FX4 (Supplementary Table S2), which is on the borderline of the commonly used criterion for species delimitation. Based on dDDH values, the Type Strain Genome Server identifies that this genome does not fall within any named species and potentially represents a new species."

Provide ANI values in a table or heatmap format to make the paper more comprehensible to the reader.

We have now generated a table of ANIs for all genome sequence assemblies mentioned in this manuscript. This is a large table and so we include it as supplementary Excel spreadsheet file, in which the ANIs are coloured on a white-to-red scale.

Further, calculation of digital DNA-DNA hybridization values might help in the taxonomic refinement of these strains along with phylogeny and ANI.

Where useful to do so, we added results of dDDH as calculated by the Type Strain Genome Server and also mentioned this in the Methods section.

Line 162: There are more genomes available for X. euroxanthea in the NCBI GenBank database. Adding these genomes to the phylogeny might clear whether NCPPB 4013 falls amongst the X. euroxanthea clade or in close association with them.

We have now included all available *X. euroxanthea*genome assemblies in Figure 1 and in the ANI calculations in the supplementary Excel file.

Line 177: The authors need to provide ANI and dDDH values of X. campestris pv. cannae with X. sontii and X.sacchari to support their statement that "There is also a case to incorporate "X. sontii" into this species, given its close

phylogenetic affinity"

On reflection, the case for incorporating *X. sontii*into *X. sacchari*is very weak and so we have deleted this suggestion.

I hope this newly improved manuscript is now acceptable for publication.

Yours sincerely,

David J Studholme

VERSION 1

Editor recommendation and comments

https://doi.org/10.1099/acmi.0.000532.v1.6

Duncan Y. K. Ng; Quadram Institute Bioscience, UNITED KINGDOM

Date report received: 07 March 2023
Recommendation: Minor Amendment

Comments: Overall, the study was well received. From my understanding of the reviewer's comments, most highlighted a need to calculating dDDH and visualise ANI results in a heatmap. The others should be relatively minor to address (ie. spelling/phrasing/formatting). This is a study that would be of interest to the field and community. The reviewers have highlighted minor concerns with the work presented. Please ensure that you address their comments.

Reviewer 3 recommendation and comments

https://doi.org/10.1099/acmi.0.000532.v1.4

Anonymous.

Date report received: 01 February 2023
Recommendation: Major Revision

Comments: 1. Methodological rigour, reproducibility and availability of underlying data - 2. Presentation of results 3. How the style and organization of the paper communicates and represents key findings 4. Literature analysis or discussion 5. Any other relevant comments The study/resource is important and timely for the community working in the area of plant pathogens in general and Xanthomonas in particular. Below are the major comments that authors need to address to strengthen the manuscript DNA gyrase gene sequence is available for all the strains sequenced in the study. The authors need to make sure that gyrase gene sequence is identical to the corresponding sequence in the genome of that particular strai Line 49- 34 valid species available. Figure 1: Species names in the phylogenetic tree are not in italics Table 1: Adding genome assembly statistics such as genome coverage and contamination, N50 value will make the genome assembly more reliable. Line 108, Line 124, Line 151: Provide ANI data and dDDH values to strengthen the statement. Provide ANI values in a table or heatmap format to make the paper more comprehensible to the reader. Further, calculation of digital DNA-DNA hybridization values might help in the taxonomic refinement of these strains along with phylogeny and ANI. Line 162: There are more genomes available for X. euroxanthea in the NCBI GenBank database. Adding these genomes to the phylogeny might clear whether NCPPB 4013 falls amongst the X. euroxanthea clade or in close association with them. Line 177: The authors need to provide ANI and dDDH values of X. campestris pv. cannae with X. sontii and X.sacchari to support their statement that "There is also a case to incorporate "X. sontii" into this species, given its close phylogenetic affinity"

Please rate the manuscript for methodological rigour
Good

Please rate the quality of the presentation and structure of the manuscript
Good

To what extent are the conclusions supported by the data?
Partially support

Do you have any concerns of possible image manipulation, plagiarism or any other unethical practices?
No

Is there a potential financial or other conflict of interest between yourself and the author(s)?
No

If this manuscript involves human and/or animal work, have the subjects been treated in an ethical manner and the authors complied with the appropriate guidelines?
Yes

Reviewer 2 recommendation and comments

https://doi.org/10.1099/acmi.0.000532.v1.3

Anonymous.

Date report received: 29 December 2022
Recommendation: Major Revision

Comments: The manuscript provides very useful information on genome sequences for unique strains that represent strains that are not widely distributed and represent underrepresented strains within several species. I do not see that it is necessary to include species level clades as headings as they are only informative relating to the article by Parkinson et al, 2009 and in my opinion are somewhat confusing. I believe the manuscript would benefit from including average nucleotide identities throughout the manuscript and in table format. The authors point to >96% ANI but >95% is ofter referred to as the cut-off for species. X cannabis is not a species and apparently the article by Jacobs et al. is not sufficient for species status. I have made many comments directly on the pdf. 1. Methodological rigour, reproducibility and availability of underlying data - Fine 2. Presentation of results - discussed above 3. How the style and organization of the paper communicates and represents key findings - Could be improved 4. Literature analysis or discussion - Brief but fine 5. Any other relevant comments. No

Please rate the manuscript for methodological rigour
Satisfactory

Please rate the quality of the presentation and structure of the manuscript
Satisfactory

To what extent are the conclusions supported by the data?
Strongly support

Do you have any concerns of possible image manipulation, plagiarism or any other unethical practices?
No

Is there a potential financial or other conflict of interest between yourself and the author(s)?
No

If this manuscript involves human and/or animal work, have the subjects been treated in an ethical manner and the authors complied with the appropriate guidelines?
No: not applicable

Reviewer 1 recommendation and comments

https://doi.org/10.1099/acmi.0.000532.v1.5

Anonymous.

Date report received: 20 December 2022
Recommendation: Minor Amendment

Comments: Jamie Harrison and co-workers present ten novel genome sequences of underexplored taxa in the genus Xanthomonas, which will be useful for follow-up studies using phylogenomic approaches. The short descriptions of the sequenced representatives of them are interesting to read. Methods are well described at protocols.io. There are a few points that the authors should consider when revising their manuscript: (i) L25 and elsewhere: There seems to be some confusion in the literature over the proper spelling of the pathovar name "phormicola". In the original description by Takimoto (1933), the name "Bacterium phormicola" was used, and the spelling "phormicola" was later also used for Xanthomonas. However, in its first Comprehensive List of Names of Plant Pathogenic Bacteria, the Committee corrected several pathovar names according to Latinization rules, and added a connecting vowel "i" in compound names. Since then, the spelling used in publications has been "phormiicola". We therefore suggest to use the correct name following Latin rules, i.e. "phormiicola". (ii) Line (L) 25/26: Is X. dyei pv. eucalypti a type or a pathotype strain? (iii) L61/62: When looking at GenBank, it is not really true that there are no genome sequences available for Slc 2, 3, 4, 5 and 6. It's probably more precise to say that they have not been formally described yet. (iv) L93: It was not Jacobs and colleagues who for the first time proposed "Xanthomonas cannabis" as a new species name. In fact, Netsu et al. (J. Gen. Plant Pathol. 80: 164-168) speculated in 2014 that their Japanese strain of Xanthomonas campestris pv. cannabis caused the same disease as observed earlier by Watanabe ("Sen-i sakumotsu byo gaku". Asakura Publishing, Tokyo, pp 112-113). In 1947, Watanabe had named the causal bacterium as "Pseudomonas cannabis", and subsequently, Okabe and Goto proposed in 1955 to reclassify it into the genus Xanthomonas as "X. cannabis" ("Bacterial plant diseases in Japan: I. A list of bacterial diseases and their pathogens". Rep. Fac. Agric. Shizuoka Univ. 5: 63-71). Based on this historical account, Jacobs and collaborators used the name suggested by Okabe and Goto, "X. cannabis". (v) L98: esculenti in italics (vi) L101: NCPPB 2877 (with space in between) (vii) L164: When ANI values are not conclusive, i.e. between 94% and 96%, it is advised to complement the analysis by calculating digital DNA-DNA hybridization (dDDH) values (e.g. https://www.dsmz.de/services/online-tools/genome-to-genome-distance-calculator-ggdc). Please add and discuss this information here. (viii) L177/178: Please provide ANI and dDDH values when suggesting to incorporate "X. sontii" in the species X. sacchari. (ix) L200: What is meant with "species X. hortorum lineage"? (x) References 1, 5, 7, 11, 20, 23, 33, 34, 53, 55, 61, 62, 64 and 65 are not correct. (xi) For ref. 45, one may better cite: Daub ME, Ehrenshaft M (2000). The Photoactivated Cercospora Toxin Cercosporin: Contributions to Plant Disease and Fundamental Biology. Annu. Rev. Phytopathol. 38: 461-490. doi: 10.1146/annurev.phyto.38.1.461.

Please rate the manuscript for methodological rigour
Very good

Please rate the quality of the presentation and structure of the manuscript
Very good

To what extent are the conclusions supported by the data?
Strongly support

Do you have any concerns of possible image manipulation, plagiarism or any other unethical practices?
No

Is there a potential financial or other conflict of interest between yourself and the author(s)?
No

If this manuscript involves human and/or animal work, have the subjects been treated in an ethical manner and the authors complied with the appropriate guidelines?
Yes

SciScore report

https://doi.org/10.1099/acmi.0.000532.v1.1

iThenticate report

https://doi.org/10.1099/acmi.0.000532.v1.2