

A discovery down under: decoding the draft genome sequence of *Pantoea stewartii* from Australia's Critically Endangered western ground parrot/kyloring (*Pezoporus flaviventris*)

Rhys T. White¹, William Taylor², Natalie Klukowski³, Rebecca Vaughan-Higgins⁴, Ernest Williams⁵, Steve Petrovski³, Jayson J. A. Rose³ and Subir Sarker^{3,6,*}

Abstract

Pantoea stewartii, a plant pathogen, is primarily transmitted through contaminated seeds and insect vectors, with the corn flea beetle (*Chaetocnema pulicaria*) being the primary carrier. *P. stewartii* is a bacterium belonging to the order *Enterobacterales* and can lead to crop diseases that have a significant economic impact worldwide. Due to its high potential for spread, *P. stewartii* is classified as a quarantine organism in numerous countries. Despite its impact on agriculture, the limited genome sequences of *P. stewartii* hamper understanding of its pathogenicity and host specificity, and the development of effective control strategies. In this study, a *P. stewartii* strain (C10109_Jinnung) was discovered in the faecal matter of the Critically Endangered western ground parrot/kyloring (*Pezoporus flaviventris*) in Australia, which to our knowledge is the first reported *P. stewartii* genome from a bird source. Whole-genome sequencing and phylogenomic analysis of strain C10109_Jinnung, obtained from a captive psittacine, provides new insights into the genetic diversity and potential transmission route for the spread of *P. stewartii* beyond insects and plants, where *P. stewartii* is typically studied. Our findings provide new insights into the potential transmission route for spread of *P. stewartii* and expand the known transmission agents beyond insects and plants. Expanding the catalogue of *P. stewartii* genomes is fundamental to improving understanding of the pathogenicity, evolution and dissemination, and to develop effective control strategies to reduce the substantial economic losses associated with *P. stewartii* in various crops and the potential impact of endangered animal species.

DATA SUMMARY

The study sequences are available in the National Center for Biotechnology Information (NCBI) under BioProject accession number PRJNA967810. The raw Illumina sequence read data generated in this study have been deposited at the NCBI sequence read archive [SRA (<https://www.ncbi.nlm.nih.gov/sra>)] under accession number SRR24455997. The draft assembly has been deposited in GenBank under accession number JASCYM000000000. The software used to analyse raw sequence reads for polymorphism discovery and whole-genome sequencing-based phylogenetic reconstruction are available as described in the Methods. The authors confirm that all supporting data protocols have been provided in the article or supplementary data files.

Received 25 May 2023; Accepted 13 August 2023; Published 04 September 2023

Author affiliations: ¹Institute of Environmental Science and Research, Wellington, New Zealand; ²Institute of Environmental Science and Research, Christchurch, New Zealand; ³Department of Microbiology, Anatomy, Physiology and Pharmacology, School of Agriculture, Biomedicine and Environment, La Trobe University, Melbourne, Victoria, Australia; ⁴Perth Zoo, South Perth, Western Australia, Australia; ⁵Institute of Environmental Science and Research, Wallaceville, New Zealand; ⁶Biomedical Sciences & Molecular Biology, College of Public Health, Medical and Veterinary Sciences, James Cook University, Townsville, QLD, Australia.

*Correspondence: Subir Sarker, subir.sarker@jcu.edu.au

Keywords: Australia; birds; *Enterobacterales*; *Pantoea stewartii* subsp. *indologenes*; phylogenomics; strain C10109_Jinnung.

Abbreviations: AGRF, Australian Genome Research Facility; BWA, Burrows–Wheeler Aligner; CDS, coding sequence; contigs, contiguous sequences; NCBI, National Center for Biotechnology Information; PBFd, psittacine beak and feather disease; RefSeq, Reference Sequence; SNV, single-nucleotide variant; SRA, Sequence Read Archive.

National Center for Biotechnology Information (NCBI) BioProject accession number PRJNA967810. NCBI sequence read archive (SRA) accession number SRR24455997. GenBank accession number JASCYM000000000. More information about sequence data is in the data summary section below.

Data statement: All supporting data, code and protocols have been provided within the article or through supplementary data files. Four supplementary figures and three supplementary tables are available with the online version of this article.

001101 © 2023 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution License. This article was made open access via a Publish and Read agreement between the Microbiology Society and the corresponding author's institution.

Impact Statement

This study reports the draft genome sequence of a *Pantoea stewartii* strain, C10109_Jinnung, found in the faecal matter of a Critically Endangered bird species, the western ground parrot/kyloring (*Pezoporus flaviventris*), in Australia. Genome-wide taxonomic analysis showed that *P. stewartii* genome CCUG 26359 was the nearest taxonomic neighbour to C10109_Jinnung within the *P. stewartii* population. Phylogenomic analyses revealed that C10109_Jinnung was most genomically similar to the *P. stewartii* subsp. *indologenes* genome ZJ-FGZX1, isolated from lucky bamboo (*Dracaena sanderiana*) in China. However, a substantial genomic divergence was indicated by a distance of 18697 pairwise single-nucleotide variants between the genomes of ZJ-FGZX1 and C10109_Jinnung. These findings expand the known transmission pathways for the spread of *P. stewartii* beyond insects and plants to birds, which may have important implications for developing effective control strategies. The results of this study provide new insights into the genetic diversity of *P. stewartii* and highlight the need for further genomic studies to better understand the molecular basis of pathogenicity, evolution and dissemination of this important plant pathogen.

INTRODUCTION

Pantoea stewartii (syn. *Erwinia stewartii*) is a Gram-negative bacterium that causes Stewart's wilt (young seedlings) and leaf blight (mature plants) of economically important crops such as jackfruit (*Artocarpus heterophyllus*) [1], rice (*Oryza sativa*) [2], sudangrass (*Sorghum drummondii*) [3] and corn (*Zea mays*) [4]. In particular, the corn industry has suffered substantial economic losses due to the impact of *P. stewartii* [5–7]. For instance, in the USA in 1999, the state of New York alone reported losses exceeding USD\$1 million in corn yield attributed to Stewart's wilt [8]. Consequently, many countries classify *P. stewartii* as a quarantine organism [9].

P. stewartii can spread through contaminated seeds and insect vectors. *P. stewartii* has been associated with gut colonization and pathogenicity in various insect hosts such as the corn flea beetle (*Chaetocnema pulicaria*) [10], pea aphid (*Acyrtosiphon pisum*) [11] and yellow fever mosquito (*Aedes aegypti*) [12]. The corn flea beetle is the primary vector for *P. stewartii* [13], the spread of *P. stewartii* between plants depends on the presence of this beetle, and the incidence of Stewart's wilt/leaf blight is directly related to the density of the corn flea beetle population in corn fields [10, 14, 15]. Although *P. stewartii* substantially impacts agriculture, there is limited information on its genome sequence. Only 29 draft and three complete genome sequences have been deposited in the public National Center for Biotechnology Information (NCBI) databases (as of March 2023), including the complete genome sequences *P. stewartii* subsp. *indologenes* strain ZJ-FGZX1 [16], *P. stewartii* subsp. *stewartii* strain DC283 [17] and *P. stewartii* strain HR3-48 [18]. Therefore, expanding the catalogue of *P. stewartii* genomes is fundamental for an improved understanding of the genetic and molecular basis of pathogenicity, evolution, dissemination and potential host specificity, and for developing effective control strategies.

The main culprit responsible for causing Stewart's wilt of corn is *P. stewartii* subsp. *stewartii*, as opposed to *P. stewartii* subsp. *indologenes* [19–21]. Despite the relatively few known genotypic differences between these subspecies, they can still be distinguished based on variations in 'electrophoretic protein profiles, fatty acid composition, and biochemical properties' [19, 22]. Furthermore, specific single-nucleotide variants (SNVs) identified in housekeeping genes, such as *recA* (encoding protein RecA) and *galE* (encoding UDP-glucose 4-epimerase), provide additional markers for differentiation [20, 21]. Nevertheless, these analyses are not always completely reliable in distinguishing between different species of *Pantoea* or subspecies of *P. stewartii* [20]. However, by exploring and comparing the genomes of different *P. stewartii* subspecies, researchers can gain valuable insights into their distinct characteristics and virulence factors.

P. stewartii has yet to be documented in Australia. However, there is a considerable risk of its introduction through maize grains imported from the USA. Given that Australia has extensive maize cultivation and suitable environmental conditions for Stewart's wilt disease, there is potential for substantial yield losses. Therefore, the entry of *P. stewartii* poses a quarantine risk to Australia. The extent of this risk depends on evaluating local insects as potential vectors of *P. stewartii*, making quantification challenging. While investigating unknown circulating pathogens, *P. stewartii* was discovered in the faecal matter of the Critically Endangered western ground parrot/kyloring (*Pezoporus flaviventris*). To the best of our knowledge, there are no publicly available genomic studies of *P. stewartii* genomes from birds in Australia, which raises questions about the genetic diversity of circulating strains and their potential transmission pathways. This finding was unexpected but not unprecedented, as a similar discovery was made with *Pantoea eucalypti* in goose (*Branta canadensis*) faeces from Canada [23]. Due to the novelty of discovering *P. stewartii* from an avian source, we conducted whole-genome sequencing and phylogenomic analyses on a *P. stewartii* strain obtained from a captive psittacine in Australia. This allowed us to place the strain within the larger global phylogenetic context of *P. stewartii*. By employing metagenomic analyses, our study has yielded valuable insights into the potential mechanisms underlying the dissemination of *P. stewartii*. While our study revealed the presence of *P. stewartii* in bird faeces, seeds may serve as a more probable reservoir or source of infection for the spread of *P. stewartii*, rather than the bird itself acting as the primary host. These findings broaden the

spectrum of known transmission routes associated with *P. stewartii* beyond the conventional focus on insects and plants, thus enhancing our understanding of its transmission dynamics.

METHODS

Clinical history

The clinical history provided pertains to a male western ground parrot/kylloring captured in April 2021 as part of a translocation programme in Western Australia. The capture and subsequent transport to camp were uneventful, but a single damaged secondary feather was noted on the right wing. As only one feather was affected, the veterinarian interpreted the damage as follicular rather than feather disease. After several unsuccessful attempts to fly, the bird was deemed unfit for release to the wild and transported to Perth Zoo (South Perth, Western Australia) for further assessment and care. Upon arrival at Perth Zoo, the parrot was observed to have a right-wing droop and a small alopecic area on the crown of his head where feather loss was evident, possibly from trauma during transport. The bird tested negative for psittacine beak and feather disease (PBFDF) virus, chlamydia, polyoma virus, psittacine herpesvirus and parasites in May 2021. The bird had lost 9% of its body weight since capture, probably reflecting a combination of true weight loss, dehydration and a smaller amount of seed (native food plants with other grains) being offered (Fig. S1, available in the online version of this article). The bird's diet consisted of either a zoo mix or imported parrot pellets from the USA (which may have contained corn). The veterinarians believed the right-wing droop indicated a musculoskeletal injury and treated the western ground parrot/kylloring with the anti-inflammatory medication meloxicam at 0.5 mg kg⁻¹ via intramuscular injection. Later, the bird was observed drinking from a water bowl and eating seeds from a tray. However, the right-wing droop was still intermittently evident.

Sampling, ethical considerations and extraction of DNA

Faecal samples were obtained from the western ground parrot/kylloring declared unfit for translocation at Perth Zoo in 2021. To the best of our knowledge, there are no other samples in the world from this species/sample type as it is endemic to Western Australia, exceedingly rare, and critically endangered. Faecal samples were obtained during routine animal husbandry practice with no handling or other intervention required. The psittacine was housed individually in the aviary, and the faecal sample was collected from the temporary holding box (Fig. S1). The attending veterinarian collected the faecal samples and shipped samples to La Trobe University (Melbourne, Victoria, Australia) to investigate circulating unknown pathogens. The Animal Ethics Committee at La Trobe University was informed that findings from the material (with no bird intervention) were to be used in a publication, and a formal waiver of ethics approval was granted. Faecal materials were aseptically resuspended and homogenized vigorously in sterile PBS (1:10), and the DNA was extracted as per previous methods [24, 25] using a ReliaPrep gDNA Tissue Miniprep System (Promega).

Library construction and next-generation sequencing

A total of 7 ng of extracted genomic DNA (from the faecal materials) was used to prepare the library using the protocol adapted previously using the Illumina DNA Prep (Illumina) for Illumina short-read sequencing [26]. The quality and quantity of the prepared library were assessed using an Agilent Tape Station (Agilent Technologies) by the Australian Genome Research Facility (AGRF) (Parkville, Victoria, Australia). The prepared library was sequenced at the AGRF, producing 2×150 bp paired-end reads using the NovaSeq 6000 platform, according to the manufacturer's instructions. Subsequently, an attempt was made to prepare a long-read library using the SQK-LSK109 ligation sequencing kit with native barcoding expansion EXP-NBD104 (Oxford Nanopore Technologies). Regrettably, we encountered difficulties in generating a successful library due to the limited quantity of input DNA available. The protocol was further modified to accommodate a low gDNA input (approximately 53.0 ng) using half reaction volumes as per the stated protocol [27]. The final DNA yield in the prepared long-read library was too low to detect when using a Qubit 3.0 Fluorometer (Thermo Fisher Scientific), and thus analyses were not progressed.

Quality trimming and *de novo* assembly of the sequence read data

Raw reads were checked for quality using FastQC v0.11.9 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>, accessed on 10 March 2023). Raw sequence reads were *de novo* assembled using Shovill v1.1.0 (<https://github.com/tseemann/shovill>, accessed on 10 March 2023), which utilizes: Seqtk v1.3-r106 (<https://github.com/lh3/seqtk>, accessed on 10 March 2023); Trimmomatic v0.36 [28]; Lighter v1.1.2 [29]; FLASH v1.2.11 [30]; SKESA v2.4.0 [31, 32]; Samclip v0.4.0 (<https://github.com/tseemann/samclip>, accessed on 10 March 2023); SAMtools v1.16.1 [33], Burrows-Wheeler Aligner (BWA) v0.7.17 [34]; and Pilon v1.24 [35]. Shovill was used with parameters set to: (i) estimate the genome size to 5 Mb; (ii) remove contiguous sequences (contigs) with a sequence coverage below 20-fold; and (iii) enable single-cell mode. Assembly metrics were assessed using QUAST v5.0.2 [36]. A genome sequencing coverage plot was created with the script 'wgscovplotter.jar' from the Java utilities for Bioinformatics (JVARKIT) (<https://github.com/lindenb/jvarkit>).

In silico species identification and antibiotic resistance genotyping

To perform taxonomic profiling and identify *P. stewartii* by classifying the trimmed paired-end DNA sequences, we used Kraken v2.1.3 [37] with default parameters and an NCBI Reference Sequence (RefSeq) database [38], PlusPFP (<https://benlangmead.github.io/aws-indexes/k2>). The database contained references for archaea, bacteria, human, viruses, plasmids, protozoa, fungi, plant, and the 'UniVec core' subset of the UniVec database (a database of vector, adaptor, linker and primer sequences). Full taxonomic profiles of paired-end reads can be found in Table S1. ABRicate v1.0.1 (<https://github.com/tseemann/abricate>, accessed on 13 March 2023) was used to screen the draft assembly for acquired antibiotic resistance genes using the ARG-ANNOT [39] database (last updated 1 November 2022). The genome sequence for a putative plasmid (named pC10109_Jinnung) was annotated using RAST v2.0 [40–42]. The genome map for pC10109_Jinnung was visualized with MacVector v16.0 (<https://macvector.com>).

Taxonomic classification of the C10109_Jinnung genome

To identify the nearest taxonomic neighbours to the genome of C10109_Jinnung, the draft assembly was input into a Genome Taxonomy Database Toolkit (GTDB-Tk) genome-based taxonomy (GTDB-Tk v2.1.1 with GTDB package R207_v2 [43, 44]). Only sequences within the genus *Pantoea* (g_ genus) were extracted from the GTDB-Tk and used to construct a concatenated reference alignment of 119 bacterial marker genes. The taxonomic tree was inferred using maximum-likelihood approximation with FastTree v2.1.7 [45] under the WAG model [46] of protein evolution with gamma-distributed rate heterogeneity [47] (+GAMMA).

Variant detection and phylogenetic analyses

To estimate the number of SNVs and phylogenetic relationships between the genome of C10109_Jinnung and other *P. stewartii* genomes sequenced elsewhere, these were compared to a global and publicly available dataset (Table S2). Parsnp v1.7.4 [48] was used to generate a core-genome alignment (alignment of the syntenic regions across all genomes). The resulting SNV alignments were used to reconstruct phylogenies. Maximum-parsimony trees were reconstructed using the heuristic search feature of PAUP v4.0a [49], before adding 1000 bootstrap replicates. The resulting phylogenetic trees were visualized using FigTree v1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>, accessed on 10 March 2023).

Metagenomic analysis

Metagenomic data assemblies were produced using SPAdes v3.15.5 (with the --meta flag) [50], and the identification of taxa based on paired-end reads and contigs was performed using Kraken v2.1.3 and the aforementioned PlusPFP database. Additionally, identified contigs were extracted using KrakenTools [51] and subjected to further investigation through alignment using BWA v0.7.15 and SAMtools v1.15.1. For metagenomic taxonomy investigation and analysis, R v4.0.5 [52] was utilized along with the phyloseq v1.34.0 [53] and Pavian v1.2.0 [54] packages. Trimmed reads were also aligned using BWA to the C10109_Jinnung assembly to further quantify their respective abundance, using mapQ>30 counts produced with SAMtools. Full taxonomic profiles of assembled reads can be found in Table S3.

RESULTS

The genome of *P. stewartii* strain C10109_Jinnung

The clinical metadata for the sample sequenced in this study (C10109_Jinnung) are listed in Table 1. Whole-genome sequencing resulted in the draft genome sequence of C10109_Jinnung, which comprised 4645191 bp (split over 20 contigs) with an average GC content of 53.72%. Mapping the short reads from C10109_Jinnung to its draft genome shows a consistent and high median coverage exceeding 4500-fold, with a uniform distribution across the genome (i.e. a thorough representation of the genetic information in the genome of C10109_Jinnung) (Fig. S2). Table 1 also summarizes the sequence read data quality metrics for C10109_Jinnung. GTDB-Tk was used to initially classify the C10109_Jinnung strain with the nearest taxonomic neighbours within the *P. stewartii* population and verify the taxonomic assignment. Genome-wide taxonomic analysis with GTDB-Tk revealed that the C10109_Jinnung strain clusters with the draft chromosome of *P. stewartii* strain CCUG 26359 (GenBank: GCA_008801695) (Fig. 1). Fig. 1 represents a clustering analysis based on a concatenated alignment of 119 conserved genes and is therefore not a 'go-to' phylogenetic tree for focused clustering based on SNVs. While branches appear short in Fig. 1 they actually represent a large difference.

Despite being a draft genome, the chromosome of C10109_Jinnung displayed characteristics commonly found in both *Enterobacteriales* and plant-associated bacteria. Nucleotide–nucleotide BLAST (BLASTn) comparisons between C10109_Jinnung, the reference chromosome for strain ZJ-FGZX1 (GenBank: CP049115) and other complete *P. stewartii* genomes [DC283 (GenBank: CP017581) and HR3-48 (GenBank: CP099540)] revealed the presence of various genetic elements (Fig. 2). These included the cellulose biosynthesis gene cluster, enterobacterial common antigen, a type VI secretion system, flagella and the colanic acid biosynthesis gene cluster.

Table 1. Genome metrics and descriptions for *Pantoea stewartii* strain C10109_Jinnung

Strain characteristics	
Collection year	2021
Location	Perth, Australia
Source	<i>Pezoporus flaviventris</i>
Clinical presentation	Soft tissue injury
Anatomical site	Faecal sample
Sequence data quality metrics	
<i>Raw data</i>	
Average read length	150 bp
Total reads (pairs)	93033611
<i>Quality trimming</i>	
Average read length (range)	148 bp (36–150 bp)
Total reads (pairs) survived trimming	81078492 (87.1%)
<i>Taxonomic classification (PlusPFP database)*</i>	
Number of unclassified reads	3854321 (4.8%)
Number of reads species 1	73662008 (90.9%), <i>Pantoea stewartii</i>
Number of reads species 2	849732 (1.1%), <i>Pantoea ananatis</i>
Number of reads species 3	196462 (0.2%), <i>Homo sapiens</i>
Genomic description	
Average genome coverage	~ 4675x
Chromosome length	4645191 bp
Number of contigs	20
Number of contigs >1000 bp	18
Largest contig	1636662 bp
GC content	53.72%
N50	366162 bp
Number of reads mapped to assembly (mapQ>30)	77463154

*Full taxonomic profiles of paired-end reads can be found in Table S1.

Additionally, BLASTn analysis of the draft genome of C10109_Jinnung revealed multiple regions with similarities (>70% nucleotide identity) to genes associated with virulence relative to the complete chromosome of ZJ-FGZX1. These regions contain genes encoding: (i) the flagellum, with the extracellular components secreted by a type III secretion system, which contribute to bacterial motility; (ii) the O-antigen, a component of lipopolysaccharide linked to immune evasion; (iii) chromobacterium haem utilization and aerobactin operon, crucial for acquiring iron; and (iv) a P fimbriae operon involved in bacterial adhesion. Among the seven prophages in ZJ-FGZX1, prophage 6 was the sole prophage found to be shared with the draft genome of C10109_Jinnung (Fig. 2). In the complete chromosome of ZJ-FGZX1, prophage 6 has integrated into the tRNA-*metH* site. ZJ-FGZX1 prophage 6 comprises a total of 33 hypothetical protein-coding sequences (CDS), along with two other CDSs identified as a phage tail sheath protein (locus tag: G5574_10520) and a recombinase family protein (locus tag: G5574_10530). Caution should be exercised when interpreting the presence or similarity of ZJ-FGZX1 prophage 6 with regions in C10109_Jinnung, because the corresponding region in C10109_Jinnung is fragmented across three contigs of varying sizes: 356933 bp (GenBank: JASCYM010000004), 1849 bp (GenBank: JASCYM010000005) and 181569 bp (GenBank: JASCYM010000006).

The genome of C10109_Jinnung is genomically distinct from other *P. stewartii* genomes

In this study, the genomic diversity of the draft genome sequence of C10109_Jinnung, obtained from an Australian Critically Endangered psittacine, was assessed by comparing it with 29 draft and three complete publicly available *P. stewartii* genomes

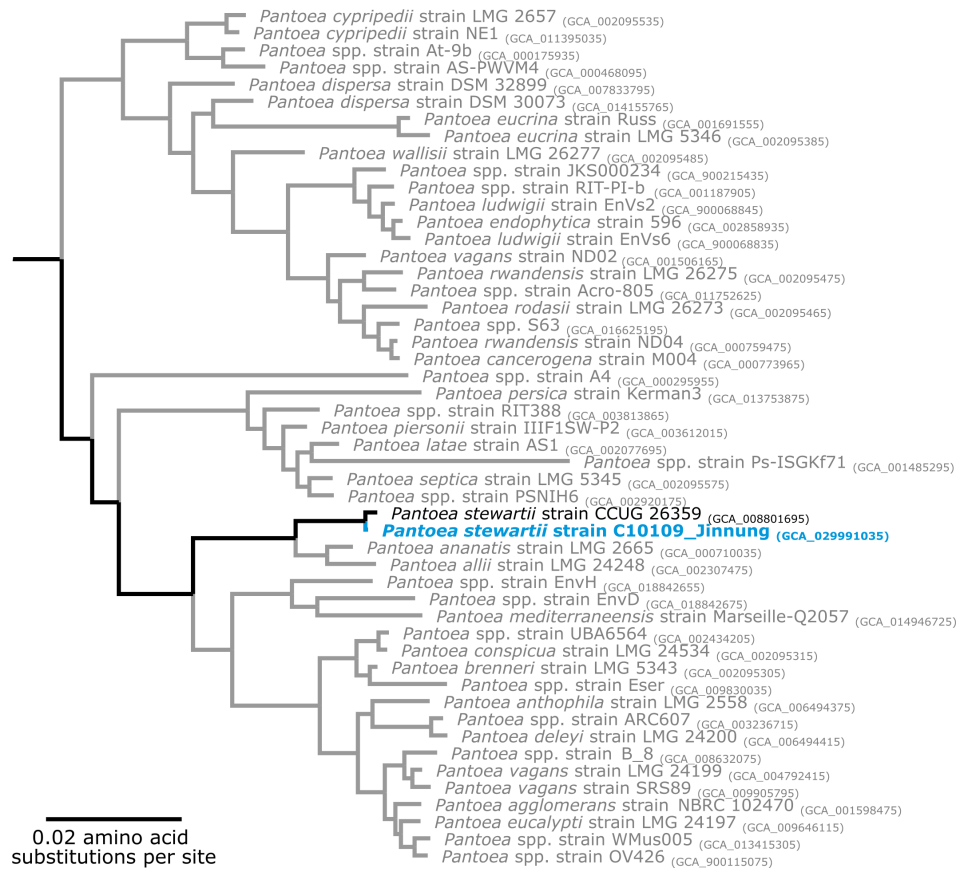


Fig. 1. Taxonomic tree of the genus *Pantoea*. The maximum-likelihood approximation was reconstructed using a concatenated alignment of 119 conserved bacterial markers. Taxonomy is shown according to the GTDB (s., species) from NCBI taxonomy. GenBank assembly accession numbers are displayed after the genome names. The blue taxon represents the strain sequenced in this report. The tree is rooted according to the outgroup genus *Erwinia*, which has been omitted for visualization.

(Fig. 3). Excluding the newly sequenced C10109_Jinnung obtained from an Australian Critically Endangered psittacine, the dataset comprises genomes of *P. stewartii* strains that have been identified from various sources, namely food ($n=26$), plants ($n=3$) and the environment ($n=2$), with one genome having an unknown source. Furthermore, the dataset consisted of *P. stewartii* genomes that were obtained from different geographical locations, including the USA ($n=14$), India ($n=10$), Malaysia ($n=3$), China ($n=2$) and Costa Rica ($n=2$), with two genomes having an unknown origin. The genomes in this collection were obtained over a period ranging from 1956 to 2020.

The alignment of 100720 core-genome SNVs from 33 *P. stewartii* genomes was used to construct a maximum-parsimony phylogenetic tree (Fig. 3). However, the low consistency index of 0.44 indicates that there could be incongruence or homoplasy in the dataset, potentially affecting the accuracy of the inferred relationships. To address this, 1000 bootstrapped replicates were generated to increase confidence in the inferred relationships. Although the deeper branching patterns in the *P. stewartii* phylogeny remain more uncertain, the higher bootstrap values for more recent nodes suggest greater reliability in the inferred relationships between more 'closely' related taxa. However, the results of the phylogenetic analysis revealed that the 33 genomes exhibited a remarkable genomic diversity within the *P. stewartii* phylogeny. For example, the terminal branch leading to strain C10109_Jinnung is defined by 8226 SNVs, as determined by pairwise SNV distances, relative to the reference chromosome ZJ-FGZX1. Additionally, C10109_Jinnung was most genomically 'similar' to strain ZJ-FGZX1, which was isolated from lucky bamboo (*Dracaena sanderiana*) in China in 2017 [16]. However, this should be interpreted cautiously because C10109_Jinnung was separated from ZJ-FGZX1 by a pairwise SNV distance of 18 697 SNVs.

While C10109_Jinnung shares the same branch with strain ZJ-FGZX1 (subsp. *indologenes*), this does not necessarily contradict the possibility of it representing subsp. *stewartii*. To address this, we conducted an *in silico* analysis of the *galE* (encoding UDP-glucose 4-epimerase) and *recA* (encoding protein RecA) genes in the sequence of C10109_Jinnung. We then compared these genes with those of subsp. *stewartii* (DC283) and subsp. *indologenes* (ZJ-FGZX1) (Fig. S3). Based on *in silico* PCR and sequence alignments of the *recA* and *galE* genes from the complete DC283 and ZJ-FGZX1 genomes, it appears that C10109_Jinnung shares

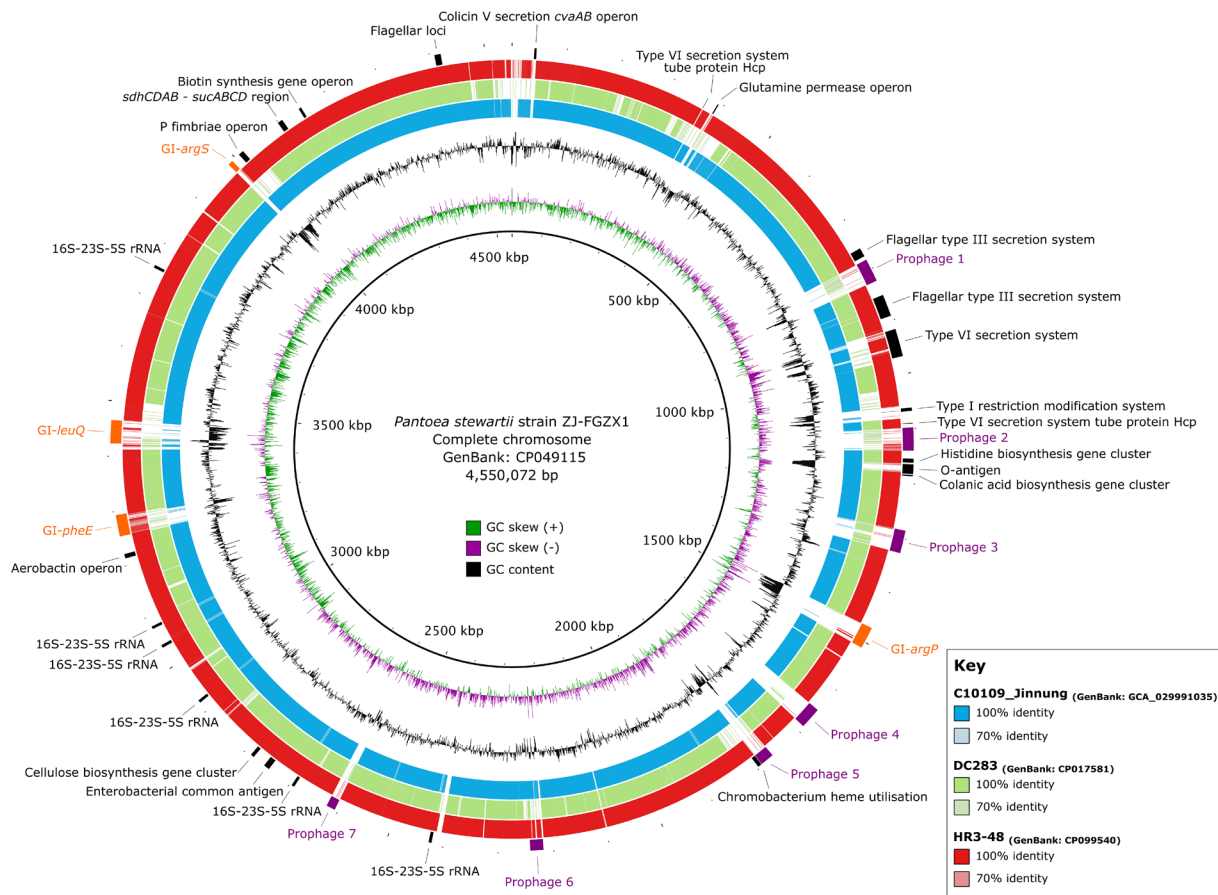


Fig. 2. Genomic comparisons between *Pantoea stewartii*. Circular representation of the chromosome of *P. stewartii* subsp. *indologenes* strain ZJ-FGZX1 (GenBank: CP049115). The innermost rings represent the position in the genome in base pairs, GC skew and GC content. Rings 4–6 represent nucleotide identity between sequences according to BLASTn (70–100%) between the draft genome of C10109_Jinnung (GenBank: GCA_029991035) and the complete genomes *P. stewartii* subsp. *stewartii* DC283 (GenBank: CP017581) and *P. stewartii* HR3-48 (GenBank: CP099540), as shown in the key. The outermost ring highlights regions of interest (black), genomic islands (orange) and prophages (purple). Image created using BRIG v0.95 [72].

the same SNVs with ZJ-FGZX1 in most cases. However, one exception is an SNV C246T in the *recA* gene compared to ZJ-FGZX1 (matching subsp. *stewartii* DC283). This evidence suggests that C10109_Jinnung is probably a member of subsp. *indologenes*.

Metagenomic diet

Taxonomic analysis using Kraken was conducted to investigate the composition in the C10109_Jinnung faecal sample. Analysis revealed a diverse range of taxa from part of the parrot's diet. Metagenomic analysis revealed that the plant genera detected comprised mainly sunflowers (*Helianthus* spp.), panicgrass (*Panicum* spp.), ryegrass (*Lolium* spp.), eucalyptus (*Eucalyptus* spp.), *Rhodamnia* spp., foxtail millet (*Setaria* spp.) and wheat (*Triticum* spp.). The species designation for the contigs were the common sunflower (*Helianthus annuus*), Hall's panicgrass (*Panicum hallii*), switchgrass (*Panicum virgatum*), flooded gum tree (*Eucalyptus grandis*), white myrtle (*Rhodamnia argentea*), millet (*Setaria italica*), annual ryegrass (*Lolium rigidum*) and common wheat (*Triticum aestivum*) (Fig. 4).

Identification of a putative plasmid with cloning vector features

During the *de novo* assembly process of the sequence data for C10109_Jinnung, a putative plasmid with a length of 5839 bp was identified (named pC10109_Jinnung). This putative plasmid was sequenced to a coverage of approximately 1340-fold, indicating a potentially substantial presence within the faecal sample. The significance of this finding is currently unclear, and contamination cannot currently be ruled out. The sequence of the pC10109_Jinnung plasmid displays typical characteristics of a cloning vector, including T7 promoters and a multi-cloning site featuring seven restriction enzymes (Fig. S4). Detailed analysis and supplementary information regarding the putative plasmid pC10109_Jinnung, encompassing gene annotations, sequence similarities and distinctive features, can be found in the online supporting materials.

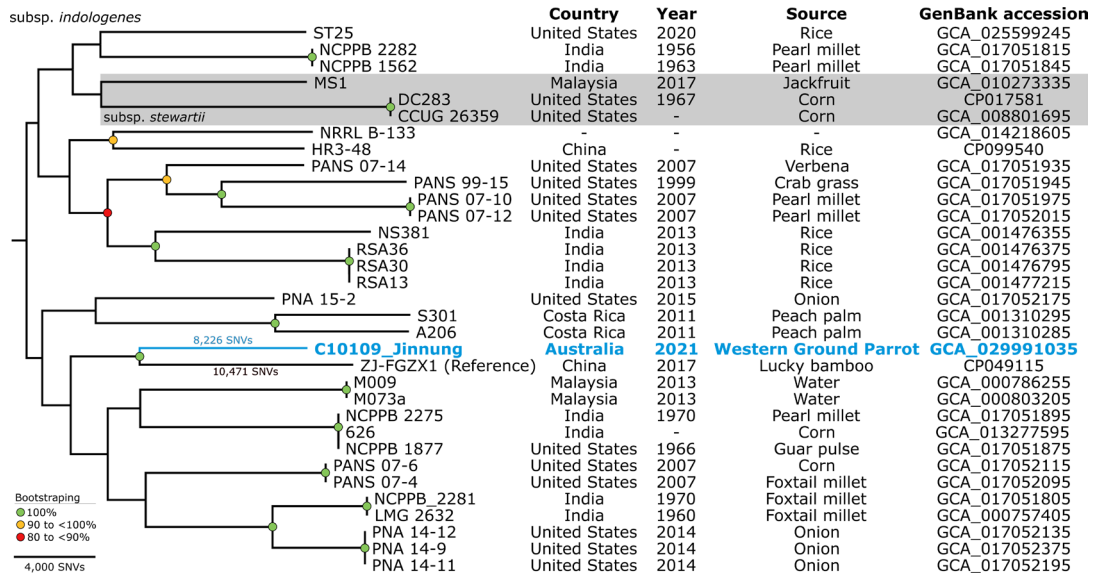


Fig. 3. Maximum-parsimony phylogeny of *Pantoea stewartii*. The phylogeny was inferred from 100720 core-genome single-nucleotide variants (SNVs) from 33 genomes and rooted at the midpoint. SNVs were derived from a core-genome alignment of 3433126bp and are called against the reference chromosome of *P. stewartii* subsp. *indologenes* strain ZJ-FGZX1 (GenBank: CP049115). The blue taxon represents the strain sequenced in this report. Bootstrap values (using 1000 replicates) are shown. The consistency index is 0.44.

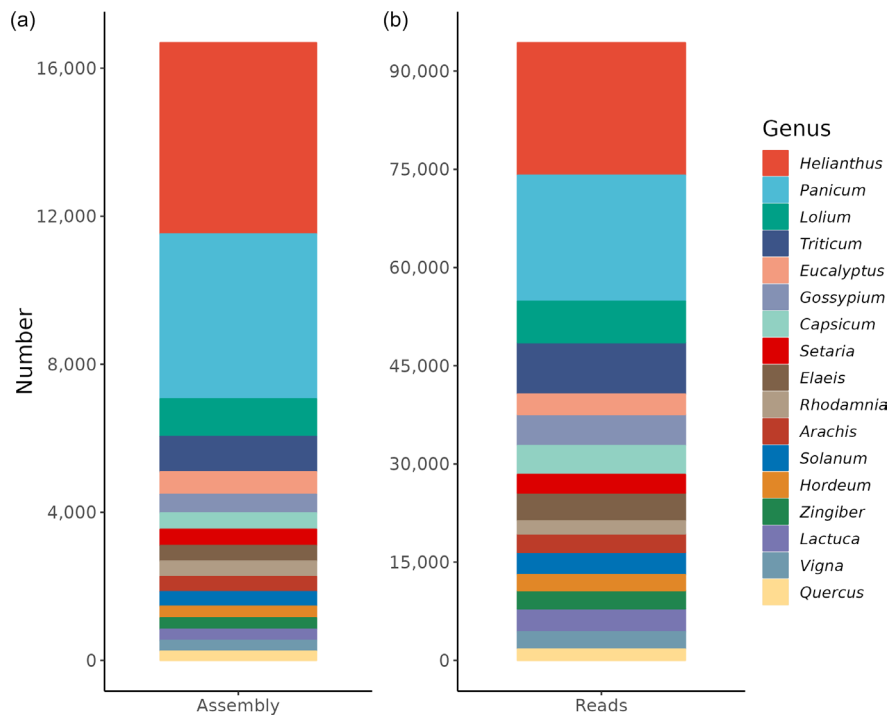


Fig. 4. Potential *Pantoea stewartii* host genera. (a) Number of contigs in meta-assembly; (b) number of unassembled reads. Genera with over 1% abundance are shown.

DISCUSSION

There have been no reported instances of *P. stewartii* strains being found from a bird source (to our knowledge). Despite the substantial economic losses incurred due to the impact of *P. stewartii* [5–7], with many countries classifying it as a quarantine organism [9], more global genomic data are needed for this bacterium. This study introduces a draft genome of a *P. stewartii* strain (C10109_Jinnung) obtained from Australia's Critically Endangered western ground parrot/kyloring. It is important to clarify that our study is based on an exceptional occurrence that came about unexpectedly through our collaboration with Perth Zoo, focusing on a translocation project for the endangered western ground parrot/kyloring population in Western Australia. The bird in question, unfit for release into the wild, was sent to Perth Zoo for evaluation. This research does not encompass the relationship between *P. stewartii* and western ground parrots/kylorings across Western Australia. Rather, we present the first documented instance of *P. stewartii* in a bird source, supported by genomic analysis.

Genome analysis revealed the presence of multiple genetic elements within the chromosome, some associated with virulence. These included the type VI secretion system (cellular aggression mechanism [55]), flagella (mobility [56]), chromobacterium haem utilization and aerobactin operons (iron acquisition [57, 58]), as well as a P fimbriae operon (bacterial adhesion [59]). Unfortunately, there was insufficient DNA to perform long-read sequencing [60], and benefit from the improvements in draft assembly and resolving genomic complexity that this would have provided. We are exploring using Oxford Nanopore Technologies (long-read sequencing) with additional samples. Although the currently available genomes of *P. stewartii* display genetic diversity in the core genome, they present an underrepresented population. Moreover, this study shows that the genome of C10109_Jinnung is genomically distinct from other *P. stewartii* genomes.

Insights from phylogenomic analysis and the call for broader genomic exploration

Previous studies have explored potential insect vectors of *P. stewartii*, but only in controlled environments [10–12]. The corn flea beetle stands out as an efficient vector due to its behaviour of depositing faecal matter directly on the leaf wounds it creates while feeding on the plant [13]. Therefore, unlike the corn flea beetle, the deposition of bird faecal matter on plants seems unlikely to cause infection in healthy plants unless there are pre-existing wounds. This is an important observation, in particular concerning the biosecurity risks associated with *P. stewartii*, but must be viewed in the context of the limitation that this study only investigated one genome from a *P. stewartii* strain obtained from a western ground parrot/kyloring. Although we acknowledge the importance of a larger sample size, practical challenges hindered our efforts. Restricted access to critically endangered parrot species in Australia and financial constraints prevented us from obtaining additional samples from various locations. Despite these challenges, publishing our findings establishes a groundwork for future sampling endeavours and funding requests. We eagerly encourage and anticipate that upcoming research will build upon our discoveries, fostering further exploration in this field. Therefore, additional research is necessary to investigate the role of pre-existing wounds in the transmission of plant infections through bird faecal matter and to sequence more *P. stewartii* genomes.

The true weight of our research emerged when we conducted phylogenetic analysis, revealing the evolutionary relationships and genetic context of C10109_Jinnung, which is probably classified as *P. stewartii* subsp. *indologenes* (Fig. 3). This study represents the most extensive phylogenomic investigation into *P. stewartii* (to our knowledge), encompassing a comprehensive framework with 32 additional genomes from diverse sources and locations. However, it is worth noting that while insightful, these 33 genomes offer a limited representation of the broader *P. stewartii* population. Nevertheless, this study has led to a better understanding of the genetic diversity and evolutionary dynamics within the species *P. stewartii*. Additionally, utilizing genomic DNA extracted directly from faecal material, rather than relying on cultured isolates, demonstrated the potential of culture-independent whole-genome sequencing for genotypic *P. stewartii* subspecies typing. This approach bypasses the limitations of traditional culturing methods and allows for the direct analysis of environmental samples. Future studies may also benefit from incorporating phenotypic strain culturing and biochemical tests for definitive subspecies identification [19]. These complementary techniques enable the isolation and characterization of specific strains, providing additional phenotypic information to enhance subspecies identification efforts.

This study's core-genome SNV-based phylogenetic analysis revealed that C10109_Jinnung shared a most recent common ancestor with strain ZJ-FGZX1 (also subsp. *indologenes*), isolated from lucky bamboo in China in 2017 [16]. However, caution must be exercised when interpreting this finding, as C10109_Jinnung and ZJ-FGZX1 exhibit a substantial genetic divergence with an 18697 pairwise SNV distance. The genetic variation and independent changes observed between C10109_Jinnung and ZJ-FGZX1 show substantial gaps in available *P. stewartii* genomes. Moreover, detecting *P. stewartii* in a bird host (albeit a single strain in a single host) prompts further investigation into its prevalence among other bird species in Western Australia. If additional instances of *P. stewartii* are discovered in Western Australia, replicating the current study would provide insights into the similarities or differences between any new findings and C10109_Jinnung. A broader analysis incorporating a more comprehensive range of *P. stewartii* strains and more extensive genomic datasets will provide a deeper understanding of the evolutionary relationships and population dynamics of *P. stewartii*.

C10109_Jinnung's virulence toolkit: flagella, fimbriae and iron acquisition systems

The availability of complete reference genomes played a crucial role in our study, enabling the identification of virulence factors in the draft genome of C10109_Jinnung, including flagella, fimbriae and siderophore/iron acquisition systems. These characteristics are not exclusive to C10109_Jinnung and are commonly shared among other *P. stewartii* genomes (Fig. 2). *P. stewartii* demonstrates impressive agar surface motility facilitated by flagella-dependent movement [61], and comparative analysis of complete genomes suggests the conservation of flagella-encoding genes among genomically distinct different *P. stewartii* strains (Fig. 2). The presence of P fimbriae genes in *P. stewartii* also raises intriguing possibilities for virulence, where P fimbriae can promote adherence to host tissues, promote colonization, and potentially contribute to urinary tract infections [62]. Moreover, genes encoding siderophore/iron acquisition systems in *P. stewartii* indicate an ability to compete for iron resources [63]. Consequently, the expression of flagella, P fimbriae and iron acquisition systems in *P. stewartii* may suggest a role in facilitating the adhesion to, and colonization of, host tissues, potentially enhancing its pathogenic potential and capacity to initiate infections and endure over time [64]. However, it is essential to acknowledge that these observations are based on a limited dataset, as only three complete genomes are currently accessible in the public domain [16–18].

What about the expression vector?

Plasmids play a crucial role in molecular cloning, and various cloning vectors have been developed to facilitate the identification of plasmids with DNA inserts [65]. The Tet repressor (encoded by *tetR*) and specific aminophosphotransferases, such as *aph(3')*-Ia and *aph(3')*-IIa, have been widely used as selectable markers for tetracycline resistance [66–68] and kanamycin resistance in prokaryotic vectors, respectively [65]. The presence of pC10109_Jinnung, which shares similarities with an *Escherichia coli* expression vector and contains common *E. coli* protein expression vector sequences, suggests the potential for horizontal gene transfer [69]. Further investigation is needed to determine the origin of the possible horizontal gene transfer leading to the presence of the expression vector in the genome sequence data for C10109_Jinnung, whether it occurred naturally or because of laboratory contamination during experimental procedures. Furthermore, the high sequence coverage of pC10109_Jinnung (approximately 1340-fold) indicates a likely high abundance in the faecal sample if it is not due to contamination.

Insights into the potential role of *P. stewartii* in seed-feeding bird-mediated transmission and implications for host expansion

Taxonomic analyses unveiled the intricate composition of the C10109_Jinnung faecal sample extending beyond the identified *P. stewartii*. We identified multiple matches to plant species that form part of the parrot's diet, which could mainly reflect the diet while in the zoo. Nevertheless, employing this analysis on non-captured (wild) birds has the potential to yield invaluable insights regarding their geographical distribution and dietary patterns. *P. stewartii* can exploit a diverse range of plant and insect hosts, utilizing the transportation of host plants (including seeds) and insect vectors from neighbouring countries [70]. This, alongside our observations, suggests that seed-feeding birds might act as reservoirs and facilitate the spread of the bacteria, potentially leading to host shifts. Moreover, the crops and guts of birds could serve as a reservoir for transmitting *P. stewartii* to previously uninfected species. Additionally, avian dispersal could contribute to spreading contaminated seeds, increasing the risk of pathogenic host expansion.

Of the eight plant genera with the most contigs, two have already been identified as hosts for *P. stewartii* (wheat and millet, further supporting the identification of C10109_Jinnung as representing *P. stewartii* subsp. *indologenes*) [22], while others may also become susceptible to infection. For instance, flooded gum trees (*Eucalyptus grandis*) are a significant forestry crop, and the closely related *Erwinia psidii* has been found to have undergone a host shift to infect Myrtaceae and *Eucalyptus* in South America [71]. The metagenomic diet analyses provides a potential snapshot of the western ground parrot/kyloring diet, but it is not comprehensive and may not pick up everything consumed. For instance, while the western ground parrot/kyloring was provided with corn (Fig. S1), none was detected in the metagenomics diet analysis (Fig. 4). Furthermore, the substantial implications of *P. stewartii* host expansion facilitated by birds on food security and economic productivity cannot be overstated.

CONCLUSION

This study highlights the need for more comprehensive genomic data on *P. stewartii*, a bacterium causing significant economic losses and classified as a quarantine organism. While no previous reports exist of *P. stewartii* strains isolated from a bird source, the draft genome analysis of C10109_Jinnung from one of the Critically Endangered Australian parrots reveals genetic elements associated with virulence, including cellular aggression, mobility, iron acquisition and bacterial adhesion. However, it is important to acknowledge the limited representation of *P. stewartii* genomes in the current dataset, potentially missing intermediate evolutionary lineages. Future research should investigate the specific genetic changes and mechanisms driving this divergence, explore functional implications of genetic variations, and expand the analysis to encompass a broader range of *P. stewartii* strains and larger genomic datasets. Incorporating diverse strains from various sources and locations would enable a comprehensive exploration of the genomic landscape of *P. stewartii*, facilitating insights into the epidemiology, host interactions, and the potential

for antibiotic resistance acquisition and dissemination. Additionally, identifying plant species in the C10109_Jinnung sample suggests potential hosts or carriers for *P. stewartii* spread, necessitating investigation into bird-mediated transmission and its implications for plant host expansion and food security.

Funding information

This research received no external funding.

Acknowledgements

We express our sincere gratitude to the Perth Zoo and Parks, and the Wildlife Western Ground Parrot team, specifically Arthur Ferguson and Allan Burbidge, for their invaluable support, facilities, and scientific and technical expertise. We also recognize the facilities and the scientific and technical assistance of the Australian Genome Research Facility staff. We thank Alison Mackie (Department of Primary Industries and Regional Development, Government of Western Australia), Vinko Besic, Xiaoyun Ren and Donia Macartney-Coxson (Institute of Environmental Science and Research, Wellington) for their valuable feedback. Subir Sarker is a recipient of the Australian Research Council Discovery Early Career Researcher Award (Grant Number DE200100367) funded by the Australian Government. The Australian Government had no role in study design, data collection and analysis, the decision to publish, or the preparation of the manuscript.

Author contributions

Conceptualization: RTW and SS; investigation: RTW, WT, EW and SS; funding acquisition: SS; formal analysis: RTW, WT, EW and SS; wet-lab experiments: RV-H, NK, JJAR and SS; data analysis: RTW, WT, EW and SS; data curation: RTW; Illumina sequencing was done at the Australian Genome Research Facility (Parkville, Australia); supervision: SS; writing (original draft preparation): RTW, WT, EW and SS; writing (review and editing): RTW, RV-H, DP, WT, EW, NK, SP, JJAR and SS. All authors have read and agreed to the published version of the manuscript.

Conflicts of interest

The authors declare that there are no conflicts of interest

Ethical statement

Animal ethics approval for capturing wild western ground parrot/kylloring (*Pezoporus flaviventris*) in 2021 was obtained from the Western Australian Department of Biodiversity, Conservation and Attractions Animal Ethics Committee. The project, titled 'Western Ground Parrot Recovery' (Approval No.: 2018-11A), aimed to contribute to the conservation of this Critically Endangered species. The faecal sampling was conducted during routine animal husbandry practice without handling animals. The Animal Ethics Committee at La Trobe University was informed that findings from the material were to be used in a publication, and a formal waiver of ethics approval was granted.

References

- Gapasin RM, Garcia RP, Advincula CT, Cruz CS, Borines LM. Fruit bronzing, a new disease affecting jackfruit caused by (Smith) mergaert *Pantoea stewartii* et al. *Ann Trop Res* 2014;36:17–31.
- Kini K, Agnimonhan R, Afolabi O, Soglonou B, Silué D, et al. First report of a new bacterial leaf blight of rice caused by *Pantoea ananatis* and *Pantoea stewartii* in Togo. *Plant Dis* 2017;101:241.
- Azad HR, Holmes GJ, Cooksey DA. A new leaf blotch disease of sugarcane caused by *Pantoea ananas* and *Pantoea stewartii*. *Plant Dis* 2000;84:973–979.
- Braun EJ. Ultrastructural investigation of resistant and susceptible maize inbreds infected with *Erwinia stewartii*. *Phytopathology* 1982;72:159–166.
- Anderson TR. An outbreak of Stewart's bacterial wilt of corn in Ontario, Canada. *Plant Dis* 1986;70:603.
- Dillard HR. An outbreak of Stewart's bacterial wilt of corn in New York state. *Plant Dis* 1989;73:273.
- Roper MC. *Pantoea stewartii* subsp. *stewartii*: lessons learned from a xylem-dwelling pathogen of sweet corn. *Mol Plant Pathol* 2011;12:628–637.
- Kuhar TP, Stivers-Young LJ, Hoffmann MP, Taylor AG. Control of corn flea beetle and Stewart's wilt in sweet corn with imidacloprid and thiamethoxam seed treatments. *Crop Protection* 2002;21:25–31.
- Jerald P, Robert I. Pest Risk Analysis. In: *The Risk of Introducing Erwinia Stewartii in Maize Seed*. Switzerland: ISF Secretariat, 2003.
- Ammar E-D, Correa VR, Hogenhout SA, Redinbaugh MG. Immunofluorescence localization and ultrastructure of Stewart's wilt disease bacterium *Pantoea stewartii* in maize leaves and in its flea beetle vector *Chaetocnema pulicaria* (Coleoptera: Chrysomelidae). *J Microsc Ultrastruct* 2014;2:28–33.
- Stavrinides J, No A, Ochman H. A single genetic locus in the phytopathogen *Pantoea stewartii* enables gut colonization and pathogenicity in an insect host. *Environ Microbiol* 2010;12:147–155.
- Terenius O, Lindh JM, Eriksson-Gonzales K, Bussi ere L, Laugen AT, et al. Midgut bacterial dynamics in *Aedes aegypti*. *FEMS Microbiol Ecol* 2012;80:556–565.
- Pepper E. *Stewart's Bacterial Wilt of Corn*. St Paul, MN: The American Phytopathological Society; 1967.
- Claflin LE. Diseases caused by prokaryotes. In: Munkvold GP and White DG (eds). *Compendium of Corn Diseases*, 3rd edn. Minnesota, United States: The American Phytopathological Society; 1999. pp. 3–9.
- Menelas B, Block CC, Esker PD, Nutter FW. Quantifying the feeding periods required by corn flea beetles to acquire and transmit *Pantoea stewartii*. *Plant Dis* 2006;90:319–324.
- Zhang S, Xu Z, Le R, Hu H. First report of leaf blight wilt on *Dracaena Sanderiana* by *Pantoea Stewartii* subsp. *Indologenes* in China. *Plant Dis* 2020;104:1854.
- Duong DA, Stevens AM, Jensen RV. Complete genome assembly of *Pantoea stewartii* subsp. *stewartii* DC283, a corn pathogen. *Genome Announc* 2017;5.
- Zhang X, Ma Y-N, Wang X, Liao K, He S, et al. Dynamics of rice microbiomes reveal core vertically transmitted seed endophytes. *Microbiome* 2022;10:216.
- Mergaert J, Verdonck L, Kersters K. Transfer of *Erwinia ananas* (synonym, *Erwinia Uredovora*) and *Erwinia stewartii* to the genus *Pantoea emend. As Pantoea ananas* (Serrano 1928) comb. nov and *Pantoea stewartii* (Smith 1898) comb nov, respectively, and description of *Pantoea stewartii* subsp. *indologenes* subsp. *Int J Syst Evol Microbiol* 1993;43:162–173.
- Wensing A, Zimmermann S, Geider K. Identification of the corn pathogen *Pantoea stewartii* by mass spectrometry of whole-cell extracts and its detection with novel PCR primers. *Appl Environ Microbiol* 2010;76:6248–6256.
- Gehring I, Wensing A, Gernold M, Wiedemann W, Coplin DL, et al. Molecular differentiation of *Pantoea stewartii* subsp. *indologenes* from subspecies *stewartii* and identification of new isolates from maize seeds. *J Appl Microbiol* 2014;116:1553–1562.

22. Koirala S, Zhao M, Agarwal G, Gitaitis R, Stice S, et al. Identification of two novel Pathovars of *Pantoea stewartii* subsp. *Indologenes* affecting *Allium* sp. and millets. *Phytopathology*® 2021;111:1509–1519.
23. Nadarasah G, Stavriniades J. Quantitative evaluation of the host-colonizing capabilities of the enteric bacterium *Pantoea* using plant and insect hosts. *Microbiology* 2014;160:602–615.
24. Sarker S, Isberg SR, Moran JL, Araujo RD, Elliott N, et al. Crocodilepox virus evolutionary genomics supports observed poxvirus infection dynamics on saltwater crocodile (*Crocodylus porosus*). *Viruses* 2019;11:1116.
25. Sarker S, Das S, Helbig K, Peters A, Raidal SR. Genome sequence of an Australian strain of canid alphaherpesvirus 1. *Aust Vet J* 2018;96:24–27.
26. Sarker S. Metagenomic detection and characterisation of multiple viruses in apparently healthy Australian Neophema birds. *Sci Rep* 2021;11:20915.
27. Heavens D, Choonea D, Giolai M, Cuber P, Aanstad P, et al. How low can you go? Driving down the DNA input requirements for nanopore sequencing. *Genomics* 2021. DOI: 10.1101/2021.10.15.464554.
28. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114–2120.
29. Song L, Florea L, Langmead B. Lighter: fast and memory-efficient sequencing error correction without counting. *Genome Biol* 2014;15:509.
30. Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 2011;27:2957–2963.
31. Souvorov A, Agarwala R. SAUTE: sequence assembly using target enrichment. *BMC Bioinformatics* 2021;22:375.
32. Souvorov A, Agarwala R, Lipman DJ. SKESA: strategic k-mer extension for scrupulous assemblies. *Genome Biol* 2018;19:153.
33. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. The sequence alignment/Map format and SAMtools. *Bioinformatics* 2009;25:2078–2079.
34. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 2009;25:1754–1760.
35. 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.
36. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 2013;29:1072–1075.
37. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 2014;15:R46.
38. Sayers EW, Barrett T, Benson DA, Bryant SH, Canese K, et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 2009;37:D5–D15.
39. Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, et al. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob Agents Chemother* 2014;58:212–220.
40. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, et al. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 2008;9:75.
41. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, et al. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res* 2014;42:D206–D214.
42. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, et al. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* 2015;5:8365.
43. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* 2018;36:996–1004.
44. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* 2022;38:5315–5316.
45. Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 2009;26:1641–1650.
46. Whelan S, Goldman N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol* 2001;18:691–699.
47. Yang Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J Mol Evol* 1994;39:306–314.
48. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol* 2014;15:524.
49. Wilgenbusch JC, Swofford D. Inferring evolutionary trees with PAUP*. *Curr Protoc Bioinformatics* 2003;6.
50. 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.
51. Lu J, Rincon N, Wood DE, Breitwieser FP, Pockrandt C, et al. Metagenome analysis using the Kraken software suite. *Nat Protoc* 2022;17:2815–2839.
52. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2021. <https://www.R-project.org/>
53. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013;8:e61217.
54. Breitwieser FP, Salzberg SL. Pavian: interactive analysis of metagenomics data for microbiome studies and pathogen identification. *Bioinformatics* 2020;36:1303–1304.
55. Russell AB, Hood RD, Bui NK, LeRoux M, Vollmer W, et al. Type VI secretion delivers bacteriolytic effectors to target cells. *Nature* 2011;475:343–347.
56. Berg HC, Anderson RA. Bacteria swim by rotating their flagellar filaments. *Nature* 1973;245:380–382.
57. Carbonetti NH, Boonchai S, Parry SH, Väisänen-Rhen V, Korhonen TK, et al. Aerobactin-mediated iron uptake by *Escherichia coli* isolates from human extraintestinal infections. *Infect Immun* 1986;51:966–968.
58. de Lima VM, Batista BB, da Silva Neto JF. The regulatory protein ChuP connects heme and siderophore-mediated iron acquisition systems required for *Chromobacterium violaceum* virulence. *Front Cell Infect Microbiol* 2022;12:873536.
59. Källénius G, Möllby R, Svenson SB, Helin I, Hultberg H, et al. Occurrence of P-fimbriated *Escherichia coli* in urinary tract infections. *Lancet* 1981;2:1369–1372.
60. Batinovic S, Rose JJA, Ratcliffe J, Seviour RJ, Petrovski S. Cocultivation of an ultrasmall environmental parasitic bacterium with lytic ability against bacteria associated with wastewater foams. *Nat Microbiol* 2021;6:703–711.
61. Herrera CM, Koutsoudis MD, Wang X, von Bodman SB. *Pantoea stewartii* subsp. *stewartii* exhibits surface motility, which is a critical aspect of Stewart's wilt disease development on maize. *Mol Plant Microbe Interact* 2008;21:1359–1370.
62. Buckles EL, Luterbach CL, Wang X, Lockatell CV, Johnson DE, et al. Signature-tagged mutagenesis and co-infection studies demonstrate the importance of P fimbriae in a murine model of urinary tract infection. *Pathog Dis* 2015;73:ftv014.
63. Chu BC, Garcia-Herrero A, Johanson TH, Krewulak KD, Lau CK, et al. Siderophore uptake in bacteria and the battle for iron with the host; a bird's eye view. *Biometals* 2010;23:601–611.
64. Correa VR, Majerczak DR, Ammar E-D, Merighi M, Pratt RC, et al. The bacterium *Pantoea stewartii* uses two different type III secretion systems to colonize its plant host and insect vector. *Appl Environ Microbiol* 2012;78:6327–6336.

65. Green MR, Sambrook J. Cloning and transformation with plasmid vectors. *Cold Spring Harb Protoc* 2021;2021:db.
66. Bochner BR, Huang HC, Schieven GL, Ames BN. Positive selection for loss of tetracycline resistance. *J Bacteriol* 1980;143:926–933.
67. Maloy SR, Nunn WD. Selection for loss of tetracycline resistance by *Escherichia coli*. *J Bacteriol* 1981;145:1110–1111.
68. Craine BL. Novel selection for tetracycline- or chloramphenicol-sensitive *Escherichia coli*. *J Bacteriol* 1982;151:487–490.
69. Nguyen TH, Barnes CL, Agola JP, Sherazi S, Greene LH, et al. Demonstration of horizontal gene transfer from genetically engineered *Thermosynechococcus elongatus* BP1 to wild-type *E. coli* DH5 α . *Gene* 2019;704:49–58.
70. HealthE, Jeger M, Bragard C, Candresse T, Chatzivassiliou E, et al. Pest categorisation of *Pantoea stewartii* subsp. *stewartii*. *EFSA Journal* 2018;16:e05356.
71. Coutinho TA, Brady CL, van der Vaart M, Venter SN, Telechea N, et al. A new shoot and stem disease of *Eucalyptus* species caused by *Erwinia psidii*. *Australasian Plant Pathol* 2011;40:55–60.
72. Alikhan N-F, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 2011;12:402.

Five reasons to publish your next article with a Microbiology Society journal

1. When you submit to our journals, you are supporting Society activities for your community.
2. Experience a fair, transparent process and critical, constructive review.
3. If you are at a Publish and Read institution, you'll enjoy the benefits of Open Access across our journal portfolio.
4. Author feedback says our Editors are 'thorough and fair' and 'patient and caring'.
5. Increase your reach and impact and share your research more widely.

Find out more and submit your article at microbiologyresearch.org.