INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY

RESEARCH ARTICLE

McKnight *et al.*, *Int. J. Syst. Evol. Microbiol.* 2024;74:006294

DOI 10.1099/ijsem.0.006294

MICROBIOLOGY SOCIETY

OPEN ACCESS

Xanthomonas rydalmerensis sp. nov., a non-pathogenic member of Group 1 *Xanthomonas*

Daniel J. E. McKnight[1,2,*], Johanna Wong-Bajracharya[1], Efenaide B. Okoh[1,3], Fridtjof Snijders[1], Fiona Lidbetter[1], John Webster[1], Mathew Haughton[1], Aaron E. Darling[2], Steven P. Djordjevic[2], Daniel R. Bogema[1,2] and Toni A. Chapman[1,2,*]

Abstract

Five bacterial isolates were isolated from *Fragaria* × *ananassa* in 1976 in Rydalmere, Australia, during routine biosecurity surveillance. Initially, the results of biochemical characterisation indicated that these isolates represented members of the genus *Xanthomonas*. To determine their species, further analysis was conducted using both phenotypic and genotypic approaches. Phenotypic analysis involved using MALDI-TOF MS and BIOLOG GEN III microplates, which confirmed that the isolates represented members of the genus *Xanthomonas* but did not allow them to be classified with respect to species. Genome relatedness indices and the results of extensive phylogenetic analysis confirmed that the isolates were members of the genus *Xanthomonas* and represented a novel species. On the basis the minimal presence of virulence-associated factors typically found in genomes of members of the genus *Xanthomonas*, we suggest that these isolates are non-pathogenic. This conclusion was supported by the results of a pathogenicity assay. On the basis of these findings, we propose the name *Xanthomonas rydalmerensis*, with DAR 34855^T = ICMP 24941 as the type strain.

INTRODUCTION

The members of the genus *Xanthomonas* are Gram-negative proteobacteria that cause disease in over 400 plant species. Their hosts include 268 dicotyledonous and 124 monocotyledonous plants and they cause significant agricultural impact [1–3]. Members of the genus *Xanthomonas* infect a wide variety of economically significant crops, including citrus, cabbage, tomato, capsicum, bean, rice, wheat and barley [1, 2, 4]. Pathogenic species of the genus *Xanthomonas* often consist of a suite of pathovars adapted to infect one host species or closely related species within the same genus, with many exhibiting tissue specificities for the vascular system or mesophyll [5]. Symptoms of infections include vascular wilts, blights, cankers, black chaff, black rot and spots on stems, leaves and fruit [6]. Members of the genus *Xanthomonas* are responsible for significant economic losses worldwide, as they lower the quality of produce and can cause crop loss [7].

Species of the genus *Xanthomonas* use a range of strategies to cause disease in plants. These involve the deployment of virulence-associated factors, including secretion systems, flagella, lipopolysaccharides (LPS) and extracellular polysaccharides (EPS). Secretion system types I to VI have been identified in species of the genus *Xanthomonas*, with the type III secretion system (T3SS) considered the primary virulence determinant. The T3SS is composed of a needle-like macromolecular protein that enables the translocation of exoenzymes, or type 3 (T3) effector proteins, directly into the cytoplasm of the host cell [8]. These effector proteins cause cytotoxicity, haemolysis, proteolysis and phosphorylation/dephosphorylation within the host [9, 10].

Author affiliations: [1]NSW Department of Primary Industries, Elizabeth Macarthur Agricultural Institute, Woodbridge Rd, Menangle NSW 2568, Australia; [2]University of Technology Sydney, 15 Broadway, Ultimo NSW 2007, Australia; [3]Western Sydney University, Penrith, NSW, Australia.
***Correspondence:** Daniel J. E. McKnight, daniel.j.mcknight@student.uts.edu.au; Toni A. Chapman, toni.chapman@dpi.nsw.gov.au
Keywords: bioinformatics; biosecurity; nonpathogenic; phylogenetics; *Xanthomonas*.
Abbreviations: ANI, average nucleotide identity; dDDH, digital DNA–DNA hybridisation; EPS, extracellular polysaccharides; SNP, single nucleotide polymorphism; T1SS, type I secretion system; T2SS, type II secretion system; T3SS, type III secretion system; T4SS, type IV secretion system.
The raw reads sequences for the isolates examined are available in the GenBank database with the following accession numbers: SRR25298660, SRR25298661, SRR25298662, SRR25298664 and SRR25298665 for the Oxford Nanpore sequences of DAR 34883, DAR 34882, DAR 34881, DAR 34857 and DAR 34855^T, respectively, SRR25298655, SRR25298656, SRR25298657, SRR25298658 and SRR25298659 for the Illumina sequences of DAR 34883, DAR 34882, DAR 34881, DAR 348757 and DAR 34855^T, respectively. Accession numbers for assembled genomes are available in the GenBank database with the following accession numbers: GCF_033170385.1 (biosample number SAMN35004907), GCF_033170365.1 (biosample number SAMN35004908), GCF_033170335.1, (biosample number SAMN35004910), GCF_033170315.1 (biosample number SAMN35004911) and GCF_033170295.1 (biosample number SAMN35004912 for DAR 34855^T, DAR 34857, DAR 34881, DAR 34882 and DAR 34883, respectively.
Five supplementary tables are available with the online version of this article.

Additionally, members of the genus *Xanthomonas* use a separate T3SS to secrete extracellular components of the flagella, that it uses for swimming motility. LPS play a critical role in cell wall integrity in Gram-negative bacteria and adherence to plant tissues [11, 12]. EPS, usually xanthan gum, contributes to the formation of biofilms that enhance adhesion to plant cell surfaces and improve stress tolerance [13].

The genus *Xanthomonas* predominantly consists of pathogenic species. Non-pathogenic species are often overlooked as they are difficult to detect in plants and lack immediate economic importance. The first report of non-pathogen strains of members of the genus *Xanthomonas* isolated from healthy plant tissue was in 1996, but they were only delineated to the genus level [14]. Over the past decade, several non-pathogenic species such as *X. bonasiae*, *X. maliensis*, '*X. massiliensis*', '*X. sontii*' and *X. youngii* have been identified [15–18]. Furthermore, non-pathogenic strains of known pathogenic species (i.e. *X. arboricola* and *X. euroxanthea*) have also been identified [19, 20]. The study of these non-pathogens increases our understanding of how pathogenic species cause disease, as it allows for the identification of specific genes, proteins or mechanisms that contribute to pathogenicity.

The identification of species within the genus *Xanthomonas* has a long history. In 1921, Ethel Mary Doidge identified a bacterial plant pathogen responsible for causing canker in tomato and capsicum plants [21]. In 1939, Walter John Dowson renamed multiple isolates, previously named *Bacterium vesicatorium* by Doidge, and proposed the genus *Xanthomonas* [22]. Isolates were originally classified as members of the genus *Xanthomonas* on the basis of phenotypic characteristics. The classification of species of the genus *Xanthomonas* has been a subject of debate throughout its history due to the limitations of phenetic characterisation methods. Subsequently, these methods have been proven to be inaccurate. Many isolates that were believed to represent the same species were delineated as technological advances were made, such as DNA–DNA hybridisation and protein electrophoresis. As a result, the taxonomy of the genus *Xanthomonas* has been systematically amended, leading to the reclassification of 20 nomenspecies based on DNA–DNA hybridisation by Vauterin *et al.* [23]. Genomic sequencing advances have resulted in substantial further reclassifications [24]. To designate a distinct novel species, it is necessary to employ diverse methodologies that yield consistent outcomes. Here, we provide carbon utilisation, MALDI-TOF MS, genome relatedness indices and results of phylogenetic analysis as evidence to identify and classify a novel species of the genus *Xanthomonas* collected from *Fragaria* × *ananassa* in Rydalmere, Australia.

ORIGIN AND ISOLATION

Bacterial samples were collected from *Fragaria* × *ananassa*, growing in Rydalmere, Australia, and submitted to the NSW Department of Agriculture diagnostic laboratories in 1976. They were originally identified as members of the genus *Xanthomonas* and stored in the NSW Plant Pathology and Mycology Herbarium as specimens DAR 34855^{T}, DAR 34857, DAR 34881, DAR 34882 and DAR 34883. The isolates were freeze-dried and stored under vacuum in glass ampoules at 4°C. These five isolates are referred to as the Rydalmere isolates for the remainder of this article.

GROWTH AND RECOVERY FROM CULTURE COLLECTION

Bacterial cultures were recovered onto yeast dextrose carbonate (YDC) solid agar and incubated at 25°C for 48 h. They are rod-shaped, aerobic and have been identified as Gram-stain-negative through the application of traditional Gram staining techniques and microscopic analysis. Under a stereomicroscope, they were found to form yellow, slightly convex, smooth, circular, mucoid colonies that were 2–3 mm in diameter on YDC media. Isolates were deposited in NZ – ICMP 24941, noting that the isolates were originally deposited in the NSW Plant Pathology and Mycology Herbarium, which is a publicly accessible collection.

DNA EXTRACTION, SEQUENCING AND GENOME ASSEMBLY

DNA extraction for the Rydalmere isolates was performed at the Elizabeth Macarthur Agricultural Institute. Library preparation and Illumina short read sequencing was performed according to the method of Bogema *et al.* [25] using a HiSeq 2500 system.

To extract DNA for nanopore sequencing, isolates on solid nutrient agar (NA) were retrieved from glycerol stocks and subsequently sub-cultured onto NA. Single colonies were collected and suspended in 1 ml phosphate buffered saline (PBS) (8.00g of NaCl, 1.15g of Na_2HPO_4, 0.20g of KH_2PO_4, and 0.20g of KCl per litre). Then samples were centrifuged at 10 000 ***g*** for 1 min and cell pellets were washed three times with 1 ml PBS. Genomic DNA was extracted with the Nanobind CBB Big DNA Kit (Circulomics) according to manufacturer's instructions. Genomic DNA concentration and extraction quality were determined using a Qubit 4 Fluorometer and Nanodrop 1000 (Thermo Fisher Scientific), respectively. Library preparation and Nanopore sequencing was provided by the Garvan Institute of Medical Research, Sydney, Australia. Sequencing libraries were generated using the high molecular weight DNA of each isolate with the Multiplex Ligation Sequencing kit (Nanopore, catalogue number SQK-MLK111-96-XL). Sequencing was conducted on a PromethION R9.4 flow cell (Nanopore, catalogue number FLO-PRO002).

Nanopore traces were basecalled with Guppy v6.3.7+532d626 [26] utilising the R9.4.1 super-high accuracy model. Raw reads were filtered to a minimum length of 1000 bp using Filtlong v0.2.1 [27] and draft assemblies were generated using Flye v2.9-b1768,

Miniasm v0.3-r179 with Minipolish v0.1.2, Raven v1.8.1 and Necat v0.0.1_update20200803 [28–32]. Draft assemblies were combined to produce consensus long-read assemblies using Trycycler v0.5.3 [33]. Long-read assemblies were polished using Nanopore reads with Medaka v1.7.1 [34] and with Illumina short reads using two sequential executions of Polypolish v0.5.0 and POLCA from MaSuRCA v4.0.9 [35, 36]. Whole genome sequences of DAR 34855^T, DAR 34857, DAR 34881, DAR 34882 and DAR 34883 were submitted to the NCBI database and their accession numbers can be found in Table S5, available in the online version of this article. Accession numbers for the Illumina short reads and Oxford Nanopore long reads of the Rydalmere isolates are listed in Table S1.

GENOME RELATEDNESS INDICES

First, comparative genomic analysis was employed to determine whether the Rydalmere isolates represent a novel species. Genome relatedness was estimated using average nucleotide identity (ANI) with FastANI v1.32 [37] and digital DNA–DNA hybridisation (dDDH) performed using formula 2 of Genome-to-Genome Distance Calculator (GGDC 3.0) [38]. Genomes with ANI values greater than the 95% threshold are considered to represent the same species [39]. Given that ANI values between the Rydalmere isolates were over 99.998%, we concluded that they are clonal and thus only one isolate (DAR 34855^T) was used for the analysis.

ANI was calculated to compare the Rydalmere isolates with all 2770 genomes labelled under the *Xanthomonas* genus in the National Centre for Biotechnology Information (NCBI) database as of the9th of January 2023. The Rydalmere isolates had ANI values less than the species delineation threshold when compared against all assemblies in the NCBI database, with the exception of three isolates: LMG9002 (GCF_009835085.1), 3307 (GCF_014199735.1) and 3498 (GCF_014199795.1). These three NCBI isolates have not previously been classified as representing a species of the genus *Xanthomonas* and do not share an ANI exceeding 95% with any isolates of known species of the genus *Xanthomonas*. Isolates 3307 and 3498 were collected from rainwater in the USA and LMG9002 was collected in 1989 from an orange tree in the USA and has been confirmed to be non-pathogenic [40, 41].

On the basis of ANI, the most closely related species to the Rydalmere isolates and 3307, 3498 and LMG9002 were '*X. sontii*', *X. sacchari* and *X. bonasiae* with ANI values of 93.8%, 93.6% and 88.8%, respectively. These findings are reflected by the dDDH results, which compared DAR 34855^T to all current type strains of members of the genus *Xanthomonas* that have a publicly available genome. Due to the clonal relationship of the Rydalmere isolates, the results obtained for DAR 34855^T are representative of the other isolates. All type strains had a dDDH value of less than 51%, which is lower than the species delineation threshold of 70% [42]. The most closely related type strains according to dDDH were '*X. sontii*', *X. sacchari* and *X. hyacinthi* with values of 50.8%, 48.8% and 33.3%, respectively. The results from ANI and dDDH indicate that the Rydalmere isolates described here and the three NCBI isolates do not represent a previously described species. Raw ANI and dDDH values can be found in Tables S2 and S3, respectively.

PHYLOGENETIC ANALYSIS

The 16S rRNA gene is a widely used molecular marker for the identification and classification of bacterial species [43]. This gene is present in all bacteria and has regions of both conserved and variable sequences that allow for differentiation between bacterial taxa. The 16S gene sequences were extracted from the whole genome sequences using Bakta v1.5.1 and Barrnap v0.9 [44, 45]. Alignment was performed using MUSCLE v5.1 [46] and a maximum likelihood tree with 100 bootstrap replicates was reconstructed using IQTREE v2.1.4 [47] with default settings and the three parameter model 2 with empirical base frequencies and a proportion of invariable sites (TPM2+F+I model). The phylogenetic tree was visualised and outgroup rooted using Interactive Tree of Life (ITOL) [48]. All assemblies analysed contained a 16S rRNA gene, except *X. arboricola* (probably a product of the short read sequencing and assembly method), which was excluded from the analysis.

The majority of the strains of members of the genus *Xanthomonas* in Fig. 1 are distributed across flat clades, making different species indiscernible. The 16S sequences of the Rydalmere isolates and closely related NCBI isolates are indistinguishable from that of *X. sontii*. This demonstrates that the 16S rRNA gene lacks discriminative resolution for species of the genus *Xanthomonas*, reflecting the results presented in previous studies [49]. As such, phylogenies generated from other genes or sets of genes are required for bacterial species classification.

Gyrase B is a subunit of DNA topoisomerase that plays a crucial role in regulating the supercoiling of DNA during replication and transcription [50]. The gene encoding the gyrase B subunit is widely used as a molecular marker for bacterial identification and phylogenetic analysis. This is due to it being highly conserved across bacterial genomes and having sufficient sequence variation to differentiate species of the genus *Xanthomonas* [51].

A gyrase B gene tree was reconstructed using sequences from the Rydalmere isolates, three closely related NCBI isolates, strains of members of the genus *Xanthomonas* and *Pseudoxanthomonas suwonensis* as an outgroup (Fig. 2). The gene sequences were extracted from the whole genome sequences using Bakta v1.5.1 and Genagr v0.1 [52]. All genomes contained a gyrase B gene, excluding '*Xanthomonas massiliensis*' (GCF_900018785.1). Gene sequences were aligned using MUSCLE v5.1 [46]

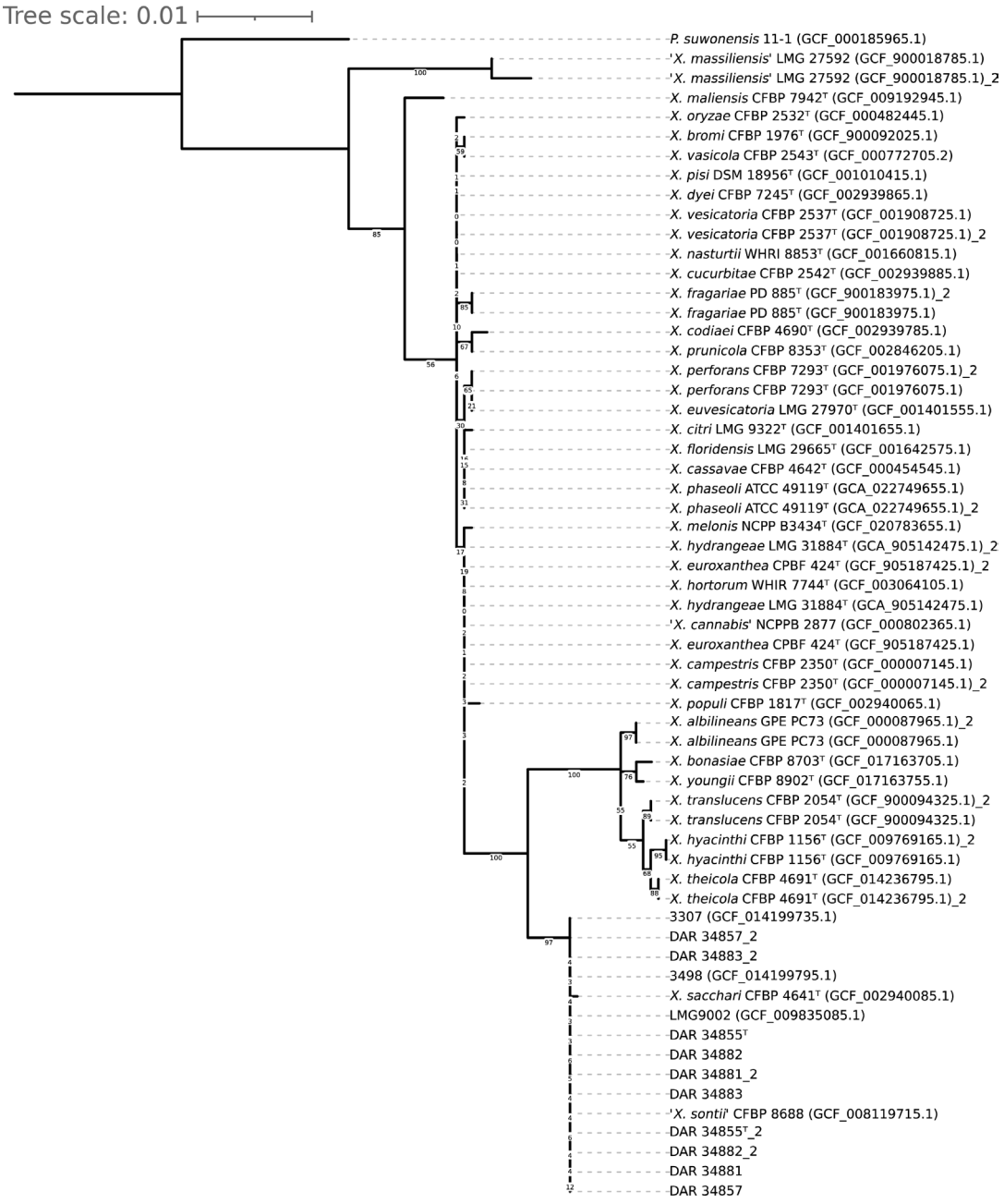

Fig. 1. Outgroup rooted maximum likelihood phylogeny of 16S rRNA gene sequences from Rydalmere isolates, three closely related NCBI isolates,strains of members of the genus *Xanthomonas* and *Pseudoxanthomonas suwonensis*. Strains obtained from NCBI are written as species name followed by strain and RefSeq database numbers (in parentheses). Genomes with duplicates of the 16S rRNA gene are denoted by '_2' for the second copy of the gene. The scale bar is in units of nucleotide substitutions per site.

with default settings and gapped positions were removed using trimaL v1.4.rev22 [53]. A maximum likelihood tree with 100 bootstrap replicates was generated using IQTREE v2.1.4 [47] with default settings and the general time reversible model with empirical base frequencies a proportion of invariant sites and and the discrete gamma model with four categories (GTR+F+I+G4 model). The tree was visualised and outgroup rooted using ITOL [48].

The gene tree formed two distinct clades, characteristic of the genus *Xanthomonas* [49, 54, 55], with the outgroup external to both groups. the Rydalmere and closely related NCBI isolates were positioned within group 1, defined Young *et al.* [54], in a monophyletic sub-clade. *X. sacchari*, '*X. sontii*' and *X. bonasiae* were the most closely related species, which reflects the findings for the genome relatedness indices.

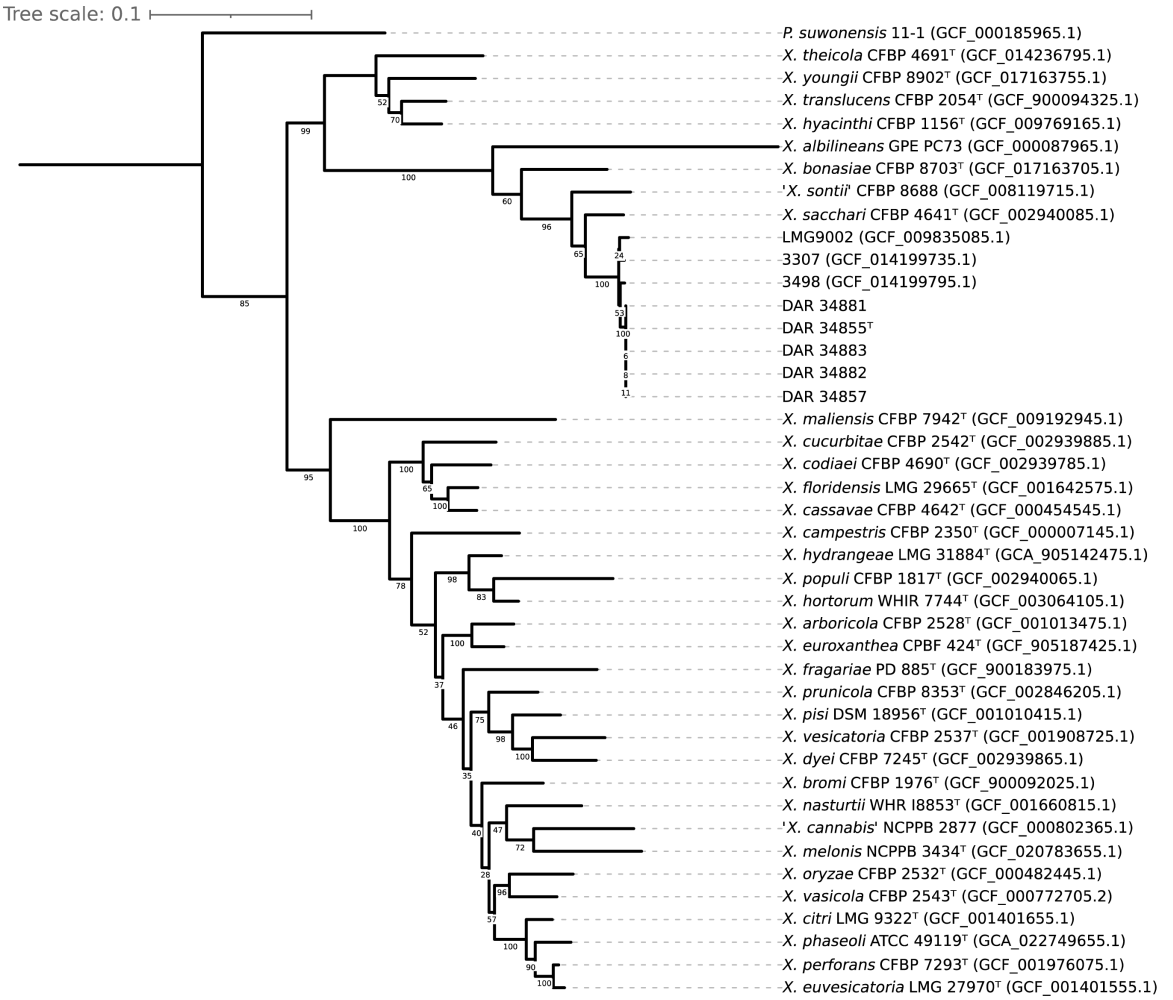

Fig. 2. Outgroup-rooted maximum likelihood tree using *gyrB* gene sequences of Rydalmere isolates, three closely related NCBI isolates, strains of members of the genus *Xanthomonas* and *P. suwonensis* as the outgroup. Strains obtained from NCBI are written as species name followed by strain designation and RefSeq database numbers (in parentheses). The two major clades that characterise the genus *Xanthomonas* genus are labelled according Young *et al.* [54]. The scale bar is in units of nucleotide substitutions per site.

Geneious Prime 2023.0.4 (https://www.geneious.com) was used to inspect the sequence alignment. There is a unique 18-nucleotide insertion between 403 and 420 bp in the *gyrB* alignment that was only detected in the Rydalmere isolates, the three related NCBI isolates and three closely related members of the genus *Xanthomonas* (*X. sacchari*, *X. bonasiae* and '*X. sontii*'). The insertion does not shift the reading frame.

To further investigate the evolutionary relationship and the accurate phylogenetic position of the Rydalmere and closely related NCBI isolates in the *Xanthomonas* genus, a 92-gene multilocus phylogeny was generated using UBCG v3 [56]. Analysis included the Rydalmere isolates, three closely related NCBI isolates, strains of members of the genus *Xanthomonas* and *Pseudoxanthomonas suwonensis* as the outgroup (Fig. 3). Concatenated core genes were aligned using MAFFT v7.310 [57] and a maximum likelihood tree was reconstructed using RaxML [58] with default settings and the general time reversible model with categories (GTRCAT model). The tree was visualised and outgroup rooted using ITOL [48]. The genomes all contained at least 89 of the 92 core genes, with the exceptions of *X. pisi* and *X. phaseoli*, which had 49 and 84, respectively. Accordingly, *X. pisi* was excluded from the analysis.

Analysis revealed that the members of the genus *Xanthomonas* grouped into two distinct clades, as reported previously [49, 54]. The most closely related species was shown to be *X. sacharri* and '*X. sontii*', which is consistent with the genomic relatedness indices and *gyrB* gene tree. The Rydalmere isolates and similar NCBI isolates formed a monophyletic sub-clade, distinct from those of closely related species. These results reflect the findings from ANI and dDDH analysis that indicate that they do not represent any known species.

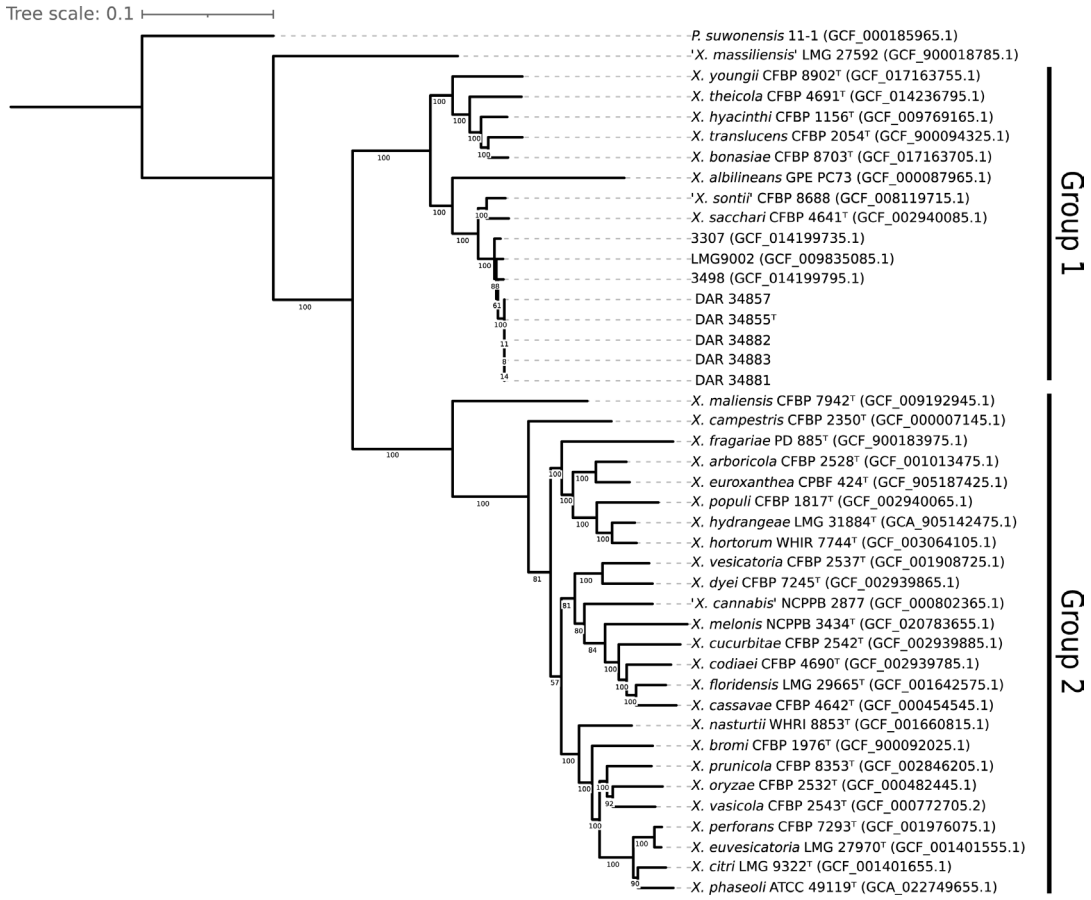

Fig. 3. Outgroup rooted maximum likelihood core gene phylogeny of Rydalmere isolates, three closely related NCBI isolates, strains of members of the genus *Xanthomonas* and *P. suwonensis*. Strains obtained from NCBI are written as species name followed by strain designation and RefSeq database numbers (in parentheses). The two major clades that characterise the genus *Xanthomonas* genus are labelled according Young *et al.* [54]. The scale bar is in units of nucleotide substitutions per site.

Single nucleotide polymorphisms (SNPs) from the core genomes of the Rydalmere isolates and three closely related NCBI isolates were identified and used to reconstruct an SNP-based phylogeny to investigate their intra-specific relationships (Fig. 4). SNPs were detected using Snippy v4.6.0 [59] with default settings and DAR 34855^{T} as the reference genome. The alignments were used to reconstruct a maximum likelihood tree with 100 bootstrap replicates using IQTREE v2.1.4 [47]. The phylogeny was minimum variance rooted using FastRoot v1.5 [60] and visualised using ITOL [48].

SNPs were only found on five nucleotides between the Rydalmere isolates, which supports the ANI results that indicated their clonality. In comparison, SNPs were detected at 103591 different nucleotide positions between the Rydalmere isolates and the

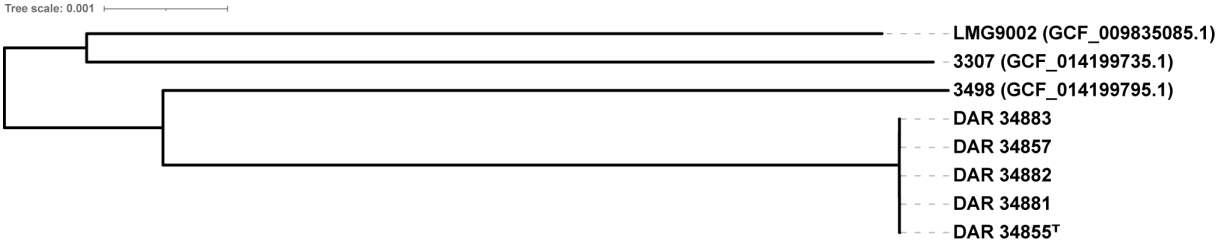

Fig. 4. Minimum variance rooted maximum likelihood core genome SNP-based phylogeny of the Rydalmere isolates and three closely related NCBI isolates. NCBI isolates written as strain name followed by RefSeq database number in parentheses. The scale bar is in units of nucleotide substitutions per site.

three closely related NCBI isolates. The Rydalmere isolates exhibit very few SNPs to differentiate them, but in comparison with the NCBI isolates, we found significant genetic variation between members of the species.

PHENOTYPIC CHARACTERISATION

Phenotypic characterisation of DAR 34883 and DAR 34855^T was performed using Biolog GEN III MicroPlates using protocol A. Selected isolates were sub-cultured from glycerol stock onto solid nutrient–yeast–glycerol agar (NYGA) and incubated at 27 °C for 48 h. Then, freshly grown colonies were suspended in inoculating fluid (IF) A using sterile cotton tipped swabs. Density of the inoculating fluid was measured and adjusted to achieve a transmission of 95–98%. A 100 µl aliquot of the bacterial suspension was transferred to each well of the MicroPlate and incubated at 30°C. Microplates were read using a MicroStation 2 Reader (Biolog) at 24 h, 48 h and 120 h. Three plates were used to characterise each of the two isolates with each well was measured three times to calculate the mean result. Readings were interpreted as per the manufacturer's instructions. Carbon utilisation wells were considered positive if they exceeded 160% of the value for the no-substrate control, negative if they were less than 130% and weak if they fell within the range in between [20, 61]. Chemical sensitivity assay wells were considered positive if the reading was greater than 50% of that of the positive control, as per the manufacturer's instructions. Additionally, if the results between triplicates were inconsistent, they were labelled 'v' to signify variability. Values measured at 24 h were the most consistent between triplicates and thus were used for analysis.

Across the three replicates, DAR 34855^T and DAR 34883 could utilise 39 substrates, weakly utilise 1 and could not utilise 17 (Table 1; full results Table S4). The remaining substrates gave inconclusive results as they were variable between isolates. Variation between the isolates was observed as DAR 34883 utilised L-rhamnose as a carbon source, whereas DAR 34855^T exhibited variable utilisation. Additionally, DAR 34883 displayed resistance to aztreonam, while DAR 34855^T was susceptible. Rydalmere isolates differed from closely related species by their ability to utilise L-galactonic acid lactone, 3-methyl glucose and D-malic acid. However, they were not unique in their susceptibility/tolerance to different growth environments.

PRESENCE AND ABSENCE OF VIRULENCE-ASSOCIATED FACTORS

The presence and absence of *Xanthomonas*-specific virulence-associated factors was examined using BLASTX in Diamond v2.0.15 [62] with no limit on maximum target sequences. A reference database was constructed using 163 previously characterised genes for T3SS, Type III effectors (T3E), flagellar T3SS, type IV secretion system (T4SS), type VI secretion system (T6SS), LPS and EPS from members of the genus *Xanthomonas* published by Gétaz *et al.* [63]. Additional proxy genes (*estA*, *fhaB*, *hlyB*, *pctB*, *raxB*, *xadA*, *xcs* and *xps*) were used to detect the presence of the type I secretion system (T1SS), type II secretion system (T2SS) and type V secretion system (T5SS) [64]. Analysis was performed on the Rydalmere isolates and closely related NCBI isolates (Table 2).

The Rydalmere isolates, being clonal, had identical virulence-associated factor profiles. The virulence gene profiles of the Rydalmere isolates and the three NCBI isolates were almost identical, with the only difference being the presence of *hlyB* in LMG9002. The presence of a range of genes that encode the biosynthesis of xanthan, a polysaccharide associated with stress tolerance, host defence inhibition, bacterial adhesion and biofilm formation, was detected [65, 66]. Two genes involved in lipopolysaccharide biosynthesis were detected. All isolates contained a wide variety of flagella T3SS genes, which export the flagella from the cytoplasm to the cell exterior, enabling motility. However, none of the isolates contained the genetic elements required for the synthesis of the virulent T3SS. One of the three proxy genes for the T1SS was detected, indicating that it lacks the ability to express the entire secretion system. Conversely, two proxy genes were used to detect the genes clusters for virulent T2SS (*xps*) and non-virulent T2SS (*xcs*). The presence of the *xps* gene cluster indicates that these isolates may have a functioning T2SS. Two of the eight T4SS subunits searched for were detected. As such, it is unlikely that these isolates can express a functional T4SS. No genes from the T3SS, T5SS or T6SS were detected in any of the assemblies.

MALDI-TOF MS ANALYSIS

A profile for DAR 34855^T was generated using a Bruker matrix-assisted laser desorption/ionisation-time of flight (MALDI-TOF) Biotyper Microflex LT (Bruker Daltonics). The isolate was inoculated onto an NA plate from glycerol stock and incubated for 24–48 h at 25–28°C. Then, a colony was transferred to an MBT Biotarget plate and an overlay of 1 µl of 70% formic acid was added and allowed to dry. Then, a second overlay of 1 µl of α-cyano-4-hydroxycinnamic acid was added and allowed to dry before analysis. Finally, the prepared target plate was inserted into the MALDI-TOF Biotyper and was run using settings recommended by Nellessen and Nehl [67] with a ratio of total-laser-shots to laser-shots-per-raster-spot of 5000:250. Analysis was conducted using the MBT Compass version 5.1.3 software and compared with isolates in our own database and the MBT Compass reference library. The resulting spectrum did not match that of any known species, with the highest score being another isolate of a member of the genus *Xanthomonas* in our database, with a log score of 1.83. Log scores between 1.7

Table 1. Phenotypic characterisation of DAR 34883 and DAR 34885 using BIOLOG GENIII microplates to determine carbon utilisation, antimicrobial resistance and growth at varying pH and salt concentrations. +, Positive, –, negative, v, variable. *X. hyacinthi, X. sacchari* and *X. theicola* data were obtained from BIOLOG GEN III database

Characterisation	DAR 34855^T	DAR 34883	*X. hyacinthi*	*X. sacchari*	*X. theicola*
Dextrin	+	+	+	+	+
Maltose	+	+	+	+	–
Trehalose	+	+	+	+	+
Cellobiose	+	+	–	+	–
Gentiobiose	+	+	–	+	+
Sucrose	+	+	+	+	–
Turanose	+	+	–	+	–
Raffinose	–	–	–	+	–
Lactose	+	+	–	+	–
Melibiose	+	+	–	+	–
Methyl β-D-glucoside	+	+	–	+	–
Gelatin	+	+	+	+	+
D-Fructose	+	+	+	+	+
D-Galactose	+	+	+	+	+
L-Fucose	+	+	–	+	+
L-Rhamnose	v	+	–	–	–
D-Sorbitol	–	–	–	–	–
D-Mannitol	–	–	–	–	–
D-Arabitol	–	–	–	–	–
L-Alanine	+	v	–	+	–
L-Aspartic acid	+	+	+	+	+
L-Glutamic acid	+	+	+	+	+
Formic acid	–	–	–	+	–
Quinic acid	+	+	–	+	+
Citric acid	+	+	+	+	+
Bromosuccinic acid	+	v	+	+	+
Propionic acid	+	+	–	+	+
Acetic acid	+	+	+	+	+
L-Galactonic acid lactone	+	+	–	–	–
3-Methyl glucose	+	+	–	–	–
D-Malic acid	+	+	–	–	–

and 2.0 are considered 'low confidence identification' and thus indicated that this species has not previously been identified and submitted to the Bruker database. The MALDI BioTyper Main Spectra (BTMSP) file for identification of DAR 34855^T is available on Figshare at 10.6084/m9.figshare.25182509.

PATHOGENICITY TESTS

Inoculation of the strawberry cv. elsanta plantlets with DAR 34855^T was adapted from the protocol by Kałużna et al. [68] in their second trial. Plantlets used in the study were derived from the stolons of 3 month-old strawberry plants. The youngest leaves of each plant were inoculated with 1×10^8 c.f.u. ml^{-1} of the bacteria in sterile water. This concentration was confirmed

Table 2. Presence and absence of virulence-associated factors in Rydalmere isolates compared with closely related NCBI isolates detected using BLASTX. The presence and absence of a virulence-associated factor is signified by a plus symbol or a minus symbol, respectively. The threshold requirements are a sequence identity similarity of 70% or greater and an *e*-value of $1e^{-10}$ or less

Gene category	Gene name	Rydalmere isolates	LMG9002	3307	3498
EPS	*gumD*	+	+	+	+
	gumK	+	+	+	+
	gumL	+	+	+	+
	gumM	+	+	+	+
	motA	+	+	+	+
	motB	+	+	+	+
LPS	*etfA*	+	+	+	+
	metB	+	+	+	+
T1SS	*hlyB*	–	+	–	–
T2SS	*xps*	+	+	+	+
Flagellar T3SS	*fleN*	+	+	+	+
	flgB	+	+	+	+
	flgC	+	+	+	+
	flgE	+	+	+	+
	flgF	+	+	+	+
	flgG	+	+	+	+
	flgH	+	+	+	+
	flgI	+	+	+	+
	flhA	+	+	+	+
	flhB	+	+	+	+
	fliA	+	+	+	+
	fliF	+	+	+	+
	fliG	+	+	+	+
	fliH	+	+	+	+
	fliI	+	+	+	+
	fliJ	+	+	+	+
	fliL	+	+	+	+
	fliM	+	+	+	+
	fliN	+	+	+	+
	fliP	+	+	+	+
	fliQ	+	+	+	+
	fliR	+	+	+	+
	fliS	+	+	+	+
T4SS	*virB4*	+	+	+	+
	virB11	+	+	+	+

with c.f.u. plating on NA, incubation and enumeration. The negative control for each group was inoculated using the same method with sterile water. Each group used a different method of inoculation, totalling three separate methods. Each group (Fig. 5) contained four replicate plants and each plant was inoculated three times on each of four different leaves. Negative control plants received similar treatments, but sterile distilled water was used instead of the bacterial suspension. The

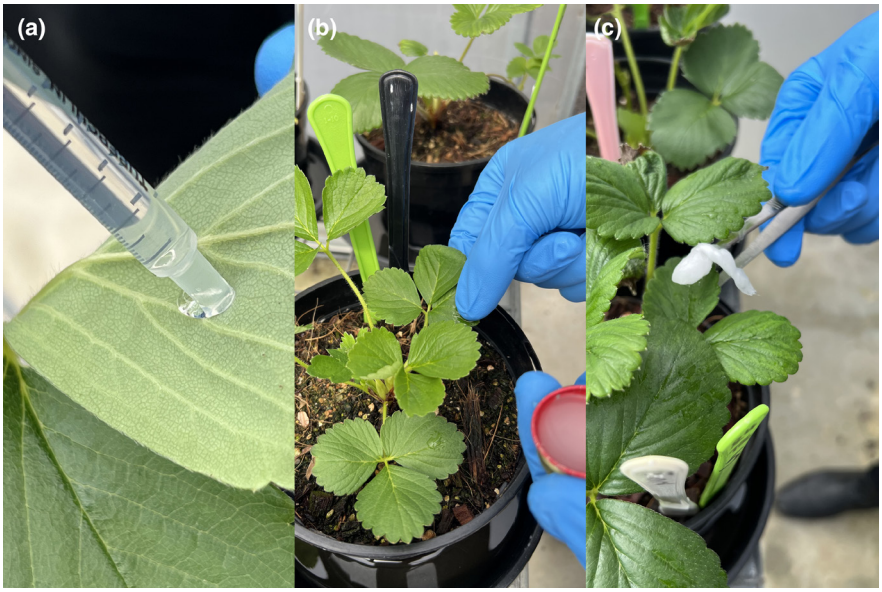

Fig. 5. Plant pathogenesis assay inoculation methods. (a) Inoculation of the bacterial suspension using a syringe on the abaxial and adaxial leaf surfaces. (b) Rubbing of leaf surface with a mixture carborundum powder and bacterial suspension. (c) Leaf infiltration using forceps and sterile cotton soaked in the bacterial suspension.

greenhouse was maintained at 25–30°C with a relative humidity level of 70–90% under natural light conditions and plants were systematically watered. The pathogenesis trial was run for 8 weeks and checked twice weekly for signs of infection, no hypersensitivity or disease symptoms were observed.

GENOME FEATURES

Each of the Rydalmere isolates were analysed using QUAST v5.1.0rc1 [69] and Bakta v1.5.1 [44]. DAR 34855^{T} was designated as the type strain due to its clonal relationship with the remaining four isolates, with the added distinction of being the first collected. The genome of DAR 34855^{T} is complete and comprised of a circular chromosome with no plasmids. It is 5.04 Mbp in length, similar to other species in the genus, with a total number of 4161 coding sequences (CDS). The DNA guanine–cytosine content (GC-content) is 68.7 mol%, which is within the typical range for species of the genus *Xanthomonas*.

The core genome of the Rydalmere isolates was defined using Roary v3.13.0 [70] with default parameters, using genome annotations predicted by Bakta v1.5.1 [44]. The Rydalmere isolates contain 4165 core genes with 11 shell genes and no cloud genes. When reanalysed with the inclusion of the three closely related NCBI isolates, there were 3524 core genes, 757 shell genes and 633 cloud genes. No isolates in the analysis were found to contain plasmids.

DESCRIPTION OF *XANTHOMONAS RYDALMERENSIS* SP. NOV

Xanthomonas rydalmerensis (ry.dal.mer.en'sis, O.L. fem. adj. of or belonging to Rydalmere, a suburb of Sydney, where it was first isolated)

Cells are Gram-stain-negative, aerobic and rod shaped. When grown on yeast dextrose carbonate (YDC) solid agar and incubated at 25°C for 48 h formed slightly convex, smooth, circular, mucoidal, yellow-pigmented colonies that were 2–3 mm in diameter. Able to utilise α-D-glucose, gelatin, pectin, Tween 40, dextrin, lactose, D-mannose, D-galacturonic acid, maltose, melibiose, D-fructose, L-galactonic acid lactone, trehalose, methyl β-D-glucoside, D-galactose, L-lactic acid, cellobiose, D-salicin, 3-methyl glucose, glycerol, L-aspartic acid, D-glucuronic acid, citric acid, gentiobiose, *N*-acetyl-D-glucosamine, D-fucose, L-glutamic acid, glucuronamide, α-ketoglutaric acid, acetoacetic acid, sucrose, L-fucose, D-fructose 6-phosphate, D-malic acid, propionic acid, turanose, quinic acid, L-malic acid, L-alanine, bromosuccinic acid, and acetic acid. It is not pathogenic to strawberry cv. elsanta, its host plant. Can be differentiated from other species of the genus *Xanthomonas* by dDDH and ANI calculations and with MALDI-TOF spectra.

This species includes the type strain DAR 34855^T = ICMP 24941 (NCBI accession number GCF_033170385.1) and strains DAR 34857 (GCF_033170365.1), DAR 34881 (GCF_033170335.1), DAR 34882 (GCF_033170315.1) and DAR 34883 (GCF_033170295.1). The genome size of the type strain is approximately 5.04 Mbp with a DNA G+C-content of 68.7 mol%. The four strains were collected from *Fragaria* × *ananassa* in 1976 in Rydalmere, Australia.

Funding information
This research was funded by the Australian Research Council Linkage project (LP180100593) and Rural Research and Development for Profit/Grains Research and Development Corporation (9177866).

Acknowledgement
We thank Nandita Pathania of the Department of Agriculture, Fisheries and Forestry (Queensland) for help with Biolog data acquisition and the NSW Plant Pathology and Mycology Herbarium for providing cultures. Also, we acknowledge AusGem DPI/UTS for producing high quality Illumina sequencing of the Rydalmere isolates. We acknowledge the financial support provided by the Australian Research Council Linkage project and Rural Research and Development for Profit/Grains Research and Development Corporation. We would like to thank the NSW DPI Advanced Gene Technology Centre for the space and resources for Nanopore sequencing. Finally, we thank Krista Plett and Bernie Dominiak for reviewing this manuscript.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References

1. Hayward AC. The hosts of *Xanthomonas*. In: Swings JG and Civerolo EL (eds). *Xanthomonas*. Dordrecht: Springer Netherlands; 1993. pp. 1–119.
2. Leyns F, De Cleene M, Swings J-G, De Ley J. The host range of the genus *Xanthomonas*. *Bot Rev* 1984;50:308–356.
3. Starr MP. The genus Xanthomonas. In: Starr MP, Stolp H, Trüper HG, Balows A and Schlegel HG (eds). *The Prokaryotes*. Berlin, Heidelberg: Springer; 1981. pp. 742–763.
4. Ryan RP, Vorhölter F-J, Potnis N, Jones JB, Van Sluys M-A, *et al*. Pathogenomics of *Xanthomonas*: understanding bacterium–plant interactions. *Nat Rev Microbiol* 2011;9:344–355.
5. Bogdanove AJ, Koebnik R, Lu H, Furutani A, Angiuoli SV, *et al*. Two new complete genome sequences offer insight into host and tissue specificity of plant pathogenic *Xanthomonas* spp. *J Bacteriol* 2011;193:5450–5464.
6. An S-Q, Potnis N, Dow M, Vorhölter F-J, He Y-Q, *et al*. Mechanistic insights into host adaptation, virulence and epidemiology of the phytopathogen *Xanthomonas*. *FEMS Microbiol Rev* 2020;44:1–32.
7. Denancé N, Lahaye T, Noël LD. Editorial: genomics and effectomics of the crop killer *Xanthomonas*. *Front Plant Sci* 2016;7:71.
8. Coburn B, Sekirov I, Finlay BB. Type III secretion systems and disease. *Clin Microbiol Rev* 2007;20:535–549.
9. Hueck CJ. Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol Mol Biol Rev* 1998;62:379–433.
10. White FF, Potnis N, Jones JB, Koebnik R. The type III effectors of *Xanthomonas*. *Mol Plant Pathol* 2009;10:749–766.
11. Davey ME, O'toole GA. Microbial biofilms: from ecology to molecular genetics. *Microbiol Mol Biol Rev* 2000;64:847–867.
12. Zhang G, Meredith TC, Kahne D. On the essentiality of lipopolysaccharide to Gram-negative bacteria. *Curr Opin Microbiol* 2013;16:779–785.
13. Sutherland IW. Biofilm exopolysaccharides: a strong and sticky framework. *Microbiology* 2001;147:3–9.
14. Vauterin L, Yang P, Alvarez A, Takikawa Y, Roth DA, *et al*. Identification of non-pathogenic *Xanthomonas* strains associated with plants. *Syst Appl Microbiol* 1996;19:96–105.
15. 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.
16. 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.
17. 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.
18. Ndongo S, Beye M, Dubourg G, Nguyen TT, 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.
19. Essakhi S, Cesbron S, Fischer-Le Saux M, Bonneau S, Jacques M-A, *et al*. Phylogenetic and variable-number tandem-repeat analyses identify nonpathogenic *Xanthomonas arboricola* lineages lacking the canonical type III secretion system. *Appl Environ Microbiol* 2015;81:5395–5410.
20. 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.
21. Doidge EM. A tomato canker. *An Appl Biol* 1921;7:407–430.
22. Dowson WJ. On the systematic position and generic names of the Gram-negative bacterial plant pathogens. *Zentralblatt fur Bakteriologie, Parasitenkunde, Infektionskrankheiten und Hygiene* 1939:177–193.
23. Vauterin L, Hoste B, Kersters K, Swings J. Reclassification of *Xanthomonas*. *Int J Syst Bacteriol* 1995;45:472–489.
24. Harrison J, Hussain RMF, Aspin A, Grant MR, Vicente JG, *et al*. Phylogenomic analysis supports the transfer of 20 pathovars from *Xanthomonas campestris* into *Xanthomonas euvesicatoria*. *Taxonomy* 2023;3:29–45.
25. Bogema DR, Micallef ML, Liu M, Padula MP, Djordjevic SP, *et al*. Analysis of *Theileria orientalis* draft genome sequences reveals potential species-level divergence of the Ikeda, Chitose and Buffeli genotypes. *BMC Genomics* 2018;19:298.
26. Wick RR, Judd LM, Holt KE. Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol* 2019;20:129.
27. Wick R, Menzel P. Filtlong; 2021. https://github.com/rrwick/Filtlong [accessed 26 April 2023].
28. Chen Y, Nie F, Xie S-Q, Zheng Y-F, Dai Q, *et al*. Efficient assembly of nanopore reads via highly accurate and intact error correction. *Nat Commun* 2021;12:60.
29. Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, *et al*. metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nat Methods* 2020;17:1103–1110.
30. Li H. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics* 2016;32:2103–2110.

31. Vaser R, Šikić M. Time- and memory-efficient genome assembly with Raven. *Nat Comput Sci* 2021;1:332–336.
32. Wick RR, Holt KE. Benchmarking of long-read assemblers for prokaryote whole genome sequencing. *F1000Res* 2019;8:2138.
33. Wick RR, Judd LM, Cerdeira LT, Hawkey J, Méric G, *et al*. Trycycler: consensus long-read assemblies for bacterial genomes. *Genome Biol* 2021;22:266.
34. Wright C, Wykes M. Medaka; 2020. https://github.com/nanoporetech/medak [accessed 2 February 2023].
35. Wick RR, Holt KE. Polypolish: short-read polishing of long-read bacterial genome assemblies. *PLoS Comput Biol* 2022;18:e1009802.
36. Zimin AV, Salzberg SL. The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies. *PLoS Comput Biol* 2020;16:e1007981.
37. 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.
38. Meier-Kolthoff JP, Carbasse JS, Peinado-Olarte RL, Göker M. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res* 2022;50:D801–D807.
39. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A* 2009;106:19126–19131.
40. Bansal K, Kumar S, Patil PB. Phylogenomic Insights into diversity and evolution of nonpathogenic *Xanthomonas* strains associated with citrus. *mSphere* 2020;5:e00087-20.
41. Pena MM, Bhandari R, Bowers RM, Weis K, Newberry E, *et al*. Genetic and functional diversity help explain pathogenic, weakly pathogenic, and commensal lifestyles in the genus *Xanthomonas*. *Ecology* 2023.
42. 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.
43. Janda JM, Abbott SL. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol* 2007;45:2761–2764.
44. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, *et al*. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. *Microb Genom* 2021;7:000685.
45. Seemann T. Barrnap 0.9: rapid ribosomal RNA prediction; 2018. https://github.com/tseemann/barrnap [accessed 20 February 2023].
46. Edgar RC. High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Bioinformatics* 2021.
47. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;32:268–274.
48. 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.
49. Hauben L, Vauterin L, Swings J, Moore ERB. Comparison of 16S ribosomal DNA sequences of all *Xanthomonas* species. *Int J Syst Bacteriol* 1997;47:328–335.
50. Reece RJ, Maxwell A. DNA gyrase: structure and function. *Crit Rev Biochem Mol Biol* 1991;26:335–375.
51. Parkinson N, Aritua V, Heeney J, Cowie C, Bew J, *et al*. Phylogenetic analysis of *Xanthomonas* species by comparison of partial gyrase B gene sequences. *Int J Syst Evol Microbiol* 2007;57:2881–2887.
52. Bogema DR. Genagr; 2022. https://github.com/bogemad/genagr [accessed 23 March 2023].
53. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 2009;25:1972–1973.
54. Young JM, Park D-C, Shearman HM, Fargier E. A multilocus sequence analysis of the genus *Xanthomonas*. *Syst Appl Microbiol* 2008;31:366–377.
55. 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.
56. Na S-I, Kim YO, Yoon S-H, Ha S-M, Baek I, *et al*. UBCG: up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. *J Microbiol* 2018;56:280–285.
57. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013;30:772–780.
58. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;30:1312–1313.
59. Seemann T. Snippy: fast bacterial variant calling from NGS reads; 2015. https://github.com/tseemann/snippy [accessed 4 June 2023].
60. Mai U, Sayyari E, Mirarab S. Minimum variance rooting of phylogenetic trees and implications for species tree reconstruction. *PLoS One* 2017;12:e0182238.
61. 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.
62. Buchfink B, Reuter K, Drost H-G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat Methods* 2021;18:366–368.
63. Gétaz M, Blom J, Smits THM, Pothier JF. Comparative genomics of *Xanthomonas fragariae* and *Xanthomonas arboricola* pv. fragariae reveals intra- and interspecies variations. *Phytopathol Res* 2020;2:17.
64. Alvarez-Martinez CE, Sgro GG, Araujo GG, Paiva MRN, Matsuyama BY, *et al*. Secrete or perish: the role of secretion systems in *Xanthomonas* biology. *Comput Struct Biotechnol J* 2021;19:279–302.
65. Bianco MI, Toum L, Yaryura PM, Mielnichuk N, Gudesblat GE, *et al*. Xanthan pyruvilation is essential for the virulence of *Xanthomonas campestris* pv. *campestris*. *Mol Plant Microbe Interact* 2016;29:688–699.
66. Yun MH, Torres PS, El Oirdi M, Rigano LA, Gonzalez-Lamothe R, *et al*. Xanthan induces plant susceptibility by suppressing callose deposition. *Plant Physiol* 2006;141:178–187.
67. Nellessen CM, Nehl DB. An easy adjustment of instrument settings ('Peak MALDI') improves identification of organisms by MALDI-ToF mass spectrometry. *Sci Rep* 2023;13:15018.
68. Kałużna M, Kuras A, Puławska J. mRNA extraction of *Xanthomonas fragariae* in strawberry and validation of reference genes for the RT-qPCR for study of bacterial gene expression. *Mol Biol Rep* 2019;46:5723–5733.
69. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 2013;29:1072–1075.
70. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, *et al*. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 2015;31:3691–3693.