INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY

TAXONOMIC DESCRIPTION

Nadal-Jimenez *et al.*, *Int. J. Syst. Evol. Microbiol.* 2022;72:005440

DOI 10.1099/ijsem.0.005440

MICROBIOLOGY SOCIETY

OPEN ACCESS

Symbiopectobacterium purcellii, gen. nov., sp. nov., isolated from the leafhopper *Empoasca decipiens*

Pol Nadal-Jimenez[1,*], Stefanos Siozios[1], Nigel Halliday[2], Miguel Cámara[2] and Gregory D. D. Hurst[1]

Abstract

Bacterial endosymbionts are found in multiple arthropod species, where they play crucial roles as nutritional symbionts, defensive symbionts or reproductive parasites. Recent work has highlighted a new clade of heritable microbes within the gammaproteobacteria that enter into both obligate and facultative symbioses, with an obligately required unculturable symbiont recently given the name *Candidatus* Symbiopectobacterium. In this study, we describe a culturable rod shaped non-flagellated bacterial symbiont from this clade isolated from the leafhopper *Empoasca decipiens*. The symbiont is related to the transovarially transmitted 'BEV' bacterium that was first isolated from the leafhopper *Euscelidius variegatus* by Alexander Purcell, and we therefore name the symbiont *Symbiopectobacterium purcellii* sp. nov., gen. nov. We further report the closed genome sequence for *S. purcellii*. The genome is atypical for a heritable microbe, being large in size, without profound AT bias and with little evidence of pseudogenization. The genome is predicted to encode Type II, III and VI secretion systems and associated effectors and a non-ribosomal peptide synthase array likely to produce bioactive small molecules. The predicted metabolism is more complete than for other symbionts in the *Symbiopectobacterium* clade, and the microbe is predicted to synthesize a range of B vitamins. However, Biolog plate results indicate that the metabolism is depauperate compared to the sister clade, represented by *Pectobacterium carotovorum*. A quorum-sensing pathway related to that of *Pectobacterium* species (containing an overlapping expI-expR1 pair in opposite directions and a "solo" expR2) is evidenced, and LC-MS/MS analysis reveals the presence of 3-hydroxy-C10-HSL as the sole *N*-acylhomoserine lactone (AHL) in our strain. This AHL profile is profoundly divergent from that of other *Erwinia* and *Pectobacterium* species which produce mostly 3-oxo-C6- and 3-oxo-C8-HSL and could aid group identification. Thus, this microbe denotes one that has lost certain pathways associated with a saprophytic lifestyle but represents an important baseline against which to compare other members of the genus *Symbiopectobacterium* that show more profound integration into host biology. The type strain of *Symbiopectobacterium purcellii* gen. nov., sp. nov. is SyEd1^T (LMG 32449^T=CECT 30436^T).

INTRODUCTION

It is now understood that microbes influence multiple aspects of animal biology [1]. Symbiont contributions extend from involvement in the process of digestion in the gut, through anabolic activities and the supply of vitamins and amino acids, to protection against natural enemies and defence against prey/hosts [2]. Conversely, other symbiotic microbes are pathogenic or parasitic, and many symbioses combine both parasitic and beneficial aspects. Levels of symbiont integration vary between symbioses [3]. On the host axis, they vary from facultative relationships where the host does not require a particular symbiont, to obligate where the individual dies or becomes sterile in the absence of symbiosis. Likewise, symbionts vary in the degree to which they rely on a host – some only replicating within hosts with others having environmental replication. The process of symbiosis formation also varies – from arising within the host lifecycle through acquisition by the host or infection by the microbe, to being present through the host lifecycle, with symbiont transfer/transmission from parent to offspring.

Author affiliations: [1]Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, UK; [2]The National Biofilms Innovation Centre, School of Life Sciences, University of Nottingham Biodiscovery Institute, University of Nottingham, Nottingham, UK.
***Correspondence:** Pol Nadal-Jimenez, polnadal@liverpool.ac.uk
Keywords: *Empoasca*; leafhopper; quorum sensing; symbiosis; *Symbiopectobacterium*.
Abbreviations: AHL, *N*-acylhomoserine lactone; ANI, average nucleotide indentity; BEV, bacterium from *Euscelidius variegatus*; BHI, brain heart infusion; CO-1, cytochrome oxidase 1; NRPS, non-ribosomal peptide synthase; PCWDE, plant cell wall-degrading enzyme; PGA, polygalacturonic acid.
The 16S rRNA sequence of the type strain is available at accession OK044380. The complete genome assembly and the raw reads have been submitted to the DDBJ/EMBL/GenBank database under the BioProject accession number PRJNA756769 (genome accession number CP081864).
One supplementary figure is available with the online version of this article.

Whilst arthropod–microbe symbioses are diverse in terms of the microbial partners, particular microbial taxa have established symbiosis with a number of host species, commonly establishing in new host species through a host switch event. Well-known 'heritable symbionts' found over a broad range of arthropods include *Wolbachia*, *Rickettsia*, *Spiroplasma*, *Cardinium* and *Arsenophonus* [4]. The interactions found in these symbioses include obligate and facultative associations, and ones which are beneficial, parasitic or have a combination of features.

Recent research has added a new clade of insect symbionts, *Candidatus* Symbiopectobacterium, to 'the big five' (the most abundant/diverse insect endosymbionts: *Wolbachia*, *Rickettsia*, *Cardinium*, *Arsenophonus* and *Sodalis*) [5]. The first member of this clade to be described was the BEV strain – an acronym for bacterium from *Euscelidius variegatus*, a planthopper host species. This strain was cultured [6], but never formally named. The symbiosis was characterized as one with vertical transmission, where the host's reproduction was negatively impacted by the microbe. In addition, there was also transmission to other insects on the plant – thus establishing the symbiosis as a pathogenic one maintained through mixed modes of transmission [7]. Experiments also suggested the symbiont facilitated the transmission of phytoplasma from its bug host to plant [8]. Later, the pest species *Cimex lectularius* (common bedbug) was observed to carry a heritable symbiont related to BEV [9]. This symbiont has not been established in cell-free culture, and symbiosis is facultative from the host perspective: the bedbug does not require the symbiont. Following this, a third hemipteran – the bulrush bug *Chilacis* was observed to carry a related vertically transmitted symbiont housed in a gut mycetome, in what appears to be an obligate association, in which the host requires the symbiont [10]. More recently, symbioses involving members of this clade have extended beyond Hemiptera hosts to nematodes, with Martinson *et al.* [5] characterizing symbionts related to BEV as obligate partners of *Howardula* nematodes. They named this microbe *Candidatus* Symbiopectobacterium, reflecting its symbiotic lifestyle and its sister relationship to the well-characterized genus *Pectobacterium*.

Candidatus Symbiopectobacterium has thus emerged as a potentially widespread and significant symbiotic associate of invertebrates. The original culturable BEV isolate, on which the genus could be formally described, was lost and genomic information for this strain is partial [11]. Recovering a model culturable member of the genus is important, as it allows formal description of the microbe, completion of a closed genome sequence against which reductive evolution in symbiosis can be measured and presents a system in which gene function may be investigated. In this paper, we report the isolation to pure culture of a member of this clade from the planthopper *Empoasca decipiens*. We further present and analyse the complete genome sequence of this microbe, assess its growth requirements compared to *Pectobacterium carotovorum* and analyse its quorum sensing–signalling system.

METHODS

Symbiont isolation, morphology *in vitro* and identification through 16s rRNA sequence

Initial Cicadellidae samples with light green coloration were collected in Prince's Park, Liverpool in April 2018, scooping different plants with an insect net at a maximum height of 2 m. Fresh specimens were transported alive to the lab and sacrificed by freezing at −20 °C for 15 min. The insect specimens were surface sterilized by immersion in 70% ethanol and washed with sterile water to remove the remaining alcohol. Insect legs were excised with a sterile surgical blade and stored at −20 °C for *post hoc* host species determination through DNA barcoding.

The remainder of the insect body was mechanically crushed and resuspended in 100 µl sterile water. An aliquot of 10 µl was plated on brain heart infusion (BHI; Oxoid,) agar and grown at 30 °C for 6 days to allow the appearance of slow-growing bacterial colonies. Morphology was examined through Gram staining and scanning electron microscopy of overnight culture. To identify the bacterial species, we performed colony PCR of the 16S rRNA gene colonies emerging on the agar plates with primers 27F (AGAGTTTGATCMTGGCTCAG) [12] and 1492R(I) (GGTTACCTTGTTACGACTT) [13], and sequenced by Eurofins Genomics (Ebersberg, Germany). Sequences were manually curated and phylogenetic analysis performed based on the 16S rRNA gene sequences which included a large assemblage of members from the *Cand*. Symbiopectobacterium clade. To this end, 16S rRNA sequences were aligned using the SSU-ALIGN software [14]. A Bayesian phylogeny was estimated with MrBayes version 3.2.6 [15] by sampling across the GTR model space (nst=mixed, rates=gamma). Two independent runs were performed for 5000000 generations and sub-sampling every 500 generations using four Markov chains. The first 25% of the samples were discarded as burn-in.

The leafhopper host was identified through DNA barcoding using the mitochondrial cytochrome oxidase I (COI) sequence. To this end, insect legs were mechanically crushed and resuspended in 50 µl of sterile water and the genomic DNA (gDNA) was extracted using a Quick-DNA Universal kit (Zymo Research). gDNA (2 µl) was added to a GoTaq Green Master Mix (Promega) and used to amplify part of CO-1 with primers C1-J-1718 (GGAGGATTTGGAAATTGATTAGTTCC) and C1-N-2191 (CCCGGTAA AATTAAAATATAAACTTC) [16]. The PCR programme consisted of an initial denaturation step at 95 °C for 5 min, followed by 30 cycles of DNA denaturation at 94 °C for 15 s, primer annealing at 55 °C for 45 s, and primer extension at 72 °C for 1 min. A final extension was carried out at 72 °C for 5 min. A few microlitres of each PCR product were run on an agarose gel to assess the success of the PCR reaction and the remains cleaned through an Isolate II PCR and Gel kit (Bioline) and sent for sequencing with primer C1-N-2191. Identity was checked through analysis against the Barcode of Life Database.

In vitro growth requirements

Biolog GEN III plates (Cat. No. 1030) were used to ascertain the physiological and biochemical characteristics of strain SyEd1^T *in vitro*, and these were conducted alongside *Pectobacterium carotovorum* subsp. *carotovorum* LMG 02404^T for comparison. Within this, we also performed the assay in the presence/absence of 0.4% polygalacturonic acid (PGA; Sigma, P3850), which is commonly used to induce the expression of plant cell wall-degrading enzymes (for preparation of PGA, see [17]. For the Biolog GEN III assays, we used IF-A inoculating fluid (Biolog, Cat. No. 72401) with or without PGA supplementation to a final concentration of 0.4% PGA. Both bacterial species were grown overnight, diluted to an OD_{600}=0.4 and 50 µl of this aliquot were added to a tube containing IF-A fluid. The aliquot in the IF-A tube was homogeneously mixed using a vortex and 100 µl of this suspension was added to each of the 96 wells of the Biolog GEN III plate. The plate was subsequently incubated at 30 °C without shaking.

Potato infection assays

Pectobacterium species are well-known plant pathogens causing soft-rot disease in several plants including potatoes, carrots and cabbages. This damage is caused by a series of secreted enzymes (cellulases, proteases, pectate lyases, pectin lyases, and polygalacturonases) commonly referred to as plant cell wall-degrading enzymes (PCWDEs). The presence of 15 putative PCWDEs and two copies of the *kdgR* regulator (associated to their expression) in strain SyEd1^T led us to assess whether this symbiont retains the plant pathogenic activity of its sister clade, *Pectobacterium*. To this aim, virulence was tested in potatoes, using a method previously described in Nadal-Jimenez *et al.* [18] with minor modifications. Briefly, potatoes were bought at local stores, washed with tap water, dried and surface sterilized with 70% ethanol. Slices about 0.5 cm thick were placed in sterile Petri dishes. Overnight cultures of strain SyEd1^T and *Pectobacterium carotovorum* LMG 02404^T were diluted to an OD600=0.4, and 20 µl were placed at the centre of the potato slice. The same amount of sterile BHI medium was added to the negative controls. The plates were sealed with parafilm to avoid moisture loss and incubated at 25 °C in dark conditions. Tissue maceration was assessed visually 24, 48 and 72 h after incubation.

Symbiont genome sequencing, assembly and annotation

The genome of the symbiont was completed using a combination of short (Illumina) and long (Nanopore) reads by MicrobesNG (Birmingham, UK) using their enhanced genome service. Briefly, Illumina sequencing was performed using the Nextera XT library prep protocol on a HiSeq platform (Illumina) using a 250 bp paired-end protocol. Reads were adapter trimmed using Trimmomatic 0.30, with a sliding window quality cutoff of Q15 [19]. Long-read genomic DNA libraries are prepared with Oxford Nanopore SQK-RBK004 kit (ONT) using 400–500 ng high molecular weight DNA and sequenced in a FLO-MIN106 (R.9.4.1) flow cell in a GridION (ONT). Hybrid genome assembly of both short and long reads was performed using Unicycler version 0.4.0 under the normal mode [20]. The final assembly was manually inspected for potential misassemblies by mapping the raw reads back to it. Genome annotation was performed with the NCBI Prokaryotic Genome Annotation Pipeline [21]. Metabolic and functional assessment of the symbiont genome was conducted using the Kyoto Encyclopaedia of Genes and Genomes database [22]. Identification of secondary metabolite biosynthesis gene clusters was performed using the antiSMASH server [23]. Finally, prophage regions were predicted using the PHAge Search Tool Enhanced Release web server [24].

Phylogenomic analysis

The phylogenetic position of the symbiont was assessed based on the concatenated analysis of 527 single-copy core proteins identified among 56 publicly available genomes. These include members of the closely related genera *Pectobacterium*, *Brenneria*, *Dickeya*, *Lonsdalea*, *Sodalis* and the recently characterized *Cand*. Symbiopectobacterium [5]. Single-copy orthologue protein sequences were identified using OrthoFinder version 2.3.11 11 [25]. A maximum-likelihood phylogeny was inferred with IQ-TREE 2.0.3 [26] using the JTT+F+R3 substitution model selected using ModelFinder according to the Bayesian information criterion [27]. Clade support was assessed based on 1000 ultrafast bootstrap replicates [28]. Comparative gene content analysis across members of the *Cand*. Symbiopectobacterium clade was performed using the anvi'o package version 7 (DOI: 10.1038/s41564-020-00834-3). Gene clustering was performed using the following parameters '--use-ncbi-blast --mcl-inflation 1.5'. The average nucleotide identity (ANI) values between the strain SyEd1^T and other *Cand*. Symbiopectobacterium genomes were computed using the PyANI method [29] as implemented in anvio package.

Analysis of *N*-acyl homoserine lactone synthesis

Strain SyEd1^T cultures were grown in 5 ml BHI medium at 30 °C for 24 h and 200 r.p.m. After incubation, the cultures were centrifuged, and the supernatant collected and filtered through a 0.2 mm filter (SLGP033RS, Millipore). Acidified ethyl acetate (500 µl) was added to 1 ml supernatant sample and the mixture was vortexed for 1–2 min. Subsequently, the mixture was centrifuged for 1 min to allow the formation of a clear interface between the aqueous and organic layer. The organic (upper) layer was transferred using a pipette (without disturbing the aqueous layer) to a new 2 ml Eppendorf. The extraction process was repeated twice more, combining the extracts for each sample into one of approximately 1.5 ml extract. Upon completion, the samples were dried under vacuum in a centrifugal evaporator.

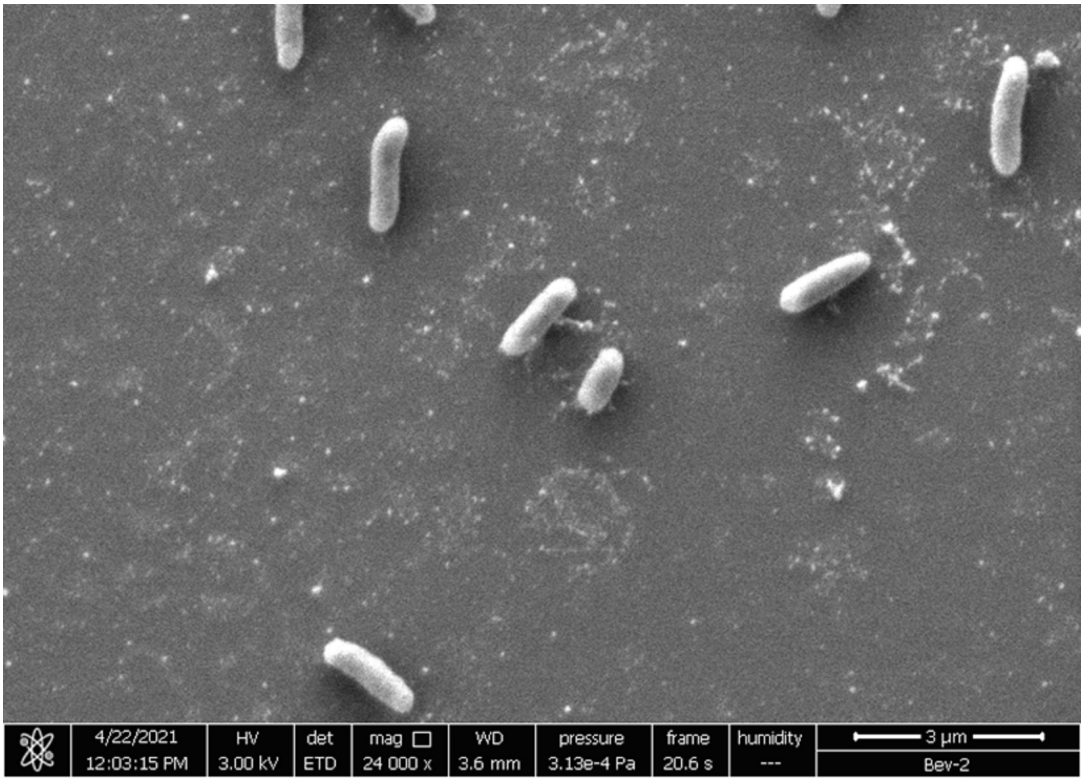

Fig. 1. SEM image of strain SyEd1^T.

Dried extract samples were reconstituted in 50 µl methanol prior to analysis. LC-MS/MS analysis of 5 µl sample injections were conducted using a Qtrap6500 hybrid triple-quadrupole linear ion trap mass spectrometer in tandem with an Exion LC system (Sciex). The overall method was a modification of that described by Ortori *et al.* [30]. Chromatography was achieved using a Phenomenex Gemini C18 column (3.0 µm, 50×3.0 mm) with a constant flow rate of 450 µl min^{-1} of mobile phase A (0.1% (v/v) formic acid) and mobile phase B (0.1% (v/v) formic acid in methanol). The LC gradient began at 10% B for 1.0 min, increased linearly to 50% B over 0.5 min, then to 99% B over 4.0 min. The composition remained at 99% B for 1.5 min, decreased to 10% B over 0.1 min, and stayed at this composition for 2.9 min. Analyte detection was conducted with the MS operating in multiple reaction monitoring mode, screening the LC eluent for specific AHLs (unsubstituted, 3-oxo and 3-OH AHLs with even chain lengths from C4 to C14).

Prevalence of symbiont in *Empoasca decipiens* leafhoppers

In order to assess the prevalence of this bacterium in *E. decipiens*, we performed a PCR screening on various specimens. Additional insect collections were completed at the same location in August 2019 and tested for strain SyEd1^T by PCR assay. Using the full-genome sequence of our cultured strain, we developed two set of specific PCR primers to amplify part of the DNA gyrase subunit (*gyrB*) gene of this bacterium: BEV_gyrB_F1 (CCGTGGTGTCGGTGAAAGTA) +BEV_gyrB_R1 (TGGTCTTCT-GTCAGCGTGTC) and BEV_gyrB_F2 (CTCGTGAAATGACACGACGC) +BEV_gyrB_R2 (CAGCAGTTCCACTTGTTCGC). The gDNA was extracted in the same manner as for the leg samples and used as a template for the PCR reactions.

RESULTS

Symbiont isolation and identification

The bacterium grows under standard aerobic conditions in brain heart infusion (BHI) medium (CM1032, Oxoid), forming circular white colonies approx. 2–3 mm in diameter on BHI agar, and cultures emitted a pronounced plant-like odour similar to that of wet straw/hay. The bacterium is Gram negative, and SEM revealed it to be a non-flagellated rod shape, of length 1–1.5 µm (Fig. 1). The bacterium will also grow in LB (Miller; 110285, Millipore/Merck KGaA) although at a slower rate, and growth is inhibited by light.

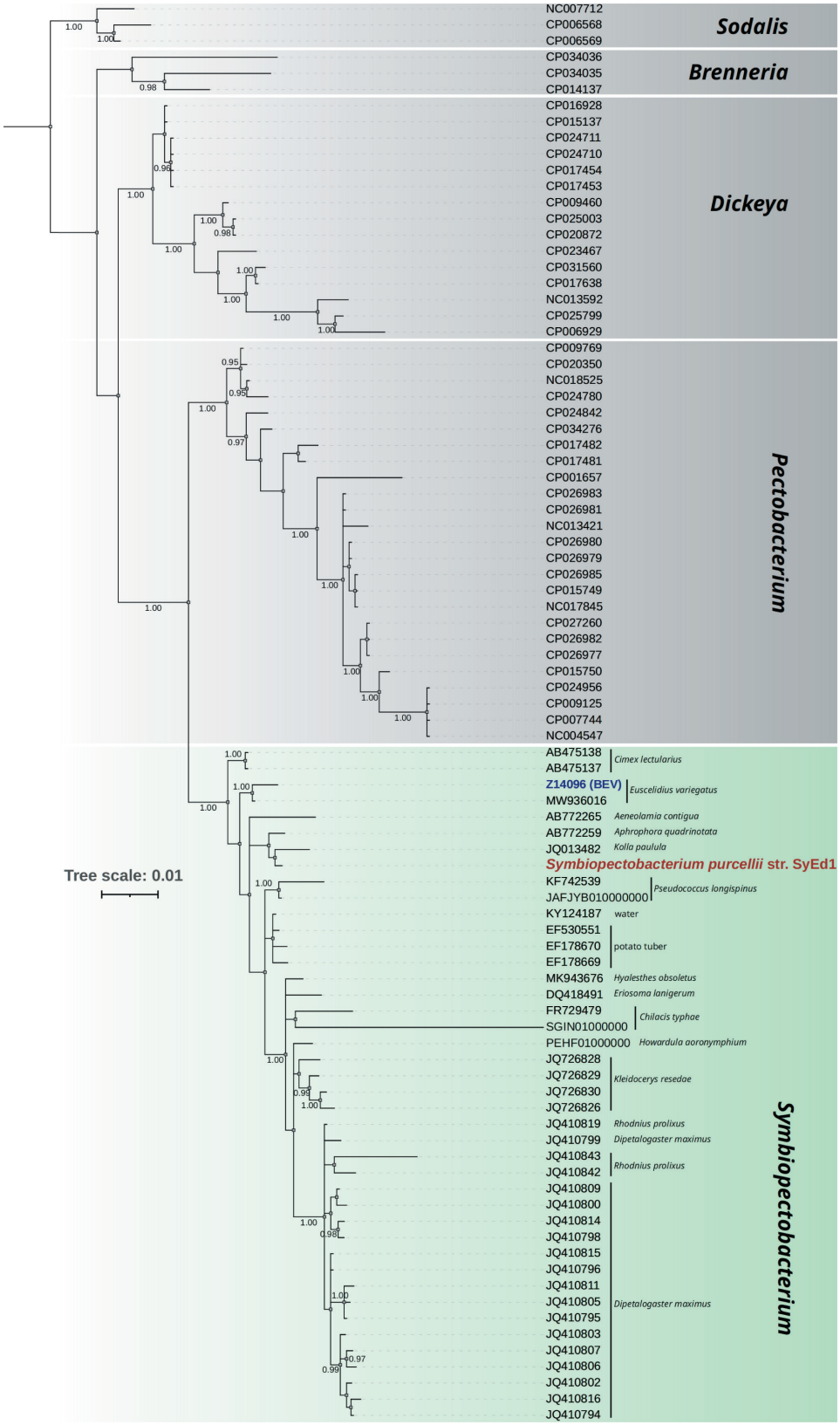

Fig. 2. Phylogenetic affiliation of the 16S rRNA of strain SyEd1^{T} (indicated in red) compared to other strains, as estimated with Mr Bayes. Numbers on nodes represent posterior probability. The *Candidatus* Symbiopectobacterium 'BEV' bacterium originally isolated from the leafhopper *Euscelidius variegatus* is also highlighted in blue.

Table 1. Utilization of carbon sources for growth by strain SyEd1^T in the presence and absence of PGA, with comparison to *Pectobacterium carotovorum* LMG 2404^T

+++, Strong growth; ++, medium growth; +, weak growth; –, no growth.

	Strain SyEd1^T	**Strain SyEd1^T+PGA**	***P. carotovorum.* LMG 2404^T**	***P. carotovorum* LMG 2404^T+PGA**
Dextrin	–	–	+++	+++
Maltose	–	–	+++	+++
Trehalose	–	–	+++	+++
Cellobiose	–	–	+++	+++
Gentiobiose	–	–	+++	+++
Sucrose	–	–	+++	+++
Turanose	–	–	+++	+++
Stachyose	–	–	+++	+++
Raffinose	–	–	+++	+++
Lactose	–	–	+++	+++
Melibiose	–	–	+++	+++
Methyl β-D-glucoside	–	–	+++	+++
D-Salicin	–	–	+++	+++
N-Acetyl-D-glucosamine	+++	+++	+++	+++
N-Acetyl-β-D-mannosamine	–	–	+++	+++
N-Acetyl-D-galactosamine	–	–	+++	+++
N-Acetyl neuraminic acid	–	–	+++	+++
α-D-Glucose	+++	+++	+++	+++
D-Mannose	+++	+++	+++	+++
D-Fructose	+++	+++	+++	+++
D-Galactose	–	–	+++	+++
3-Methyl glucose	–	–	+++	+++
D-Fucose	–	–	+++	+++
L-Fucose	–	–	+++	+++
L-Rhamnose	+++	+	+++	+++
Inosine	–	–	+++	+++
D-Sorbitol	–	–	+++	+++
D-Mannitol	–	–	+++	+++
D-Arabitol	–	–	+++	+++
myo-Inositol	–	–	+++	+++
Glycerol	+++	+++	+++	+++
D-Glucose-6-PO4	–	–	+++	+++
D-Fructose-6-PO4	–	–	+++	+++
D-Aspartic acid	–	–	+++	+++
D-Serine	–	–	+++	+++

Continued

Table 1. Continued

	Strain SyEd1^T	**Strain SyEd1^T+PGA**	***P. carotovorum.* LMG 2404^T**	***P. carotovorum* LMG 2404^T+PGA**
Gelatin	–	–	+++	+++
Glycyl-L-proline	–	–	+++	+++
L-Alanine	–	–	+++	+++
L-Arginine	–	–	+++	+++
L-Aspartic acid	–	–	+++	+++
L-Glutamic acid	–	–	+++	+++
L-Histidine	–	–	–	–
L-Pyroglutamic acid	–	–	–	–
L-Serine	–	–	+++	+++
Pectin	–	–	+++	+++
D-Galacturonic acid	–	–	+++	+++
L-Galactonic acid lactone	–	–	+++	+++
D-Gluconic acid	–	–	+++	+++
D-Glucuronic acid	–	–	+++	+++
Glucuronamide	–	–	+++	+++
Mucic Acid	–	–	+++	+++
Quinic acid	–	–	–	–
D-Saccharic acid	–	–	+++	+++
ρ-Hydroxy phenyl acetic acid	–	–	–	–
Methyl pyruvate	+	+	+++	+++
D-Lactic Acid methyl ester	–	–	–	–
L-Lactic acid	–	–	–	–
Citric acid	–	–	+++	+++
α-Keto-glutaric-acid	–	–	–	–
D-Malic acid	–	–	–	–
L-Malic acid	–	–	+++	+++
Bromo-succinic acid	–	–	+++	+++
Tween 40	–	–	–	+++
γ-Amino-butyric acid	–	–	–	+++
α-Hydroxy-butyric Acid	–	–	–	+++
β-Hydroxy-D-L-butyric acid	–	–	–	+++
α-Keto-butyric-acid	–	++	–	+++
Acetoacetic acid	–	–	+	+++
Propionic acid	–	–	–	–
Acetic acid	+	+	+++	+++
Formic acid	–	–	+++	+++

Table 2. Impact of environmental and xenobiotic stress conditions on strain SyEd1^T growth on Biolog III plates compared to *P. carotovorum* LMG 2404^T, in the presence/absence of PGA

+++, Maintains full growth under condition stated; ++, medium growth; +, weak growth; –, no growth under condition stated.

	Strain SyEd1^T	**Strain SyEd1^T+PGA**	***P. carotovorum* LMG 2404^T**	***P. carotovorum* LMG 2404^T+ PGA**
pH 6	+++	+++	+++	+++
pH 5	–	–	–	–
1% NaCl	+++	+++	+++	+++
4% NaCl	+	+	+++	+++
8% NaCl	–	–	–	–
1% Sodium lactate	+++	+++	+++	+++
Fusidic acid	+++	+++	+++	+++
D-Serine	–	+++	–	–
Troleandomycin	+++	+++	+++	+++
Rifamycin SV	+++	+++	+++	+++
Minocycline	–	+++	–	–
Lincomycin	++	++	+++	+++
Guanidine HCl	++	++	+++	+++
Niaproof 4	++	++	+++	+++
Vancomycin	+++	+++	+++	+++
Tetrazolium violet	++	++	+++	+++
Tetrazolium blue	+++	+++	+++	+++
Nalidixic acid	++	++	–	–
Lithium chloride	+++	+++	+++	+++
Potassium tellurite	+++	+++	–	–
Aztreonam	+++	+++	–	–
Sodium butyrate	++	++	+++	+++
Sodium bromate	–	–	–	–

Phylogenetic analysis based on the 16S rRNA gene (accession number OK044380) placed the isolated microbe well within the recently characterized clade *Cand*. Symbiopectobacterium (Fig. 2), a group of microbes commonly associated with arthropods and nematodes. Sequence of the CO-1 amplicons from the insect host revealed the leafhopper host to be *Empoasca decipiens* (Hemiptera, Cicadellidae), a common species of leafhopper in Europe. *Empoasca decipiens* has been implicated in the transmission of various plant pathogens [31], and is considered a pest in various crops [32].

In vitro growth requirements

Strain SyEd1^T and *P. carotovorum* LMG 2404^T were grown at 30 °C in Biolog GEN III plates. For *P. carotovorum* LMG 2404^T, the presence of the purple tetrazolium dye as a result of growth and respiration in the wells where the strains had grown was visible after 24 h, while the wells that did not support the growth of this strain remained colourless. In the case of strain SyEd1^T, the plates had to be incubated for a total of 72 h to allow bacterial growth. This is not surprising since, in our hands, strain SyEd1^T grows slowly in BHI media (requiring up to 48 h), and even slower in less rich media. Analyses indicated that strain SyEd1^T was considerably more fastidious than the comparator outgroup strain *P. carotovorum* LMG 2404^T in terms of metabolites that supported growth (Table 1) but had broader resistance to xenobiotics than this strain (Table 2). Growth conditions for strain SyEd1^T in the Biolog analysis were only modestly altered by addition of PGA.

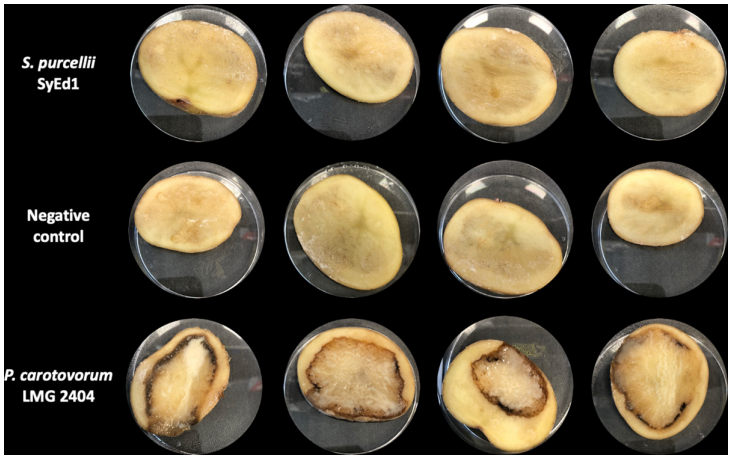

Fig. 3. Virulence assay in potatoes.

Potato infection assays

Potato slices infected with strain SyEd1^T exhibited a complete absence of infection/ tissue maceration at the different time points tested (72 h time point shown in Fig. 3) in contrast to *P. carotovorum* LMG 02404^T, used as positive control for infection. The assay was maintained for a week without any sign of infection being visible in strain SyEd1^T-infected potatoes.

Genome sequence and assembly

The genome of the symbiont presented as a single circular chromosome of circa 4.9 MB with an average G+C content of 52.5 mol% (Table 3). No plasmids were identified. The complete predicted gene set consists of 4494 protein-coding genes (including 312 predicted pseudogenes), seven ribosomal RNA operons (5S, 16S, 23S) and 76 tRNAs. The average length of the protein-coding genes is 948 bp accounting for a coding density of about 86.2%. Pseudogenization rates were estimated to be circa 7% (312 predicted pseudogenes). The main chromosome was predicted to contain six intact prophage regions and three additional incomplete fragments. The complete genome assembly and the raw reads have been submitted to the DDBJ/EMBL/GenBank database under the BioProject accession number PRJNA756769 (genome accession number CP081864).

Phylogenomic and functional analysis

To confirm the phylogenetic position of the *E. decipiens* symbiont we conducted a phylogenomic analysis base on the concatenated set of 527 single-copy orthologue proteins across 56 related genomes (Fig. 4). These results further support the placement of the symbiont in the *Cand*. Symbiopectobacterium clade. Within the existing clade *Cand*. Symbiopectobacterium, ANI of strain SyEd1^T to other sequenced genomes was greater than 90% but below 95% in all cases, indicating that these other strains belong within the same genus as strain SyEd1^T, but can be considered distinct species from it (Table 4). Comparative gene content within the

Table 3. Genome features of the strain SyEd1^T isolated from the leafhopper *Empoasca decipiens*

Feature	
Chromosome size (bp)	4942431
Plasmids	no
G+C content (mol%)	52.5
Number of predicted CDS	4494
Number of pseudogenes	312
Average CDS length (bp)	948
Coding density (%)	86.2
Number of rRNA operons	7 (5S, 16S, 23S)
Number of tRNA genes	76

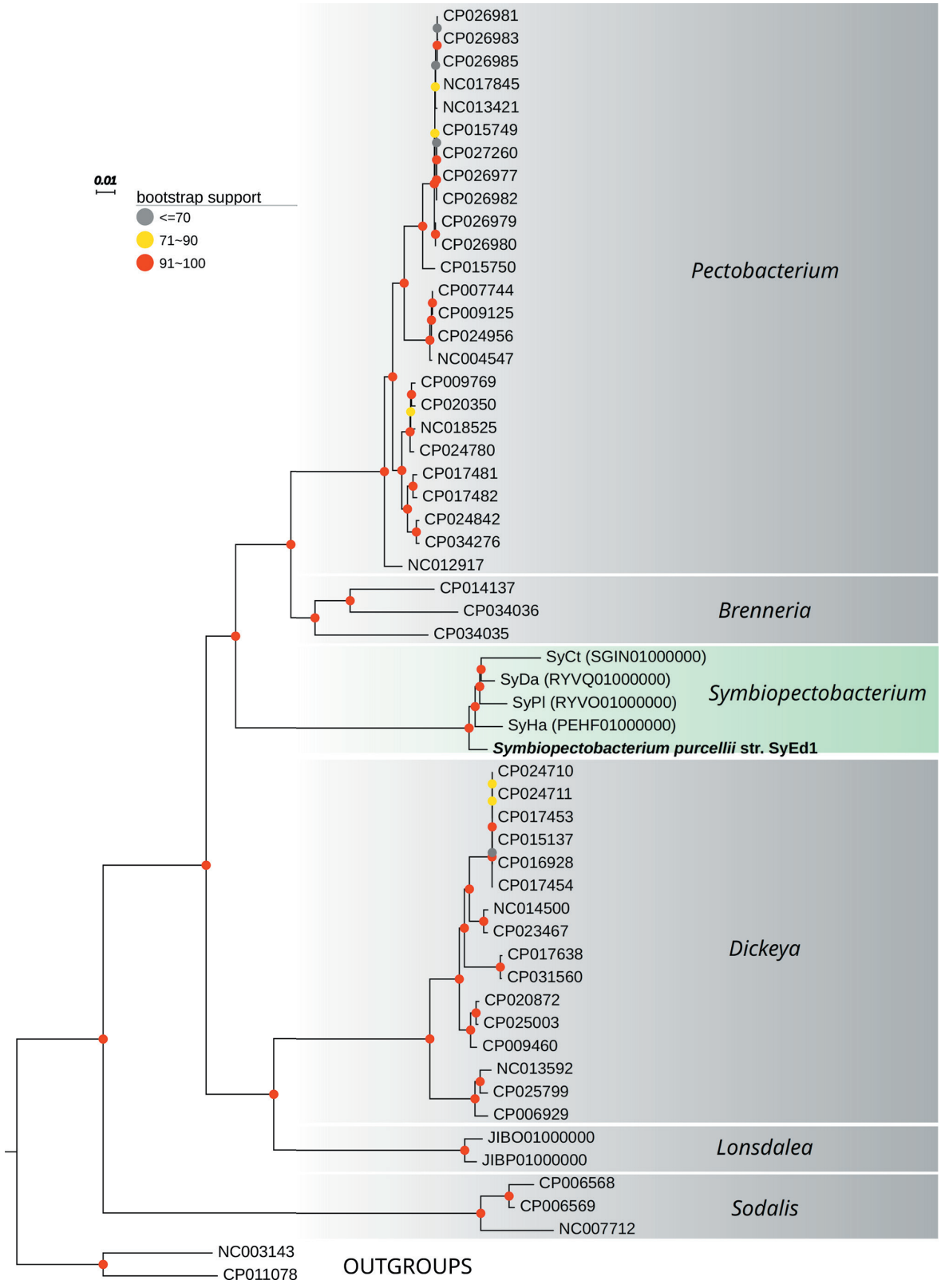

Fig. 4. Affiliation of strain SyEd1^{T} with other bacteria as estimated using IQ-TREE based on 527 shared single copy orthologues. Coloured dots on nodes represent bootstrap support. SyCt, *Candidatus* Symbiopectobacterium symbiont of *Chilacistyphae*; SyDa, *Candidatus* Symbiopectobacterium symbiont of *Diachasma alloeum*; SyPl: *Candidatus* Symbiopectobacterium symbiont of *Pseudococcus longispinus*; SyHa, *Candidatus* Symbiopectobacterium symbiont of *Howardula aoronymphium*.

Table 4. Average nucleotide identity (%) of SyEd (from *Empoasca decipiens*) to other sequenced genomes in *Candidatus* Symbiopectobacterium

SyCt, *Cand.* Symbiopectobacterium symbiont of *Chilacistyphae*; SyDa, *Cand.* Symbiopectobacterium symbiont of *Diachasma alloeum*; SyHa, *Cand.* Symbiopectobacterium symbiont of *Howardula aoronymphium*; SyPl, *Cand.* Symbiopectobacterium symbiont of *Pseudococcus longispinus*.

	SyCt	SyDa	SyEd1^T	SyHa	SyPl
SyCt	100	96.17	91.5	93.98	95.74
SyDa	96.17	100	92.05	94.63	96.42
SyEd	91.5	92.05	100	91.92	91.5
SyHa	93.98	94.63	91.92	100	94.2
SyPl	95.74	96.42	91.5	94.2	100

existing clade *Cand*. Symbiopectobacterium is given in figure S1, available in the online version of this article, though absence of function in strains aside SyEd1^T should be treated with caution, as the genome sequences are incomplete.

The genome is predicted to encode type II, III and VI secretion systems alongside a wide array of predicted secreted toxins, compatible with its likely status as a symbiont of its insect host. Anti-SMASH predicted five genomic regions associated with small molecule production. Notable amongst these is a non-ribosomal peptide synthase (NRPS) region. It is unclear if the NRPS system produces siderophore molecules that permit growth in iron-poor host environments, or antimicrobial compounds. In addition, anti-SMASH predicted putative gene clusters for the biosynthesis of thiopeptide, an aryl-polyene potentially providing defence against reactive oxygen species, and betalactone synthesis. There is also a predicted homoserine lactone synthesis island *expI*/*expR1* that may be involved in sensing of microbial titre (see below); the genome encodes additional conserved elements of the quorum sensing system, *expR2*, *gacA (expA)*, *gacS (expS)*, *rsmA* and *rsmB,* and *kdgR*. Finally, the genome encodes complete biosynthetic pathways for several B vitamins including thiamine (B1), riboflavin (B2), pantothenate (B5), biotin (B7), pyridoxine (B6) and folate (B9). A broad array of complete amino acid biosynthesis pathways was also observed, including serine, threonine, cysteine, methionine, valine, leucine, isoleucine, arginine, ornithine, arginine, proline, histidine, tryptophan, phenylalanine and tyrosine. Vitamin and amino acid provision are common means through which symbionts contribute to host function. There are also five *pel* genes predicted to encoded pectate lyase enzymes. The failure of the strain to utilize pectin on Biolog plates or on potato tubers may thus be context-dependent.

Analysis of homoserine lactones

Strain SyEd1^T analysis using LC-MS/MS revealed the presence of a single AHL that was characterized as *N*-(3-hydroxydecanoyl)-L-homoserine lactone (3-OH-C10-HSL). Fig. 5 shows the LC-MS/MS chromatogram obtained from the strain SyEd1^T sample compared to the 3-OH-C10-HSL standard and the uncultured BHI medium. Members of the genus *Pectobacterium* have been reported to produce 3-oxo-C6-HSL, 3-oxo-C8-HSL, C10-HSL [33], but, to the best of our knowledge, the presence of 3-OH-C10-HSL as the sole AHL in strain SyEd1^T is unreported in related genera. This trait may help to identify novel members of this genus that we presume may have been previously misidentified as *Pectobacterium* species associated with the plants where the leafhoppers feed.

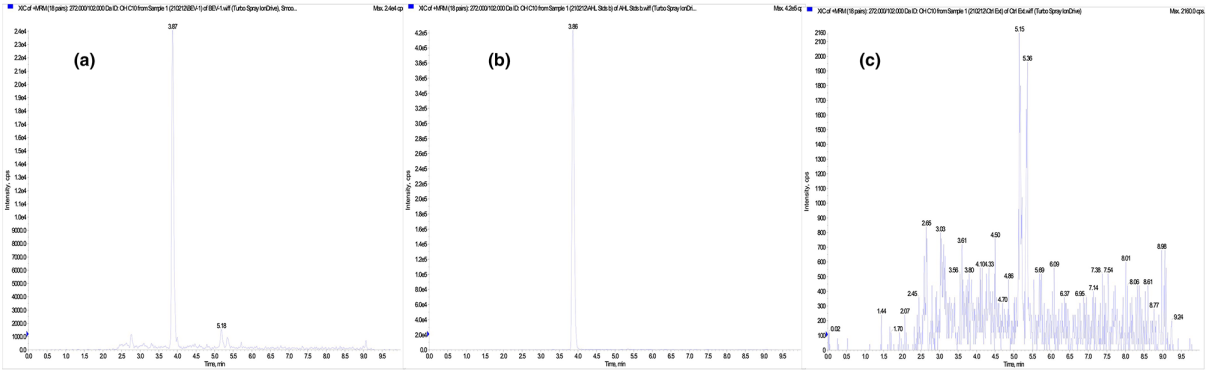

Fig. 5. LC-MS/MS of an extraction of strain SyEd1^T grown in BHI medium (a), 3-OH-C10-HSL standard (b) and negative control using a sterile BHI medium extract (c).

Prevalence of strain SyEd1^T in *E. decipiens* samples

Seven new *E. decipiens* specimens were collected and their identity confirmed by CO-1 amplification and sequencing. All samples were confirmed to be *E. decipiens* with >99.6% identity with previously deposited sequences in NCBI. The same gDNA extract was used to screen for the presence of the bacterial symbiont by PCR using our BEV_gyrB_F2 and BEV_gyrB_R2 primers. All seven samples produced an amplicon for strain SyEd1^T [10], and the identity of the amplicon was confirmed through sequencing accounting for a 100% prevalence in the population tested (95%; confidence interval 64–100%).

DESCRIPTION OF *SYMBIOPECTOBACTERIUM* GEN. NOV.

Symbiopectobacterium (Sym.bi.o.pec.to.bac.te'ri.um. Gr. masc./fem. n. *symbios*, companion; N.L. neut. n. *Pectobacterium*, a bacterial genus; N.L. neut. n. *Symbiopectobacterium*). *Symbiopectobacterium* references the related *Candidatus* Symbiopectobacterium that is an obligate symbiont of nematode worms [5], this name reflecting the symbiotic habit of the microbe, and the relationship of the genus as sister to *Pectobacterium*. The genus lies within the *Gammaproteobacteria* and contains a number of insect symbionts as well as an obligate symbiont of *Howardula* nematode worms. The type species is *Symbiopectobacterium purcellii*.

DESCRIPTION OF *SYMBIOPECTOBACTERIUM PURCELLII* SP. NOV.

Symbiopectobacterium purcellii [pur.cel'li.i. N.L. gen. n. *purcellii* in reference to Alexander Purcell, who isolated the first member of this clade, which he named the BEV symbiont (bacterium from *Euscelidius variegatus*)].

Symbiopectobacterium is an obligate symbiont of nematode worms, this name reflecting the symbiotic habit of the microbe, and the relationship of the genus as sister to *Pectobacterium. S. purcellii* is a Gram-negative rod-shaped non-flagellated bacterium of 1–1.5 µm in length. Grows optimally at 30 °C in BHI medium in the dark forming colonies within 24–48 h. Using Biolog GENIII plates, strain SyEd1^T responded positively to the following carbon sources: D-glucose, D-mannose, D-fructose, glycerol, *N*-acetyl glucosamine and L-rhamnose, and weakly to methyl pyruvate and acetic acid. Growth was inhibited at pH5, by 4 and 8% NaCl, and by D-serine, minocycline and sodium bromate. Growth was not impaired by 1% sodium lactate, fusidic acid, troleandomycin, rifamycin S, lincomycin, guanidine HCl, vancomycin, tetrazolium violet, tetrazolium blue, potassium tellurite, nalidixic acid, lithium chloride, aztreonam, sodium butyrate. The microbe does not cause macerations on potato slices.

S. purcellii forms a cluster with a variety of uncultured symbionts of insects and nematodes, as well as the previously cultured strain BEV.

The type strain is SyEd1^T (LMG 32449^T=CECT 30436^T) and was isolated from *Empoasca decipiens* (Hexapoda: Hemiptera: Cicadellidae) from Liverpool, UK (53.3868° N, 2.9565° W). The genome consists of a single circular chromosome of size 4.9 MB and DNA G+C content is 52.5 mol%.

Funding information
This work was funded by a BBSRC grant to GH (grant BB/S017534/1). Miguel Cámara is partly funded by the National Biofilms Innovation Centre (NBIC) which is an Innovation and Knowledge Centre funded by the Biotechnology and Biological Sciences Research Council, Innovate UK and Hartree Centre (Award Number BB/R012415/1).

Acknowledgements
We are thankful to Alison Beckett (University of Liverpool, UK) for SEM imaging, Sam Edwards (University of Copenhagen, DK) for assistance in *E. decipiens* collection, Nigel Gotts (University of Liverpool, UK) for assistance evaporating the AHL-containing extracts, and Rita Valente (Instituto Gulbenkian de Ciência, PT) for useful discussions about the *Pectobacterium* virulence assays

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

References
1. McFall-Ngai M, Hadfield MG, Bosch TCG, Carey HV, Domazet-Lošo T, *et al*. Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad Sci USA* 2013;110:3229–3236.
2. Hurst GDD. Extended genomes: symbiosis and evolution. *Interface Focus* 2017;7:20170001.
3. Gerardo N, Hurst G. Q&A: friends (but sometimes foes) within: the complex evolutionary ecology of symbioses between host and microbes. *BMC Biol* 2017;15:216.
4. Duron O, Bouchon D, Boutin S, Bellamy L, Zhou L, *et al*. The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. *BMC Biol* 2008;6:27.
5. Martinson VG, Gawryluk RMR, Gowen BE, Curtis CI, Jaenike J, *et al*. Multiple origins of obligate nematode and insect symbionts by a clade of bacteria closely related to plant pathogens. *Proc Natl Acad Sci U S A* 2020;117:31979–31986.
6. Purcell AH, Steiner T, Mégraud F, Bové J. In vitro isolation of a transovarially transmitted bacterium from the leafhopper *Euscelidius variegatus* (Hemiptera: Cicadellidae). *J Invertebr Pathol* 1986;48:66–73.
7. Purcell AH, Suslow KG, Klein M. Transmission via plants of an insect pathogenic bacterium that does not multiply or move in plants. *Microb Ecol* 1994;27:19–26.
8. Purcell AH, Suslow KG. Pathogenicity and effects on transmission of a mycoplasmalike organism of a transovarially infective

bacterium on the leafhopper *Euscelidius variegatus* (Homoptera: Cicadellidae). *J Invertebr Pathol* 1987;50:285–290.

9. Hypsa V, Aksoy S. Phylogenetic characterization of two transovarially transmitted endosymbionts of the bedbug *Cimex lectularius* (Heteroptera:Cimicidae). *Insect Mol Biol* 1997;6:301–304.
10. Kuechler SM, Dettner K, Kehl S. Characterization of an obligate intracellular bacterium in the midgut epithelium of the bulrush bug *Chilacis typhae* (Heteroptera, Lygaeidae, Artheneinae). *Appl Environ Microbiol* 2011;77:2869–2876.
11. Degnan PH, Bittleston LS, Hansen AK, Sabree ZL, Moran NA, *et al*. Origin and examination of a leafhopper facultative endosymbiont. *Curr Microbiol* 2011;62:1565–1572.
12. Lane DJ. 16S/23S rRNA sequencing. In: Stackebrandt E and Goodfellow M (eds). *Nucleic Acid Techniques in Bacterial Systematics*. New York: John Wiley & Sons; 1991. pp. 115–175.
13. Turner S, Pryer KM, Miao VP, Palmer JD. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *J Eukaryot Microbiol* 1999;46:327–338.
14. Nawrocki EP. Structural RNA Homology Search and Alignment using Covariance Models [PhD]: Washington University in Saint Louis. 2009.
15. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, *et al*. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 2012;61:539–542.
16. Simon C, Frati F, Beckenbach A, Crespi B, Liu H, *et al*. Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymease chain reaction primers. *Ann Entomol Soc Am* 1994;87:651–701.
17. Nadal Jimenez P, Valente RS. Preparation of polygalacturonic acid (PGA) to study virulence in Erwinia/Pectobacterium. protocolsio [Internet]. 2021.
18. Jimenez PN, Koch G, Papaioannou E, Wahjudi M, Krzeslak J, *et al*. Role of PvdQ in *Pseudomonas aeruginosa* virulence under iron-limiting conditions. *Microbiology* 2010;156:49–59.
19. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114–2120.
20. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Comput Biol* 2017;13:e1005595.
21. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, *et al*. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 2016;44:6614–6624.
22. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000;28:27–30.
23. Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, *et al*. antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res* 2021;49:W29–W35.
24. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, *et al*. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 2016;44:W16-21.
25. Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* 2015;16:157.
26. 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.
27. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 2017;14:587–589.
28. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol* 2018;35:518–522.
29. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal Methods* 2016;8:12–24.
30. Ortori CA, Halliday N, Camara M, Williams P, Barrett DA. LC-MS/MS quantitative analysis of quorum sensing signal molecules. In: Filloux A and Ramos JL (eds). *Pseudomonas: Methods and Protocols. Methods in Molecular Biology*. Springer; 2014. pp. 55–70.:11492014.
31. Galetto L, Marzachì C, Demichelis S, Bosco D. Host plant determines the phytoplasma transmission competence of *Empoasca decipiens* (Hemiptera: Cicadellidae). *J Econ Entomol* 2011;104:360–366.
32. Fathi SAA, Nouri-Ganb G, Rafiee-Das H. Life Cycle Parameters of *Empoasca decipiens* Paoli (Hom.: Cicadellidae) on four potato cultivars (*Solanum tuberosum* L.) in Iran. *J Entomology* 2009;6:96–101.
33. Rodriguez-R LM, Konstantinidis KT. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Preprints* 2016;4:1900V1.