BMC Genomics

RESEARCH Open Access

Whole-genome characterization and analysis of *Pantoea agglomerans* R6: a genomic insight into its pathogenicity and resistance as a potential opportunistic plant pathogen

Darin Edward Holman[1*], Ashwil Klein[2] and Marshall Keyster[1*]

Abstract

Pantoea agglomerans is a Gram-negative bacterium increasingly recognised as an opportunistic pathogen, yet the molecular basis underpinning its host-interaction capacity remains poorly understood. Here, we report the whole-genome sequencing and integrative characterisation of *P. agglomerans* strain R6, isolated from *Lactuca serriola*. The 4.7 Mb draft genome (GC content 55.6%) encodes 4,349 genes, including secretion system components, siderophore clusters, adhesins, and multidrug efflux pumps. Comparative genomic analysis against previously characterised *Pantoea* strains revealed an open pan-genome shaped by horizontal gene transfer, with multiple genomic islands harbouring putative virulence- and resistance-associated loci. Notably, homologues of type VI secretion system components, iron acquisition systems, and stress response pathways suggest adaptive potential during host colonisation. Complementary phenotypic assays supported these genomic predictions, demonstrating swarming motility, biofilm formation, extracellular polysaccharide production, and enzymatic activities associated with host interaction in related strains. While R6 displayed susceptibility to β-lactams, its genomic repertoire indicates potential for adaptive resilience under selective pressure. This integrative genomic and phenotypic characterisation identifies candidate molecular features associated with opportunistic behaviour and highlights the genomic potential of R6, rather than experimentally validated causal determinants of pathogenicity.

Keywords *Pantoea agglomerans*, Comparative genomics, Pathogen

*Correspondence:
Darin Edward Holman
3965207@myuwc.ac.za
Marshall Keyster
mkeyster@uwc.ac.za
[1]Environmental Biotechnology Laboratory, Department of Biotechnology, University of the Western Cape, Bellville 7535, South Africa
[2]Plant Omics Laboratory, Department of Biotechnology, University of the Western Cape, Bellville 7535, South Africa

Introduction

Pantoea agglomerans is a Gram-negative, facultatively anaerobic bacterium within the family Enterobacteriaceae that occurs across a broad range of habitats. It is frequently recovered from plant-associated environments, including epiphytic and endophytic niches, as well as from soil and water, and it has also been reported from animal-associated sources [1–3]. This ecological breadth reflects a species capable of thriving in multiple contexts, but it complicates interpretation in agricultural and clinical settings, as strains may exhibit either beneficial or opportunistic pathogenic behaviour depending on their genetic repertoire and environmental conditions.

Despite extensive reports describing both plant-beneficial and opportunistic pathogenic *P. agglomerans* strains, the molecular features that distinguish these lifestyles remain poorly resolved, particularly for strain R6 isolated from *Lactuca serriola.*

The genus *Pantoea*, first described by Gavini, Mergaert [2], comprises yellow-pigmented bacteria that display notable adaptability and have been detected in diverse reservoirs such as seeds, insects, and clinical environments [2, 4]. Several members of the genus are of agricultural relevance, particularly *P. agglomerans*, *P. ananatis*, and *P. stewartii*, which are associated with both plant-associated lifestyles and plant disease development.

A substantial number of *Pantoea* isolates have been described as plant growth-promoting bacteria (PGPB), supporting host performance through mechanisms that improve nutrient availability and stress tolerance. These mechanisms include nitrogen fixation, phosphate solubilisation, siderophore-mediated iron acquisition, and production of phytohormones such as indole-3-acetic acid (IAA) [5, 6]. Beyond direct growth promotion, some strains suppress phytopathogens through competitive exclusion, antimicrobial production, and stimulation of host defence responses. For example, selected non-pathogenic isolates have been applied as biological control agents against fire blight caused by *Erwinia amylovora*, demonstrating their value in integrated disease management strategies [7].

In contrast, pathogenic *Pantoea* strains have been implicated in economically important diseases across multiple crops, including rice blight, maize leaf spot, sorghum dieback, onion centre rot, cotton blight, and walnut dieback [8, 9]. In *Brassica napus*, infections can present as leaf spotting and blight symptoms, vascular dysfunction, and stem lesions, ultimately reducing plant vigour and yield. These outcomes highlight the capacity of *Pantoea* to colonise and persist in diverse plant tissues, including leaves, stems, seeds, and vascular niches, which facilitates dissemination and disease establishment.

Reflecting this duality, *P. agglomerans* is particularly notable because strains occur at both ends of the interaction spectrum, from growth-promoting and biocontrol-associated isolates to opportunistic plant and human pathogens [10, 11]. This variability is supported by genetic determinants linked to host interaction. Secretion systems such as type III and type VI mediate effector delivery, while surface-associated factors, including lipopolysaccharides, extracellular polysaccharides, and adhesins, promote attachment and immune modulation. Iron acquisition systems enable survival in nutrient-limited host niches. In addition, phytohormone-related pathways, including IAA and cytokinin biosynthesis, may alter host physiology and influence the shift between beneficial colonisation and disease under specific conditions [12–14]. These features underscore the fine boundary between mutualism and pathogenicity in this species.

Advances in microbial genomics have substantially improved our ability to investigate this diversity at the molecular level. The draft genome of *P. agglomerans* strain R6, isolated from *Lactuca serriola*, revealed more than 4,000 predicted coding sequences, including loci linked to stress tolerance, metabolic versatility, and putative virulence-associated functions [15]. Comparative analyses across pathogenic and non-pathogenic strains indicate that horizontal gene transfer, often mediated by plasmids and genomic islands, strongly shapes the evolution of *P. agglomerans*. Such genomic plasticity enhances ecological fitness, yet it also limits the ability to infer risk solely from taxonomic identity.

A clearer understanding of the genetic basis underpinning beneficial versus pathogenic traits in *P. agglomerans* is therefore essential for both disease prevention and the safe development of *Pantoea*-based applications in agriculture. To address this gap, assays were selected to evaluate traits relevant to host interaction and safety. Biofilm and EPS measurements were included to assess colonisation capacity, enzymatic profiling to identify activities associated with substrate utilisation, and antibiotic susceptibility testing to determine whether genomic resistance predictions translate into phenotype.

In this study, we characterise *P. agglomerans* R6 using an integrative approach that combines phenotypic assays with genome-based analyses to evaluate determinants of host interaction and opportunistic behaviour. Phenotypic traits assessed included swarming motility, biofilm and extracellular polysaccharide production, extracellular enzymatic activities, nutrient acquisition capacities, and antibiotic susceptibility, as these features are directly linked to colonisation ability, persistence in planta, and potential pathogenicity. These observations were interpreted alongside genomic predictions of secretion systems, adhesion factors, and iron uptake loci to provide a cohesive assessment of strain-level functionality.

Materials and methods

Bacterial isolate and growth conditions

Pantoea agglomerans strain R6 was originally isolated from *Lactuca serriola* roots collected at the University of the Western Cape (Bellville, South Africa) as previously reported by Istain, Mokgokong [15]. For routine experiments, a single colony was cultured in nutrient broth at 22 °C unless stated otherwise. Strain R6 was maintained as 50% glycerol stocks at – 80 °C and revived on Reasoners2 agar prior to experiments. Working cultures were not passaged more than twice to avoid phenotypic drift.

The designation "R6" originates from the isolate coding scheme used in the original collection study by Istain et al. [15], in which bacterial isolates obtained from *Lactuca serriola* roots were assigned sequential identifiers prefixed with "R" to denote root origin. Species identification as *Pantoea agglomerans* was determined from whole-genome sequence analysis reported in that study. The same strain name has been retained in the present work to maintain continuity with the deposited genome record (GenBank accession JACRVA000000000).

Genomic DNA extraction and 16S rRNA identification

Genomic DNA was extracted from overnight cultures using the GenElute™ Bacterial Genomic DNA Kit (Sigma-Aldrich) according to the manufacturer's instructions. Species identity was confirmed by amplification of the 16S rRNA gene using universal primers E9F and 1492R. PCR amplification was performed under standard cycling conditions (initial denaturation at 95 °C, followed by 30 cycles of denaturation, annealing at 55 °C, extension at 72 °C, and a final extension step). Amplicons were sequenced at the Central Analytical Facility (Stellenbosch University), and resulting sequences were edited in BioEdit and identified using BLAST against the NCBI nucleotide database.

Gram staining and Scanning Electron Microscopy (SEM)

Cell wall type and morphology were assessed by Gram staining using a standard protocol [3] and visualised by light microscopy. For ultrastructural analysis, SEM sample preparation was performed based on Tian, Hu [16] with minor modifications. Briefly, cells were fixed with glutaraldehyde followed by osmium tetroxide, dehydrated through a graded ethanol series, treated with HMDS, mounted on aluminium stubs, sputter-coated, and imaged using scanning electron microscopy.

Motility and swarming assays

Motility was evaluated using sulfide indole motility (SIM) medium prepared according to Shields and Cathcart [17]. Cultures were stab-inoculated into SIM agar and incubated at 25 °C for 24–48 h. Motility was recorded as diffuse growth extending from the stab line.

Swarming behaviour was assessed on Mueller–Hinton agar plates by spot-inoculating a single colony at the plate centre and incubating at 25 °C for up to 5 days. Plate images were captured before and after incubation and analysed using ImageJ [18] to compare swarm expansion.

Exopolysaccharide (EPS) production

EPS production was quantified using a phenol–sulfuric acid carbohydrate assay with minor adjustments from Elnahas, Amin [19]. Cultures were grown in LB broth (supplemented with sucrose) at 25 °C for 48 h under agitation. Cells were removed by centrifugation and EPS was precipitated from the supernatant using chilled acetone, dried, and resuspended in sterile water. Total carbohydrate content was determined by measuring absorbance at 490 nm and quantified against a glucose standard curve. All experiments were performed in triplicate.

Biofilm formation assay

Biofilm formation was evaluated using a crystal violet microtiter plate assay using Coffey and Anderson [20] method with minor adjustments. Briefly, cultures were grown overnight in LB broth, the following day it was diluted 1:100 in LB broth, and 200 µl incubated in U-bottom 96-well plates at 30 °C for 48 h. Wells were gently washed to remove planktonic cells, stained with 0.1% crystal violet, and solubilised with 30% acetic acid. Absorbance was measured at 595 nm to estimate biofilm biomass. Experiments were conducted in triplicate.

Functional and biochemical characterisation

A series of plate-based assays was conducted to evaluate enzymatic activities relevant to plant-associated fitness and pathogenic potential. All plates were spot inoculated with protocols being adapted according to the reference. Amylase activity was assessed on starch agar following incubation at 25 °C for 48 h and visualised by iodine staining, where clear zones indicated starch hydrolysis [21, 22]. Catalase activity was determined by bubble formation after exposure to 3% (v/v) H_2O_2 [23]. Cellulase activity was evaluated on carboxymethylcellulose agar incubated at 25 °C for 48 h and visualised using Congo red staining and NaCl destaining to detect halo formation [24]. Chitinase activity was assessed on colloidal chitin agar supplemented with bromocresol purple and incubated at 25 °C for 48 h, with colour change and halo formation indicating positive activity [25]. Lipase activity was determined on tributyrin agar following spot inoculation and incubation at 25 °C, where clear zones indicated lipid hydrolysis as described by Carrazco-Palafox, Rivera-Chavira [26]. Oxidase activity was assessed using commercial oxidase test strips according to the manufacturer's instructions. Protease production was evaluated on skim milk agar after incubation at 25 °C for 48 h, with

clearance zones indicating casein hydrolysis [27]. Urease activity was tested on Christensen's urea agar incubated at 25 °C for 48 h, with a colour change from yellow to pink indicating positive urease activity [28]. All assays were performed in triplicate.

Carbon source utilisation and compound production

Carbon utilisation assays were performed using standard plate-based growth screening. Glucose metabolism was evaluated on glucose agar plates incubated at 25 °C for 48 h, with growth indicating utilisation [28]. Sucrose utilisation was tested on 10% sucrose agar-based media by monitoring bacterial growth following incubation at 30 °C.

Hydrogen sulphide production was evaluated using peptone water cultures with lead acetate strips suspended in the tube headspace and incubated at 25 °C for 48 h, with strip darkening indicating H_2S production. Indole-3-acetic acid (IAA) production was quantified using the Salkowski colorimetric method as described by Gordon and Weber [29], using cultures grown in yeast extract mannitol broth with and without tryptophan supplementation, followed by spectrophotometric measurement and comparison to an IAA standard curve. Siderophore production was assessed on chrome azurol S (CAS) agar as described by Alexander and Zuberer [30], where orange halo formation indicated siderophore activity. All experiments were conducted in triplicate.

Antibiotic susceptibility screening

Ampicillin susceptibility was assessed using an agar-based growth assay. R2A agar plates supplemented with filter-sterilised 1% ampicillin were prepared, and cultures standardised to approximately $OD_{600} \approx 0.5$ were spread plated (100 µL per plate). Plates were incubated at 25 °C for 48 h and monitored for growth relative to non-antibiotic controls. All experiments were conducted in triplicate.

Nutrient solubilisation assays

Phosphate solubilisation was evaluated on Pikovskaya's agar following spot inoculation and incubation at 25 °C for 48 h, with solubilisation recorded as the formation of clear halo zones around colonies as described by Ishak, Rahim [31]. Zinc solubilisation was assessed on a defined agar medium supplemented with insoluble ZnO and varying NaCl concentrations, incubated at 25 °C for 48 h, with clearance zones indicating zinc solubilisation as previously described Dinesh, Srinivasan [32]. All experiments were conducted in triplicate.

API 20E biochemical profiling

Biochemical identification and metabolic profiling were performed using the API 20E system (BioMérieux) according to the manufacturer's instructions. Overnight cultures grown on tryptic soy agar were used to prepare bacterial suspensions in 0.85% NaCl diluent, and inoculated strips were incubated at 37 °C for 24 h. Reactions were interpreted using the API reading guide and analysed with the APILAB Plus database for species-level identification.

Genome sequencing, annotation, and comparative genomics

Whole-genome sequencing and assembly of *P. agglomerans* R6 were previously performed by Istain, Mokgokong [15]. Briefly, the draft assembly comprises 366 contigs (> 1,000 bp), with a reported GC content of 55.61% and an N50 of 25,969 bp, with approximately 100× sequencing depth. The genome project is available in GenBank under accession JACRVA000000000 (BioProject PRJNA647595; BioSample SAMN15595211; SRA SRR12904959). Genome annotation was performed using Proksee [33] incorporating Prokka v1.14.6 [34] with default parameters and the RefSeq bacterial database (release 220). Pan-genome analysis was conducted using Roary v3.13.0 [35] with a 95% BLASTp identity threshold and core gene definition of ≥ 99% strain presence; genomes were first annotated with Prokka using identical settings to ensure consistency. Visualization of the pan-genome was performed with GView server (2024 release) [36] and PanX v1.6 using default clustering parameters [37].

Genomic islands were predicted using IslandViewer4 (web version 2024-03) [38] incorporating the IslandPath-DIMOB [39] and SIGI-HMM algorithms with default significance thresholds. Prophage regions were identified using PHASTER (database update 2023-12) [40] with regions classified as intact, questionable, or incomplete according to PHASTER scoring criteria. CRISPR arrays and Cas genes were detected using CRISPRCasFinder v4.2.20 [41] with evidence level ≥ 2 and default repeat similarity thresholds. Antimicrobial resistance genes were screened using CARD RGI v6.0.3 against the CARD database v3.2.9 [42], and only hits with ≥ 60% identity and ≥ 60% coverage were considered for downstream interpretation.

Statistical analysis

Biological and technical repeats (in triplicates) were conducted for all experiments, for repeatability and reproducibility purposes. All the seedling root material mentioned were collected from 4 plants per treatment. All statistical data was analyzed using one-way analysis of variance (ANOVA) and the means were tested for significance using the GraphPad Prism 8.0.1 software by applying the Tukey-Kramer test at 5% level of significance.

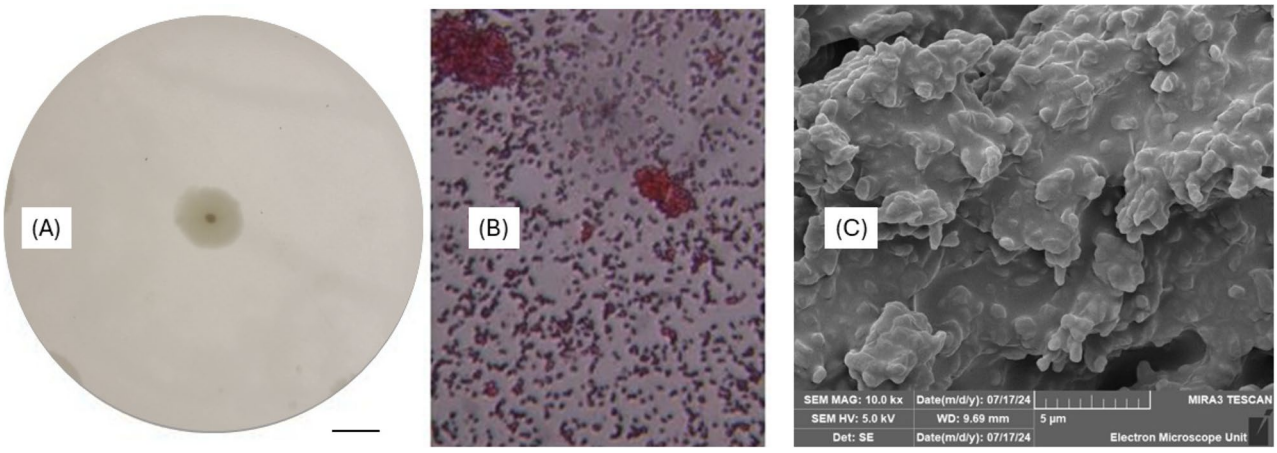

Fig. 1 Morphology and surface-associated behaviour of *Pantoea agglomerans* R6. **A** Swarming motility on Mueller–Hinton agar following central point inoculation and incubation at 25 °C; colony expansion was recorded at 24 h intervals for 48 h with a 1.5cm scale. **B** Gram-stained preparation observed at 100× oil immersion showing Gram-negative, rod-shaped cells. **C** Scanning electron micrograph (5 µm scale indicated at the bottom of the image), observed with a SEM MAG of 10.0 kx and SEM HV of 5.0 kV, illustrating bacillary morphology and cell clustering embedded within an extracellular matrix consistent with EPS and biofilm formation

Results and discussion

The swarming ability of *Pantoea agglomerans* (Fig. 1A) is a critical aspect of its motility and pathogenicity. Swarming is a coordinated, flagella-driven movement of bacterial cells across solid surfaces, allowing them to colonize and spread efficiently. This process is often regulated by genes that govern flagellar synthesis and function, such as fliC (encoding flagellin) and fliA (a sigma factor required for flagellar gene expression), which are essential for motility and swarming [43] (Supplementary Figure S1). The capacity to swarm not only facilitates rapid surface colonization but also contributes to biofilm formation, which enhances persistence in both environmental and host-associated niches.

The Gram-negative cell envelope of *P. agglomerans* provides a structural basis for several phenotypes observed in strain R6. The outer membrane, enriched in lipopolysaccharides, contributes to surface interactions that support swarming motility and initial attachment, consistent with the extensive colony expansion recorded on Mueller–Hinton agar (Fig. 1A) [44, 45]. Swarming analysis demonstrated a marked expansion of the R6 colony from an initial area of 0.22 arbitrary units at time 0 to 5.92 units after 48 h, representing a 26.9-fold increase in surface coverage as quantified using ImageJ (Table S1). This envelope architecture also facilitates the retention of extracellular polymeric substances visualised by SEM (Fig. 1C; Figure S5), promoting cell clustering and biofilm initiation [46, 47]. Furthermore, the outer membrane acts as a selective permeability barrier, which is relevant to the ampicillin susceptibility profile of R6, as β-lactams must traverse this layer to reach their peptidoglycan targets [48, 49]. These links connect cell wall structure with the measured motility, biofilm-associated morphology, and antibiotic response of the strain.

Table 1 Extracellular polymeric substance (EPS) production by *P. agglomerans* R6

Bacterium	Glucose (µg/ml)
P. agglomerans R6	420.12 ± 39.53 a
Positive control	569.28 ± 50.20 b

EPS was quantified using the phenol–sulfuric acid assay and expressed as µg glucose equivalents mL^{-1} based on a glucose standard curve. LB medium served as negative control and Microbacterium sp. as positive control. Values represent mean ± SE ($n = 10$). Different letters (a, b) indicate statistically significant differences at $p < 0.05$ according to Tukey–Kramer test

In *P. agglomerans* R6, the rod-shaped Gram-negative morphology observed by Gram staining (Fig. 1B) corresponds with the cell clustering seen in SEM images (Fig. 1C), supporting the capacity of the strain to establish surface-associated communities. Such aggregation is consistent with the quantitative EPS production measured for R6 (Table 1) and with the crystal-violet biofilm assay, traits that contribute to protection and resource utilisation within developing biofilms [50–52].

Scanning electron microscopy (SEM) provides further resolution into the structural and functional organization of *P. agglomerans*. SEM images of *P. agglomerans* R6 revealed the presence of bacterial clusters embedded within an extracellular polymeric substance (EPS) matrix. EPS production is a hallmark of biofilm formation, serving as a protective scaffold and enabling cell-to-cell adhesion. EPS production is a critical factor underlying biofilm formation in *P. agglomerans* [46]. Table 1 illustrates the EPS production of *P. agglomerans* R6, which correlates with its biofilm-forming ability. EPS facilitates bacterial adhesion to surfaces, provides a protective barrier against environmental stresses such as desiccation and predation, and enhances resistance to antibiotics [47,

53]. In addition to these protective roles, EPS biosynthetic gene clusters have been linked to virulence traits, including interactions mediated by the type VI secretion system (T6SS). The T6SS contributes to pathogenicity by enabling direct interactions with host cells as well as competitive exclusion of other microorganisms [54, 55]. Collectively, EPS production and its regulation highlight the multifaceted role of *P. agglomerans* in both persistence and pathogenesis.

Biofilm development is a multistage process regulated by genetic and environmental cues, with surface topography playing a central role. Rougher surfaces enhance bacterial attachment, facilitating subsequent biofilm maturation [56]. SEM analysis (Fig. 1C) of *P. agglomerans* revealed clustering similar to Gram-stained preparations (Fig. 1B) and highlighted the extracellular matrix that underpins biofilm integrity [57, 58]. To complement these observations, a qualitative biofilm assay was conducted using crystal violet staining (Fig. 2) with quantitative data in Table S4. When compared with *Microbacterium sp*., a known biofilm producer, *P. agglomerans* displayed similar colour patterns, in contrast to the negative control. Environmental conditions modulate biofilm formation in *P. agglomerans*, with nutrients and signaling molecules enhancing EPS production and biofilm stability [59, 60]. These modifications influence its interactions with plant hosts, thereby shaping virulence. Biofilms provide a protective niche for colonization and have been implicated in plant diseases such as gall formation and rot, underscoring the role of biofilm-mediated pathogenicity in *P. agglomerans* [3, 61].

Together, swarming motility, Gram-negative cell wall structure, and SEM-based morphological observations illustrate the interconnectedness of motility, adhesion, and biofilm formation in *P. agglomerans*. Swarming enables initial colonization, Gram-negative cell wall architecture supports structural resilience and pathogenic interactions, while SEM imaging reveals the tangible outcomes of these processes, clustered cells embedded in biofilm matrices. These features collectively underscore the bacterium's ecological adaptability and pathogenic potential.

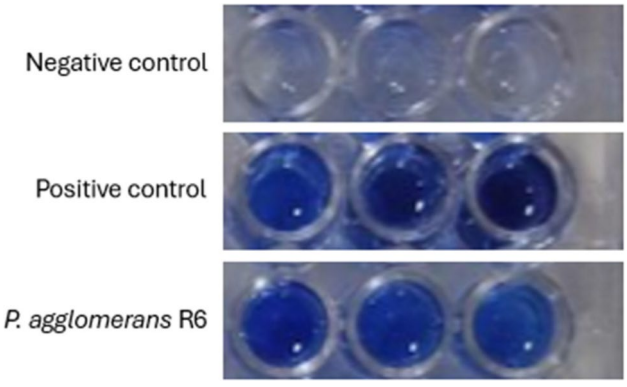

Fig. 2 Crystal violet assessment of biofilm formation by *P. agglomerans* R6. Biofilm biomass was evaluated in 96-well polystyrene plates following 48 h static incubation at 30 °C. Wells were stained with 0.1% crystal violet, washed to remove planktonic cells, and visualised after solubilisation of bound dye. LB medium served as negative control and *Microbacterium* sp. as positive biofilm control. Quantitative OD595 values are provided in the Table S2

Phenotypic characterisation of *P. agglomerans* R6

The experimental panel was selected to evaluate traits that differentiate beneficial plant-associated behaviour from opportunistic pathogenic potential in *P. agglomerans*. Biofilm and EPS assays were included to assess surface colonisation capacity and persistence on plant tissues. Enzymatic tests (amylase, lipase, oxidase) (Table S2) were chosen because such activities have been linked to degradation of host substrates and interaction with plant defences. Antibiotic susceptibility testing was performed to determine whether CARD-predicted resistance determinants confer phenotypic resistance, an important safety consideration for potential agricultural application. API 20E profiling provided metabolic context for ecological adaptation and confirmed species-level identification.

Functional assays indicated that R6 expresses multiple traits associated with plant-associated fitness and potential host interaction. The strain produced substantial biofilm biomass (crystal violet OD595 = 0.84 ± 0.07) supported by EPS levels of 420.12 ± 39.53 µg glucose equivalents mL^{-1}, consistent with the extracellular matrix visualised by SEM. Enzymatic screening detected amylase and lipase activities. Amylase production is noteworthy, as similar enzymes in phytobacteria are secreted via the type II secretion system to degrade host substrates and facilitate colonisation [62, 63]. Lipase activity may enhance adhesion to host cells, modulate immune responses, and support survival in competitive environments [64, 65]. R6 also displayed filamentous growth, a strategy associated with depletion aggregation and biofilm stabilisation through EPS, flagella, pili, and quorum sensing signals [66].

Although *P. agglomerans* is typically described as oxidase negative [67], R6 demonstrated oxidase activity, suggesting functional aerobic respiration supported by cytochrome bd oxidases that enable growth under low oxygen and contribute to oxidative stress tolerance [68, 69]. This respiratory flexibility may provide a competitive advantage in plant-associated niches with fluctuating oxygen levels. In antibiotic susceptibility testing, R6 remained susceptible to ampicillin, consistent with reports that some *P. agglomerans* strains lack β-lactam resistance [50] and indicating that CARD-predicted homologues did not confer phenotypic resistance under the conditions tested.

High-concentration sucrose (10%) was used to mimic plant-like conditions and evaluate its impact on surface-associated traits. In several bacteria, including *Pseudomonas aeruginosa* and *Streptococcus mutans*, sucrose modulates expression of virulence-related genes and stimulates EPS-mediated biofilm formation [49, 70, 71]. Although less potent than glucose, sucrose plays an important regulatory role in plant-associated bacteria [72–74].

API 20E metabolic profiling (Table S3) revealed utilisation of mannitol, rhamnose, sucrose, arabinose, and amygdalin, consistent with metabolic patterns reported for the genus [75]. Mannitol utilisation occurs via mannitol dehydrogenase, converting mannitol to fructose for entry into glycolysis [76], and disruption of this pathway can impair bacterial growth [77]. Rhamnose and arabinose, abundant in plant tissues, were metabolised through their respective isomerases [78, 79], supporting colonisation and persistence in host environments [80, 81]. Sucrose fermentation, mediated by invertase hydrolysis to glucose and fructose, further supports colonisation of sugar-rich plant surfaces [82], and loss of sucrose metabolism has been associated with reduced bacterial fitness [83, 84]. Based on API 20E analysis, R6 corresponded to profile code 1,005,133, supporting its identification as *P. agglomerans*.

Comprehensive biochemical and API 20E datasets are provided in Tables S1 and S2.

Genomic exploration of *Pantoea agglomerans* R6 for gene mining

Whole-genome sequencing and draft assembly of *P. agglomerans* R6 were previously generated using Ion Torrent technology by Istain, Mokgokong [15]. The published assembly is approximately 4.7 Mb in size (4,671,676 bp), with a GC content of 55.61%, and comprises 366 contigs (> 1,000 bp) with an N50 of 25,969 bp. The genome encodes 4,349 predicted genes, including 4,083 protein-coding sequences, together with multiple RNA features (rRNAs, tRNAs, and ncRNAs) that support core cellular functions.

In addition, the genome contains pseudogenes, including disrupted and truncated coding sequences, which may reflect genome plasticity and ongoing adaptation within plant-associated environments.

To contextualise the genetic repertoire of R6, comparative analysis was performed against publicly available *P. agglomerans* genomes (KM1, C410P1, TH81, UAEU18, and L15) and visualised using a pan-genome map (Fig. 3). Across these strains, conserved loci included key motility and chemotaxis genes, iron acquisition systems, and secretion-associated components, including T6SS-associated proteins (e.g., Hcp and VgrG), consistent with previous pan-genome findings [85]. Genes linked to adhesion, haemolysin transport, and multidrug efflux systems were also detected across strains, supporting the view that core host-interaction traits and antimicrobial tolerance mechanisms are widely distributed within *P. agglomerans* lineages.

A comparative analysis of genomic islands (GIs) was performed across six genomes, including *P. agglomerans* strains KM1, C410P1, TH81, UAEU18, and L15, with a pathogenic *Escherichia coli* genome included as a reference comparator (Fig. 4). The IslandCompare workflow was used to support GI detection and cross-genome comparison, enabling integrated visualisation of predicted islands and associated antimicrobial resistance (AMR) determinants. GI prediction was based on IslandPath-DIMOB and SIGI-HMM [86, 87], which identify candidate horizontally acquired regions using sequence composition features such as dinucleotide bias, the presence of mobility-associated genes, and atypical codon usage patterns. This comparative framework provided an overview of shared and strain-specific GI regions across the analysed genomes and guided downstream interpretation of GI-associated virulence- and AMR-linked loci discussed in subsequent sections.

IslandViewer4 (Fig. 5) presents a comparison between *P. agglomerans* R6 and *P. agglomerans* strain C410P1, highlighting predicted genomic island regions, using a data base from [88], with potential pathogenicity-associated content shared between the two strains. Several candidate loci were detected using IslandPath-DIMOB, including *cutA*, *xisR*, *pagP*, *glsB*, *yefM*, *mqo*, *gstA*, and *capV*, while *dnaB* and *capV* were additionally identified by SIGI-HMM. Notably, c*apV* was consistently detected across prediction approaches, suggesting it may represent a conserved GI-associated feature in these genomes.

The gene capV has previously been implicated in bacterial survival-related phenotypes. For example [89], reported that overexpression of a *capV* variant (*capV* Q329R) promoted bacterial filamentation, a trait that can enhance persistence under stress conditions. Because filamentation has also been linked to immune evasion, including reduced susceptibility to phagocytosis, these findings support a potential role for c*apV*-associated functions in host interaction and opportunistic pathogenicity.

Pan-genome analysis using PanX across 12 *P. agglomerans* genomes identified 8,604 predicted protein-coding genes in the species pan-genome, comprising 5,600 accessory genes (65%) and 3,004 core genes (34%) (Fig. 6). In comparison [62], analysed five *P. agglomerans* strains using a Roary-based pipeline and reported a smaller pan-genome of 6,727 genes, with 3,065 genes (45.6%) assigned to the accessory fraction and 3,662 genes (54.4%) representing the core genome. The observed increase in pan-genome size with the inclusion of additional strains,

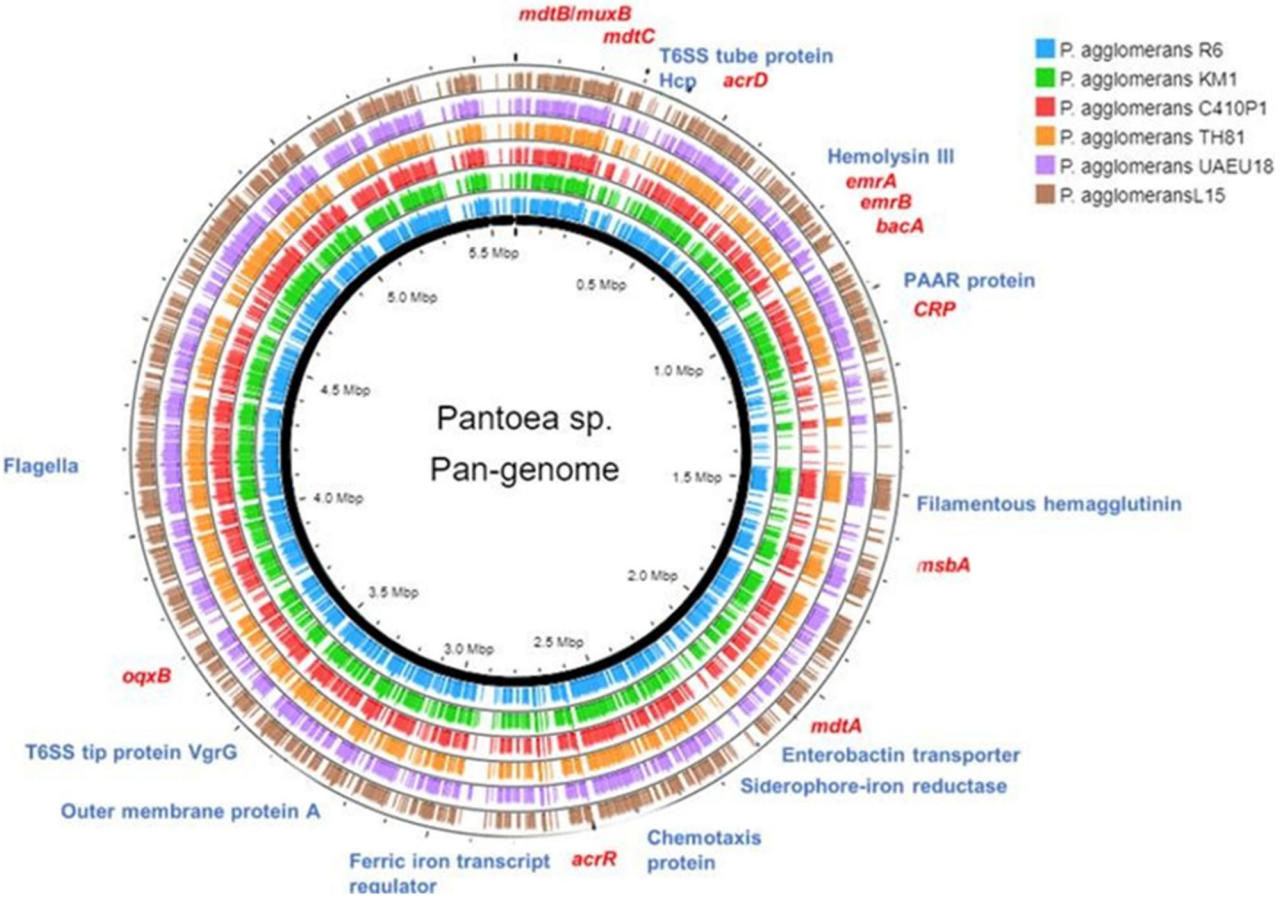

Fig. 3 Pan-genome analysis of *P. agglomerans* strains obtained using Gview server. Starting from the innermost ring: Backbone (Contigs), Ring 2: *P. agglomerans* R6, Ring 3: *P. agglomerans* KM1, Ring 4: *P. agglomerans* C410P1, Ring 5: *P. agglomerans* TH81, Ring 6: *P. agglomerans* UAEU18, Ring 7: *P. agglomerans* L15

together with a reduction in the proportional size of the core genome, is consistent with an open pan-genome structure for *P. agglomerans* [89].

An open pan-genome reflects the presence of a conserved genetic backbone alongside a variable gene pool that differs between lineages. As described by [90], this accessory fraction frequently contains genes that enhance niche adaptation and may influence host interaction phenotypes. Horizontal gene transfer (HGT) is a major driver of this variability, facilitating the acquisition and dissemination of functions such as antimicrobial resistance and virulence-associated traits [91]. In the present analysis, strain-specific gene content (Fig. 3) supports ongoing genome diversification within *P. agglomerans*, consistent with HGT contributing to ecological flexibility and potential opportunistic behaviour.

In addition to the variable gene pool identified through pan-genome analysis, *P. agglomerans* R6 was found to encode CRISPR–Cas features, including seven CRISPR arrays, four cas genes, and two predicted cas clusters (Table 2; Fig. 7) [41]. The CRISPR-associated sequences were detected across multiple contigs, which may reflect locus fragmentation due to the draft assembly rather than discrete independent arrays. CRISPR–Cas systems provide defence against invading nucleic acids such as bacteriophages and plasmids and can therefore influence genome stability and the rate at which mobile genetic elements are acquired or maintained. The presence of these loci, together with evidence of mobile genetic elements in the R6 genome (Fig. 9), suggests that R6 has the capacity to both encounter and respond to foreign DNA, supporting persistence in competitive microbial environments.

While CRISPR–Cas systems are not direct virulence determinants, they may indirectly shape strain fitness by limiting phage susceptibility and modulating the retention of horizontally acquired traits. In this context, the CRISPR–Cas repertoire in R6 may contribute to long-term survival and adaptation, which are relevant features for evaluating the ecological success and opportunistic potential of *P. agglomerans*.

The Comprehensive Antibiotic Resistance Database (CARD) and its Resistance Gene Identifier (RGI) platform were used to screen the *P. agglomerans* R6 genome for putative antimicrobial resistance (AMR)-associated

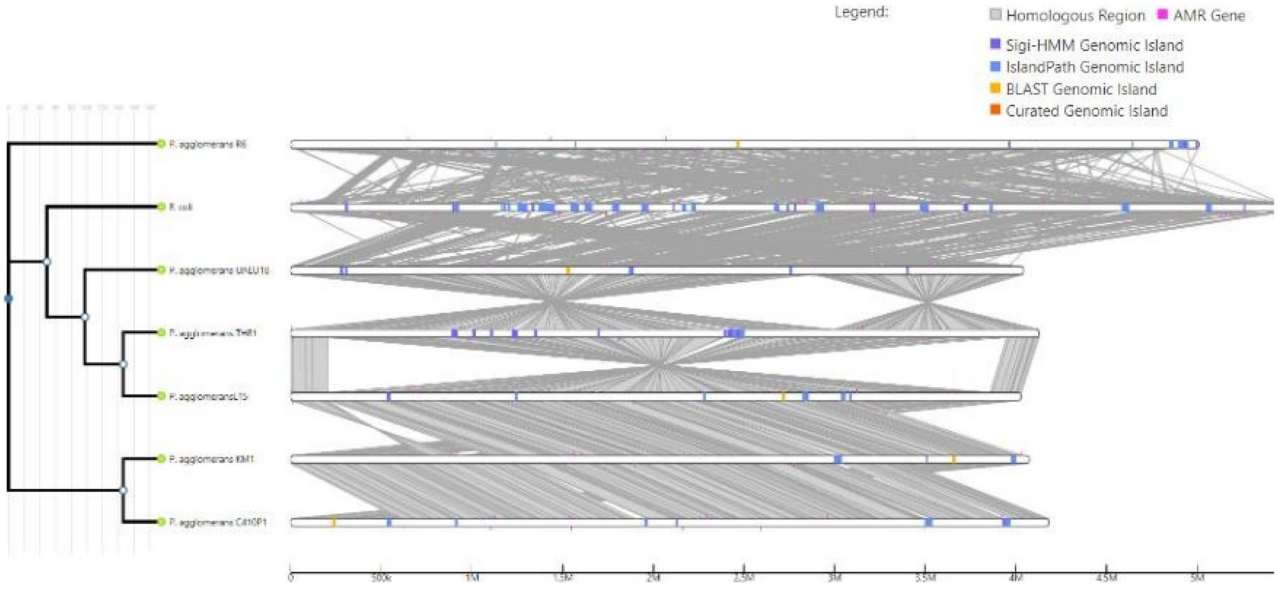

Fig. 4 Comparative genomic island analysis of *Pantoea agglomerans* genomes. Predicted genomic islands (GIs) were identified using the IslandCompare/IslandViewer4 pipeline and visualized across four *P. agglomerans* genomes. **a** Maximum-likelihood phylogenetic tree illustrating the relatedness of the analysed strains. **b** Linear genome maps showing predicted GI regions as coloured segments along each genome; grey connectors indicate homologous regions shared between genomes. **c** Gene-content view of selected GI clusters displaying annotated coding sequences associated with mobility, antimicrobial resistance, and secretion-related functions. Colours reflect GI groupings based on prediction algorithms and clustering

loci [42]. The initial RGI output returned a large number of matches (> 1,000) (Fig. 7), reflecting the broad detection scope of CARD, which includes both high-confidence resistance determinants and lower-confidence homology-based hits. To focus interpretation on the most relevant candidates, hits were further evaluated based on sequence similarity, and only a subset of genes exceeded a 60% match threshold (Table 3).

Among these higher-scoring candidates were *farA*, *gyrB*, *emrR*, *adeF*, *FEZ-1*, *ompK37*, *qnrS8*, *eptA*, and *arnT*, representing resistance mechanisms including antibiotic efflux, target alteration, target protection, and reduced permeability. These predicted AMR-associated features suggest that strain R6 may possess genetic capacity for antimicrobial tolerance under selective pressure, which could support persistence in competitive or stressed environments. Further phenotypic validation would be required to confirm whether these loci translate into clinically or agriculturally relevant resistance profiles.

Two loci, *gyrB* and *emrR*, are associated with fluoroquinolone resistance mechanisms based on CARD homology; these represent genomic predictions and do not imply phenotypic resistance. Fluoroquinolone resistance mechanisms show potential against antibiotics such as enoxacin; ciprofloxacin; levofloxacin; moxifloxacin; gatifloxacin; lomefloxacin; nalidixic acid; norfloxacin; ofloxacin; trovafloxacin; grepafloxacin; sparfloxacin and pefloxacin, with some being efficient second, third and even fourth generation antibiotics. The resistance mechanisms deployed by each gene and the drug class it falls under is shown in Table 3.

Using Bakta v1.8.2 (DB v5.0, Light) [92], additional genomic features were identified in *P. agglomerans* R6, including three short open reading frames (sORFs) and two predicted OriC protein regions (Table 4).

Short open reading frames are increasingly recognised as a common but under-annotated component of bacterial genomes. Although many sORFs encode peptides of ~ 100 codons or fewer, accumulating evidence suggests that these small proteins can contribute to cellular regulation and stress adaptation, including functions such as toxin–antitoxin components, modulators of protein stability, and regulatory factors [90, 93–95]. Because sORFs are often missed by conventional annotation workflows, their identification in R6 highlights genomic elements that may influence strain-level phenotypes but require further functional investigation.

In addition, the origin of chromosomal replication (OriC) represents a central regulatory locus controlling genome duplication. Replication initiation is typically mediated by DnaA binding to conserved motifs within OriC, and this process is coordinated to ensure accurate timing and once-per-cycle replication under varying environmental conditions [92, 96–98]. The detection of OriC features in the R6 genome supports the completeness of the assembly and provides context for genome-level regulation and stability.

Taken together, the identification of sORFs and OriC-associated features (Fig. 8) adds further resolution to the

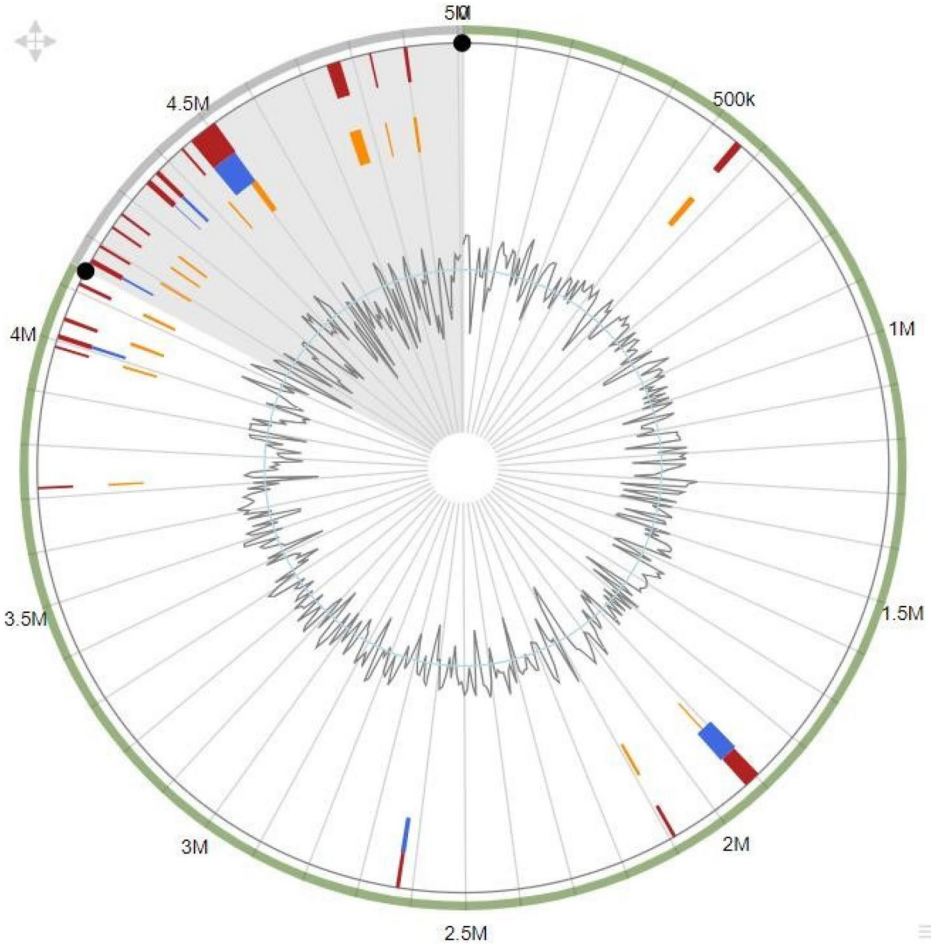

Fig. 5 *P. agglomerans* R6 aligned against reference genome *P. agglomerans* strain C410P1 for basic comparative analysis of predicted genomic islands using IslandViewer 4. The circular map represents the genome of R6. The highlighted regions indicate predicted Genomic Islands. Regions predicted by multiple methods (shown in Yellow - SIGI-HMM, Blue - IslandPath- DIMOB and Red - At least one method) are of particular interest due to their potential role in Pathogenicity Islands (PAIs), Resistance Islands, Metabolomic Islands, and or Symbiosis Islands

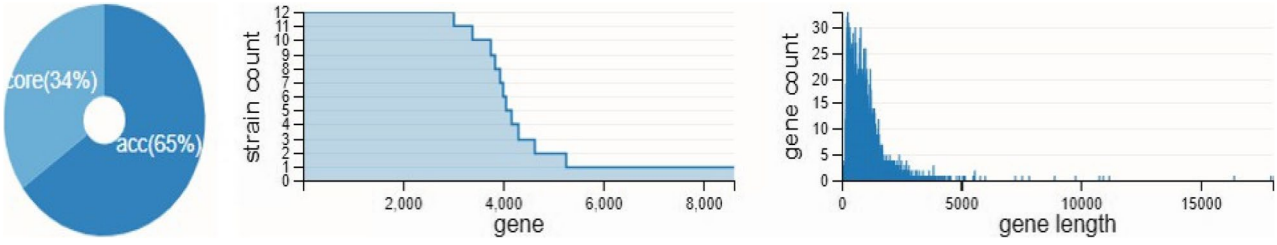

Fig. 6 PanX based gene count within 12 strains of *P. agglomerans*. The core and accessory (acc) genes making up the genome of *P. agglomerans* species

genome architecture of *P. agglomerans* R6 and highlights genetic elements that may contribute to fitness and persistence under fluctuating environmental conditions.

By interacting with larger membrane proteins, small hydrophobic proteins, which are encoded by small open reading frames, assist bacterial pathogens adapt to environmental stresses and thus regulate bacterial virulence (Fig. 8) [100]. Small membrane proteins encoded by sORFs may also represent promising targets for the development of synthetic peptides with antivirulence or antibacterial activity. However, their small size (< 100 amino acids) and often subtle contributions to cellular function make them difficult to detect and less amenable to random mutagenesis screens, meaning their roles in pathogenicity could easily be underestimated [101].

Bacteriophages, viruses that specifically infect bacterial cells, play crucial roles in microbial ecology and the dynamics of pathogenesis [99]. In *P. agglomerans* R6, four phages were identified: three double-stranded DNA (dsDNA) phages with 100% sequence matches

Table 2 CRISPR–Cas features identified in *Pantoea agglomerans* R6

Classification	Contig	Strand	length
CRISPR I	JACRVA010000002.1	Forward	92 bp
CRISPR II	JACRVA010000002.2	Forward	104 bp
CRISPR III	JACRVA010000025.1	Forward	122 bp
CRISPR IV	JACRVA010000039.1	Forward	110 bp
CRISPR V	JACRVA010000043.1	Forward	129 bp
CRISPR VI	JACRVA010000063.1	Forward	92 bp
CRISPR VII	JACRVA010000188.1	Reverse	120 bp
Cas gene I	JACRVA010000127.1	Forward	1,329 bp
Cas gene II	JACRVA010000128.1	Forward	1,636 bp
Cas gene III	JACRVA010000197.1	Forward	1,383 bp
Cas gene IV	JACRVA010000197.1	Forward	1,383 bp
Cas cluster I	JACRVA010000128.1	Forward	1,329 bp
Cas cluster II	JACRVA010000197.1	Forward	1,636 bp

CRISPR arrays and associated Cas genes were detected using CRISPRCasFinder v4.2.20. The table lists array coordinates, repeat number, consensus repeat sequence, predicted Cas gene composition, and evidence level as defined by the software

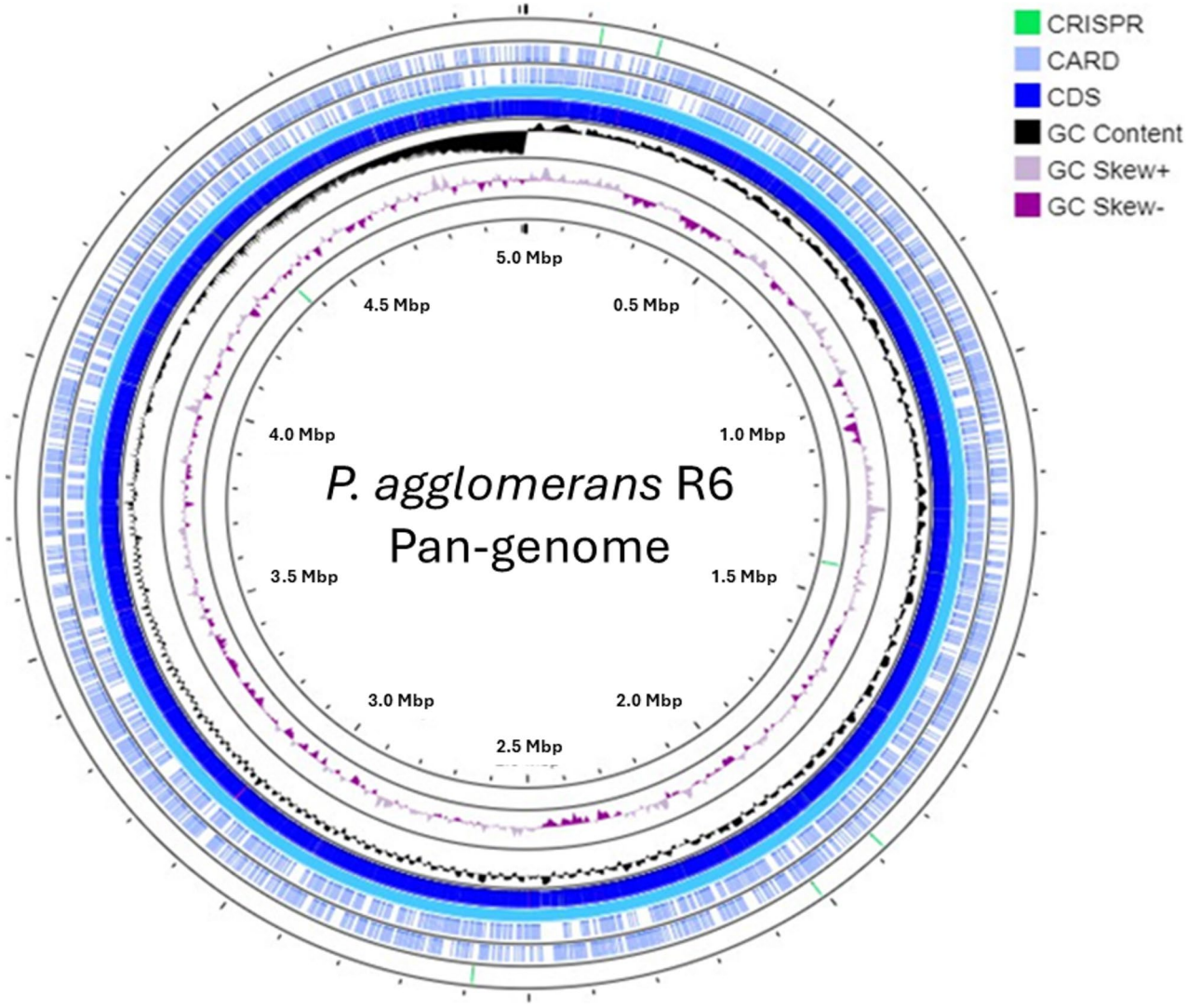

Fig. 7 Pan-genome annotation analysis of *P. agglomerans* R6 obtained using Gview server. Circles illustrate the following from outermost to innermost rings : (1) CRISPR - Showing CRISPR arrays and their associated Cas proteins [41], (2) CARD - Comprehensive Antibiotic Resistance Database (CARD) Resistance Gene Identifier (RGI) [42], (3) CDS, (4) GC content, and (4) GC skew, (5) CRISPR - Showing CRISPR arrays and their associated Cas proteins [1, 41]

Table 3 Antimicrobial resistance–associated determinants identified in *Pantoea agglomerans* R6 using CARD-RGI

Gene	Match (%)	Drug Class	Resistance Mechanism	Antibiotic
farA	60	antibacterial free fatty acids	antibiotic efflux	palmitic acid; oleic acid; linoleic acid
gyrB	80.85	fluoroquinolone antibiotic	antibiotic target alteration	enoxacin; ciprofloxacin; levofloxacin; moxifloxacin; gatifloxacin; lomefloxacin; nalidixic acid; norfloxacin; ofloxacin; trovafloxacin; grepafloxacin; sparfloxacin; pefloxacin
emrR	80.12	fluoroquinolone antibiotic	antibiotic efflux	nalidixic acid
adeF	60.7	fluoroquinolone antibiotic; tetracycline antibiotic	antibiotic efflux	tetracycline
ompK37	60.36	monobactam; carbapenem; cephalosporin; cephamycin; penam; penem	reduced permeability to antibiotic	cefoxitin; cefotaxime
qnrS8	60.87	fluoroquinolone antibiotic	antibiotic target protection	ciprofloxacin; levofloxacin; moxifloxacin; gatifloxacin; nalidixic acid; norfloxacin; sparfloxacin
eptA	60.14	peptide antibiotic	antibiotic target alteration	polymyxin B
arnT	60.46	peptide antibiotic	antibiotic target alteration	colistin A; colistin B

Predictions were generated with CARD RGI v6.0.3 (CARD database v3.2.9) using ≥ 60% sequence identity and ≥ 60% coverage thresholds. The table reports gene homologues, predicted resistance mechanisms, antibiotic classes, and model confidence. These represent genomic predictions and do not imply confirmed phenotypic resistance

Table 4 Open Reading Frames (sORFs) and predicted OriC features in the Small*Pantoea agglomerans*R6 genome

Protein name	Length (bp)	Amino Acid Sequence	Accession
LeuL	27	MFRSVRLLGLLLNALSLRGRLVGGIKH	MBE5683307.1
ThrL	22	MRKISLNTTIITTTITTGNGAG	MBE5682804.1
AzuC	25	MKRIFKHVFRAYVETFKHVPPGAMY	MBE5683414.1
OriC	18 786 – 19 215	Origin of replication	-
OriC	19 353 – 21 093	Origin of replication	-

sORFs were identified during Prokka-based genome annotation, and OriC regions were predicted using standard GC-skew and motif analyses. The table lists genomic coordinates, predicted peptide lengths, functional annotations where available, and features associated with chromosomal replication origin

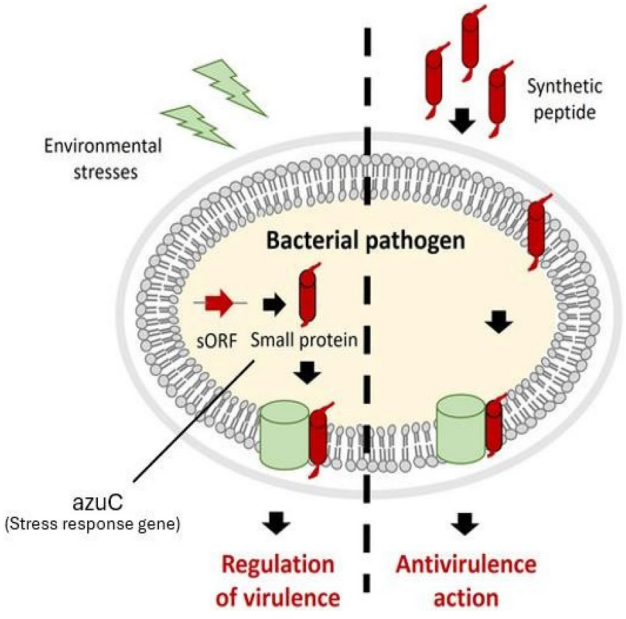

Fig. 8 Conceptual overview of potential roles of small open reading frames (sORFs) in *Pantoea agglomerans* R6. The diagram summarises predicted functions of sORF-encoded peptides identified in the R6 genome, including possible involvement in regulatory networks, modulation of stress responses, and interaction with host-associated pathways. Arrows indicate putative influences on traits related to persistence and host interaction; relationships are based on bioinformatic prediction and literature reports rather than experimental validation in R6. Adapted from [99]

to known viral phages, and one single-stranded DNA (ssDNA) phage. Their hallmark scores were 2, 0, 1, and 0, respectively.

Virulent phages typically hijack bacterial cellular machinery, replicate their genomes, and ultimately lyse the host cell [102]. While lysis itself does not directly increase pathogenicity, it can indirectly influence disease outcomes by releasing toxins, enzymes, or other virulence factors from lysed cells into the surrounding environment. Moreover, temperate or lysogenic phages can integrate into bacterial genomes, potentially carrying and spreading genes that modulate virulence, resistance, or metabolic traits.

Thus, the phages detected in *P. agglomerans* R6 may shape its ecological fitness and pathogenic potential, either through direct genetic contributions or by influencing interactions within microbial communities (Figs. 9 and 10). Bacteriophages are key contributors to bacterial genome evolution and can influence strain-level traits through lytic activity and lysogenic integration [99]. In *P. agglomerans* R6, four putative phage-associated regions were detected (Fig. 9), including three sequences with a 100% match to dsDNA phages and one associated with ssDNA phage features. The predicted regions displayed hallmark scores of 2, 0, 1, and 0, respectively, indicating variable levels of phage signature support across

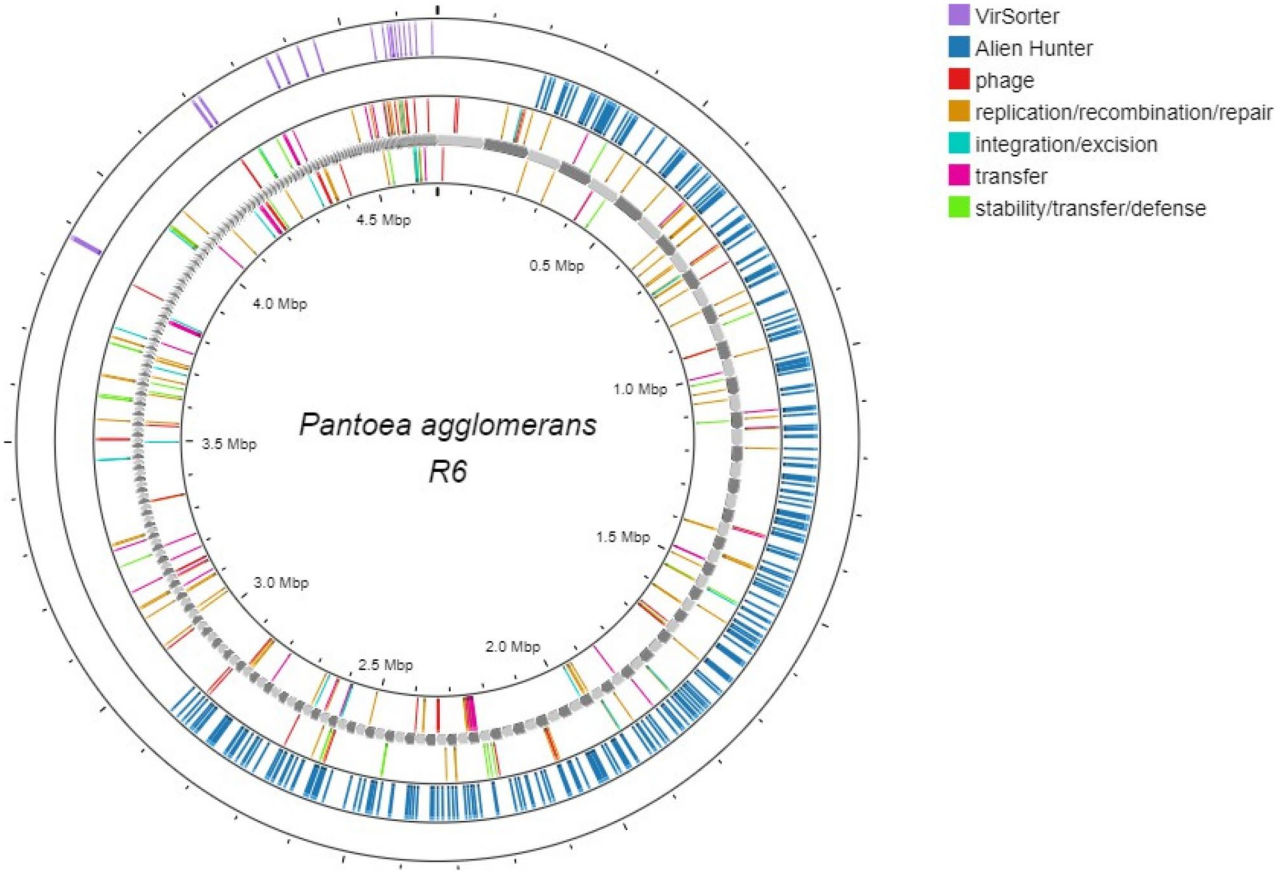

Fig. 9 Pan-genome mobile genetic element analysis of *P. agglomerans* R6 obtained using Gview server. Circles illustrate the following from outermost to innermost rings: (1) VirSorter - Detect dsDNA and ssDNA virus genomes (phage) [103], (2) Alien Hunter - Predict putative Horizontal Gene Transfer (HGT) events [104], (3) mobileOG-db - Find mobile genetic elements (MGEs) [105]

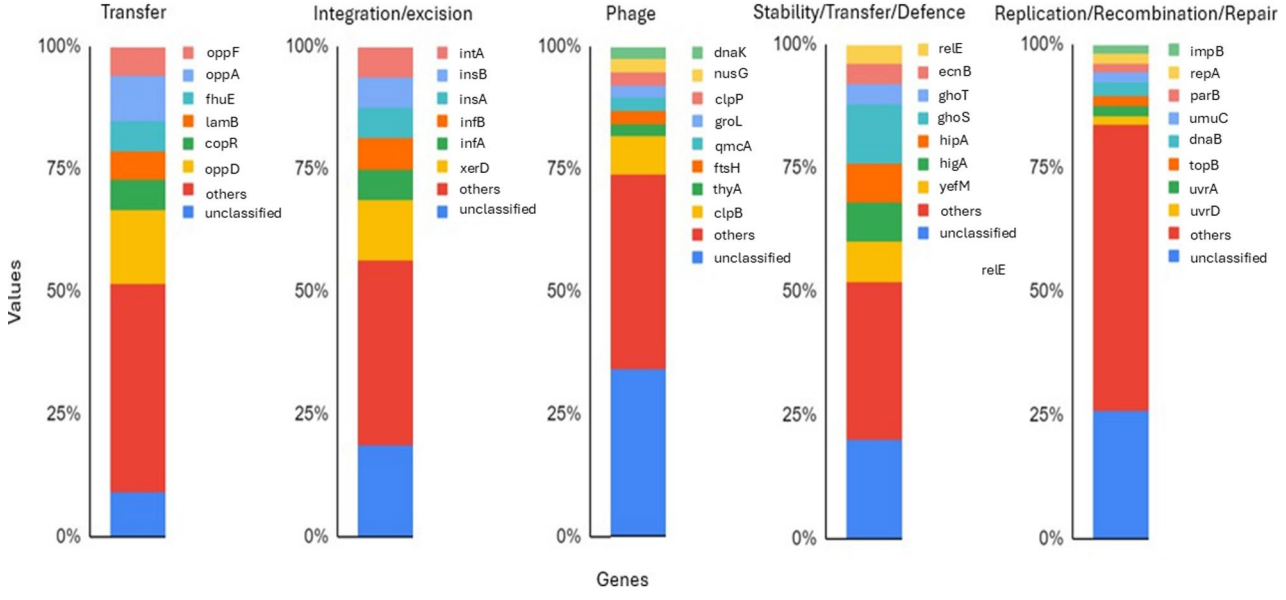

Fig. 10 Occurrence of mobile genetic elements (MGEs). The MGEs were classified in five main groups: Transfer (T), integration/excision (IE), Phage (P), stability/transfer/defense (STD), replication/recombination/repair (RRR)

the genome. While strictly lytic phages primarily impact bacterial populations through host cell lysis [102], temperate phages may persist as prophages and contribute to genome diversification by facilitating horizontal gene transfer and introducing accessory functions that can affect host interaction, stress tolerance, or antimicrobial resistance potential [104].

Consistent with this, one phage-associated sequence (JACRVA010000216.1) showed an 85.36% match to a Siphoviridae-related phage (Siphoviridae sp. ct2u94), a family frequently associated with lysogeny and gene movement between bacterial hosts [105, 106]. Putative horizontally acquired regions were additionally supported by Alien_hunter 1.7 analysis (Fig. 9), with candidate contigs listed under the Alien_hunter output [107]. Horizontal gene transfer is a major mechanism driving bacterial adaptation, enabling acquisition of traits such as resistance determinants, metabolic functions, and host-associated fitness factors that may broaden ecological range [108, 109]. In environments where microbial competition is intense, antibiotic-producing taxa can act as reservoirs for resistance-associated genes, which may subsequently be disseminated across bacterial communities [110, 111]. Together, these observations suggest that mobile DNA has likely contributed to the accessory genome of R6 and may influence its ecological persistence and opportunistic potential [112]. To further evaluate genome mobility, mobileOG-db (Beatrix 1.6 v1) analysis identified 278 mobile genetic element (MGE)-associated features in R6 (Fig. 10) distributed across five functional categories: integration/excision (24), recombination/repair (112), phage-related (46), stability/transfer/defence (39), and transfer (57) [105]. The abundance and diversity of these MGE-associated functions indicate a genetically dynamic genome with capacity for ongoing acquisition and rearrangement of accessory loci, supporting adaptation under fluctuating environmental and host-associated selective pressures.

Comparative identification of genes associated with pathogenicity and virulence in Pantoea sp

Putative virulence-associated loci in WT-R6 were identified using BLAST-based screening against the Virulence Factor Database (VFDB) together with genome annotation outputs generated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP). As summarised in Table 5, candidate genes were grouped into five functional categories: secretion systems, adhesion, motility, iron uptake/sequestration, and toxin-related factors. These assignments were based on the presence of annotated homologues within the *P. agglomerans* R6 genome and should be interpreted as genomic predictions of potential host-interaction capacity.

Within the secretion system category, WT-R6 encoded a complete set of structural components associated with a type VI secretion system (T6SS), including core Tss proteins (TssA–TssM) as well as hallmark effectors and associated components such as hemolysin-coregulated protein (Hcp), valine–glycine repeat protein (VgrG), PAAR-repeat proteins, and the ATPase ClpV, collectively supporting the presence of a functional T6SS apparatus. Adhesion-associated genes included outer membrane protein A (OmpA) and filamentous haemagglutinin (FHA), which are commonly linked to bacterial attachment and persistence in host-associated environments. Consistent with the observed motility phenotype, multiple flagellar genes were also detected, and these may additionally contribute to surface attachment, secretion-associated processes, and biofilm formation.

Iron acquisition and sequestration functions were represented by both regulatory and transport-associated loci, including the ferric uptake regulator (Fur) and iron uptake protein (EfeO). WT-R6 also encoded the iron/manganese ABC transporter permease (SitC), which has previously been reported in *P. agglomerans* KM1 (including plasmid-associated contexts) [64]. Additional siderophore-related uptake genes were detected, including ferric siderophore ABC transporter components (FepD), enterobactin transport-associated proteins (FepB, FepG, EntS), siderophore-iron reductase (FhuF), and multiple siderophore-associated transport components, including TonB-dependent receptor homologues such as FepA, collectively supporting the capacity for high-affinity iron scavenging.

Finally, toxin-associated annotations included haemolysin-related loci such as haemolysin secretion/activation components (Hha, ShlB, FhaC, HecB), haemolysin XhlA, and haemolysin III-like cytotoxic proteins. Haemolysins have been described as important contributors to host damage and inflammatory signalling in multiple bacterial systems, including *Serratia marcescens*, where haemolysin activity has been linked to induction of pro-inflammatory cytokines during infection [113–115]. Although the contribution of these predicted haemolysin-associated genes to virulence in WT-R6 remains to be experimentally determined, their presence supports the potential for cytotoxic or host-interaction phenotypes under appropriate conditions.

The characterisation of strain R6 is significant because it exemplifies the dual nature of *P. agglomerans* as both a plant-associated commensal and a potential opportunistic pathogen. Genomic features detected in R6, including type VI secretion components, adhesion-related factors, and high-affinity iron uptake systems, are commonly associated with competitive fitness and host interaction, yet these same traits may contribute to disease under

Table 5 Putative virulence-associated determinants identified in *Pantoea agglomerans* R6

Virulence Factor	Annotation	Accession number
Secretion system		
gspE	Type II secretion protein	MBE5680744.1
dotU	Type IV secretion protein	MBE5680935.1, MBE5682918.1
dotU, *hcp*, *tssA*, *tssE*, *tssF*, *tssG*, *tssH cipV*, *tssJ*, *tssK*, *tssM*, *tagF*, *tagH*, *vasK*, *vgrG.*	Type VI secretion protein	MBE5683064.1, MBE5682920.1, MBE5682755.1, MBE5683061.1, MBE5682488.1, MBE5682827.1, MBE5682489.1, MBE5682754.1, MBE5682490.1, MBE5682753.1, MBE5682491.1, MBE5682921.1, MBE5682752.1, MBE5683066.1, MBE5682917.1, MBE5682751.1, MBE5682922.1 MBE5683065.1, MBE5680934.1, MBE5683063.1, MBE5680933.1, MBE5682484.1, MBE5683062.1, MBE5682756.1,
Adhesion		
fha	Filamentous hemagglutinin	MBE5682251.1, MBE5682844.1, MBE5683682.1
flp	Prepilin peptidase-dependent pilin	MBE5680743.1
bfpP	Prepilin peptidase	MBE5681717.1
ppdB	Prepilin-type cleavage/methylation domain-containing protein	MBE5682350.1
ppdD	Prepilin-type N- terminal cleavage/ methylation domain-containing protein	MBE5682353.1
Motility		
flk, *motB*, *flgA*, *flgB*, *flgC*, *flgD*, *flgF*, *flgG*, *flgH*, *flgI*, *flgJ*, *flgK*, *flgL*, *flgN*, *flhA*, *flhB*, *flhE*, *fliE*, *fliF*, *fliG*, *fliH*, *fliJ*, *fliK*, *fliL*, *fliM*, *fliN*, *fliO*, *fliP*, *fliQ*, *fliR*, *fliS*, *fliT*, *fliZ*	Flagellar	MBE5680518.1, MBE5682272.1, MBE5680678.1, MBE5680677.1, MBE5680676.1, MBE5680675.1, MBE5680674.1, MBE5680673.1, MBE5680672.1, MBE5680671.1, MBE5680670.1, MBE5680669.1, MBE5680668.1, MBE5680680.1, MBE5682283.1, MBE5682282.1, MBE5682284.1, MBE5683288.1, MBE5682878.1, MBE5682879.1, MBE5682880.1, MBE5682881.1, MBE5682883.1, MBE5682884.1, MBE5682885.1, MBE5682886.1, MBE5682887.1, MBE5682888.1, MBE5682889.1, MBE5682890.1, MBE5682891.1, MBE5683289.1, MBE5683290.1, MBE5683435.1
pilT	Twitching motility Protein	MBE5680299.1, MBE5682587.1
Iron Uptake System		
fepA	TonB-dependent siderophore receptor	MBE5679880.1, MBE5681370.1, MBE5683351.1, MBE5683646.1
fepB	Fe2+-enterobactin ABC transporter substrate-binding protein gene	MBE5679873.1
fepD	Fe (3+)-siderophore ABC transporter permease	MBE5679875.1
fepG	Iron-enterobactin ABC transporter permease	MBE5679876.1
fhuF	Siderophore reductase	MBE5683455.1
entF	Enterobactin synthase subunit F	MBE5679877.1
entS	Enterobactin Transporter	MBE5679874.1
Toxins		
hha, *shlB*, *fhaC*, *hecB*	Hemolysin	MBE5682843.1, MBE5683450.1, MBE5683873.1
hlyIII	HemolysinIII	MBE5681530.1
Toxin-antitoxin System		

Table 5 (continued)

Virulence Factor	Annotation	Accession number
symE	Type I toxin system	MBE5683009.1, MBE5683920.1
hicA, parE, ratA, relE, yhaV,	Type II toxin system	MBE5680732.1, MBE5681208.1, MBE5681821.1, MBE5681827.1, MBE5682136.1, MBE5683011.1, MBE5683044.1, MBE5683370.1, MBE5683581.1, MBE5683582.1, MBE5683821.1
dinJ, hicB, hipB, mqsA, phd, prlF, relB, yefM	Type II antitoxin system	MBE5680731.1, MBE5682024.1, MBE5682025.1, MBE5683005.1, MBE5682135.1, MBE5680846.1, MBE5683369.1,
ghoS	Type V antitoxin system	MBE5682429.1

Genes were predicted through BLAST searches against the VFDB and genome annotation and classified according to functional categories. Adapted from Guevarra, Magez [62]

permissive conditions. At the same time, phenotypes such as EPS production and nutrient solubilisation reflect functions typically exploited in agricultural biotechnology. This combination underscores that the utility of *Pantoea*-based inoculants must be evaluated at the strain level, balancing beneficial activities with genomic indicators of risk.

The virulence- and host-interaction-associated genes identified in R6 represent homologues of loci previously reported in pathogenic *Pantoea* strains and should therefore be interpreted as indicators of genomic potential supported by comparative genomics and phenotypic characterisation, rather than experimentally validated causal determinants.

Conclusion

The integration of gene mining and comparative genomic analyses demonstrates that *Pantoea agglomerans* R6 harbours a complex repertoire of putative pathogenicity determinants. These include secretion system effectors, adhesion and motility-related genes, siderophore-mediated iron acquisition systems, and toxin-associated loci, many of which are localized within genomic islands and appear to have been horizontally acquired. The presence of these features within the accessory genome, combined with evidence of an open pan-genome, highlights the genomic plasticity of R6 and its capacity to adapt across diverse ecological niches. These findings are consistent with broader observations in *Pantoea* spp., where the balance between mutualism and pathogenicity is often mediated by horizontally transferred elements and regulatory flexibility.

While comparative analyses strongly suggest that *P. agglomerans* R6 possesses the genetic potential for opportunistic pathogenicity, functional validation is required to determine the specific contributions of individual virulence factors versus those associated with commensal or plant growth-promoting roles. This underscores the need for integrated genomic and experimental approaches to more precisely define the molecular basis of host interactions in *P. agglomerans*, providing a critical foundation for both risk assessment and the informed exploitation of beneficial strains in sustainable agriculture.

For future work, EMS mutagenesis could be employed to generate random mutations in *P. agglomerans* R6, facilitating the identification of genes involved in virulence, stress response, and host adaptation [116, 117]. Subsequent whole-genome sequencing of mutant strains, coupled with comparative genomics, would allow precise determination of the genetic determinants underpinning pathogenic potential. Complementary approaches, including metabolomic profiling and in planta functional assays, would provide insights into how these genetic factors influence bacterial physiology, host interactions,

and virulence. Together, these strategies offer a comprehensive framework for understanding and predicting the pathogenicity of *P. agglomerans* R6, while distinguishing traits that differentiate opportunistic pathogens from non-pathogenic or beneficial strains.

Overall, this work establishes a solid foundation for future studies, emphasizing both the evolutionary mechanisms that shape the ecological versatility of R6 and the experimental avenues necessary to unravel its pathogenic potential.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-026-12875-9.

Supplementary Material 1.

Authors' contributions

D.E. Holman conceived the study, performed the analyses, prepared the figures, and wrote the original manuscript. A. Klein and M. Keyster supervised the project, contributed to methodology refinement, provided critical editorial input, and secured funding. All authors read and approved the final manuscript.

Funding

D.E. Holman received funding from the National Research Foundation (NRF) of South Africa (PMDS240614227608). The laboratory of Ashwil Klein is funded by the NRF (Grant numbers: 107023, 115280). The laboratory of Marshall Keyster is funded by a Sibusiso Bengu Grant (Grant number: GB0200088). Ashwil Klein and Marshall Keyster also obtained joint funding from the Department of Science and Innovation-Technology Innovation Agency (Grant number: GB0200090).

Data availability

The whole-genome shotgun sequence of **Pantoea agglomerans** strain R6 has been deposited in the NCBI GenBank database under accession number JACRVA000000000 (version 1), BioProject PRJNA647595, BioSample SAMN15595211, and Sequence Read Archive (SRA) accession SRR12904959. All additional data generated or analysed during this study are included in this published article and its supplementary material.

Declarations

Ethics approval and consent to participate

Not applicable. This article does not contain any studies with human participants or animals performed by any of the authors.

Consent for publication

All authors have reviewed and approved the final version of the manuscript and consent to its publication.

Competing interests

The authors declare no competing interests.

Received: 13 January 2026 / Accepted: 15 April 2026
Published online: 24 April 2026

References

1. Asai N, et al. Pantoea dispersa bacteremia in an immunocompetent patient: a case report and review of the literature. J Med Case Rep. 2019;13(1):33.
2. Gavini F, et al. Transfer of Enterobacter agglomerans (Beijerinck 1888) Ewing and Fife 1972 to Pantoea gen. nov. as Pantoea agglomerans comb, nov. and Description of Pantoea dispersa sp. nov. Int J Syst Evol MicroBiol. 1989;39(3):337–45.
3. Medrano E, Bell A. Role of Pantoea agglomerans in opportunistic bacterial seed and boll rot of cotton (Gossypium hirsutum) grown in the field. J Appl Microbiol. 2007;102(1):134–43.
4. Walterson AM, Stavrinides J. Pantoea: insights into a highly versatile and diverse genus within the Enterobacteriaceae. FEMS Microbiol Rev. 2015;39(6):968–84.
5. De Maayer P, et al. The large universal Pantoea plasmid LPP-1 plays a major role in biological and ecological diversification. BMC Genomics. 2012;13:1–12.
6. Quecine M, et al. Sugarcane growth promotion by the endophytic bacterium Pantoea agglomerans 33.1. Appl Environ Microbiol. 2012;78(21):7511–8.
7. Cherif-Silini H, et al. Durum wheat stress tolerance induced by endophyte Pantoea agglomerans with genes contributing to plant functions and secondary metabolite arsenal. Int J Mol Sci. 2019;20(16):3989.
8. Cabrefiga Olamendi J. Fire blight (Erwinia amylovora) of rosaceous plants. Pathogen virulence and selection and characterization of biological control agents. Universitat de Girona. 2004.
9. Azizi MMF et al. The emergence of Pantoea species as a future threat to global rice production. J Plant Prot Res. 2020:327–35.
10. Singh M, et al. Endophytic bacteria in plant disease management, in Microbial endophytes. Elsevier. 2020:61–89.
11. Cruz AT, Cazacu AC, Allen CH. Pantoea agglomerans, a plant pathogen causing human disease. J Clin Microbiol. 2007;45(6):1989–92.
12. Wu H-J, Wang AH, Jennings MP. Discovery of virulence factors of pathogenic bacteria. Curr Opin Chem Biol. 2008;12(1):93–101.
13. Shyntum DY, et al. Pantoea ananatis utilizes a type VI secretion system for pathogenesis and bacterial competition. Mol Plant Microbe Interact. 2015;28(4):420–31.
14. Moretti C, et al. Synergistic interaction between the type III secretion system of the endophytic bacterium Pantoea agglomerans DAPP-PG 734 and the virulence of the causal agent of olive knot Pseudomonas savastanoi pv. savastanoi DAPP-PG 722. Mol Plant Pathol. 2021;22(10):1209–25.
15. Istain B, et al. Draft genome sequence of the putative endophytic bacterium *Pantoea agglomerans* R6, associated with Lactuca serriola from South Africa. Microbiol Resour Announc. 2021;10(10):10.1128/mra. 00023-21.
16. Tian L, et al. Antagonism of rhizosphere Streptomyces yangpuensis CM253 against the pathogenic fungi causing corm rot in saffron (Crocus sativus L). Pathogens. 2022;11(10):1195.
17. Shields P, Cathcart L. Motility test medium protocol. Am Soc Microbiol. 2011;214:215.
18. Abràmoff MD, Magalhães PJ, Ram SJ. Image processing with ImageJ. Biophotonics Int. 2004;11(7):36–42.
19. Elnahas MO, et al. Isolation, characterization and bioactivities of an extracellular polysaccharide produced from Streptomyces sp. MOE6. Molecules. 2017;22(9):1396.
20. Coffey BM, Anderson GG. Biofilm formation in the 96-well microtiter plate, in Pseudomonas methods and protocols. Springer. 2014:631–41.
21. Wang J-R, et al. Codon optimization significantly improves the expression level of α-amylase gene from Bacillus licheniformis in Pichia pastoris. Biomed Res Int. 2015;2015(1):248680.
22. Lee K-W, et al. Expression of alpha-amylase gene from Bacillus licheniformis in Lactobacillus brevis 2.14. Prev Nutr Food Sci. 2008;13(3):190–5.
23. Reiner K. Catalase test protocol. American Society for Microbiology. In ASM Conference for Undergraduate Educators. 2010.
24. Coniglio RO, et al. Screening of new secretory cellulases from different supernatants of white rot fungi from Misiones, Argentina. Mycology. 2017;8(1):1–10.
25. Toshy Agrawal TA, Kotasthane A. Chitinolytic assay of indigenous Trichoderma isolates collected from different geographical locations of Chhattisgarh in Central India. 2012.
26. Carrazco-Palafox J, et al. Improved method for qualitative screening of lipolytic bacterial strains. MethodsX. 2018;5:68–74.
27. Etminani F, Harighi B. Isolation and identification of endophytic bacteria with plant growth promoting activity and biocontrol potential from wild pistachio trees. plant Pathol J. 2018;34(3):208.
28. Mishra KP. Lead exposure and its impact on immune system: A review. Toxicol In Vitro. 2009;23(6):969–72.
29. Gordon SA, Weber RP. COLORIMETRIC ESTIMATION OF INDOLEACETIC ACID. Plant Physiol. 1951;26(1):192–5.
30. Alexander D, Zuberer D. Use of chrome azurol S reagents to evaluate siderophore production by rhizosphere bacteria. Biol Fertil Soils. 1991;12(1):39–45.

31. Ishak FN et al. Preliminary analysis of endophytic plant growth promoting (pgp) Methylobacterium sp. isolated from palm oil (Elaeis guineensis) leaves. in IOP Conference Series: Earth and Environmental Science. 2021. IOP Publishing.
32. Dinesh R, et al. Isolation and characterization of potential Zn solubilizing bacteria from soil and its effects on soil Zn release rates, soil available Zn and plant Zn content. Geoderma. 2018;321:173–86.
33. Grant JR, et al. Proksee: in-depth characterization and visualization of bacterial genomes. Nucleic Acids Res. 2023;51(W1):W484–92.
34. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068–9.
35. Page AJ, et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31(22):3691–3.
36. Petkau A, et al. Interactive microbial genome visualization with GView. Bioinformatics. 2010;26(24):3125–6.
37. Ding W, Baumdicker F, Neher RA. panX: pan-genome analysis and exploration. Nucleic Acids Res. 2018;46(1):e5–5.
38. Bertelli C, et al. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Res. 2017;45(W1):W30–5.
39. Bertelli C, Brinkman FS. Improved genomic island predictions with IslandPath-DIMOB. Bioinformatics. 2018;34(13):2161–7.
40. Arndt D, et al. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 2016;44(W1):W16–21.
41. Couvin D, et al. CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. Nucleic Acids Res. 2018;46(W1):W246–51.
42. Alcock BP, et al. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. Nucleic Acids Res. 2023;51(D1):D690–9.
43. Guttenplan SB, Kearns DB. Regulation of flagellar motility during biofilm formation. FEMS Microbiol Rev. 2013;37(6):849–71.
44. Zhou D, et al. Plant immune receptors interact with hemibiotrophic pathogens to activate plant immunity. Front Microbiol. 2023;14:1252039.
45. Vilas Boas D, et al. The role of flagellum and flagellum-based motility on Salmonella enteritidis and Escherichia coli biofilm formation. Microorganisms. 2024;12(2):232.
46. Czaczyk K, Myszka K. Biosynthesis of extracellular polymeric substances (EPS) and its role in microbial biofilm formation. Pol J Environ Stud. 2007;16(6).
47. Delik E, Eroğlu B, Tefon-Öztürk BE. Evaluation of the in vitro effects of concentrations of antibiotics on three Enterobacteriaceae isolates. World J Microbiol Biotechnol. 2024;40(2):73.
48. Ayoub Moubareck C. Polymyxins and bacterial membranes: a review of antibacterial activity and mechanisms of resistance. Membranes. 2020;10(8):181.
49. Mardaneh J, Dallal MMS. Isolation, identification and antimicrobial susceptibility of Pantoea (Enterobacter) agglomerans isolated from consumed powdered infant formula milk (PIF) in NICU ward: First report from Iran. Iran J Microbiol. 2013;5(3):263.
50. Büyükcam A, et al. Clinical and microbiological characteristics of Pantoea agglomerans infection in children. J Infect Public Health. 2018;11(3):304–9.
51. Hall CW, Mah T-F. Molecular mechanisms of biofilm-based antibiotic resistance and tolerance in pathogenic bacteria. FEMS Microbiol Rev. 2017;41(3):276–301.
52. Sadiq FA, et al. Community-wide changes reflecting bacterial interspecific interactions in multispecies biofilms. Crit Rev Microbiol. 2021;47(3):338–58.
53. Thomé A, et al. Dichromatic and monochromatic laser radiation effects on antibiotic resistance, biofilm formation, and division rate of Pantoea agglomerans. Laser Phys. 2018;28(6):065606.
54. Yao B, et al. Diverse virulence attributes of Pantoea alfalfae sp. nov. CQ10 responsible for bacterial leaf blight in alfalfa revealed by genomic analysis. Int J Mol Sci. 2023;24(9):8138.
55. Zamorano A, et al. Pantoea agglomerans-induced dieback in pistachio in Chile. Horticulturae. 2022;8(11):1052.
56. Maalouf M, et al. Polarization of femtosecond laser for titanium alloy nanopatterning influences osteoblastic differentiation. Nanomaterials. 2022;12(10):1619.
57. Brokate O, Papenbrock J, Turcios AE. Biofilm-forming microorganisms in the rhizosphere to improve plant growth: Coping with abiotic stress and environmental pollution. Appl Soil Ecol. 2024;202:105591.
58. Danhorn T, Fuqua C. Biofilm formation by plant-associated bacteria. Annu Rev Microbiol. 2007;61(1):401–22.
59. Goel C, et al. Antibiofilm Potential of Alpha-Amylase from a Marine Bacterium, Pantoea agglomerans. Can J Infect Dis Med Microbiol. 2022;2022(1):7480382.
60. An X-H, et al. Physiological and transcriptomic analyses of response of walnuts (Juglans regia) to Pantoea agglomerans infection. Front Plant Sci. 2023;14:1294643.
61. Zhang X, et al. Colonization of endophyte Pantoea agglomerans YS19 on host rice, with formation of multicellular symplasmata. World J Microbiol Biotechnol. 2010;26(9):1667–73.
62. Guevarra RB, et al. Comprehensive genomic analysis reveals virulence factors and antibiotic resistance genes in Pantoea agglomerans KM1, a potential opportunistic pathogen. PLoS ONE. 2021;16(1):e0239792.
63. Benali S, Mohamed B, Eddine HJ. Virulence strategies of phytopathogenic bacteria and their role in plant disease pathogenesis. Afr J Microbiol Res. 2014;8(30):2809–15.
64. Lin Y, et al. Extracellular amylase is required for full virulence and regulated by the global posttranscriptional regulator RsmA in Xanthomonas campestris Pathovar campestris. Phytopathology®. 2021;111(7):1104–13.
65. Sharma P, et al. Nucleases of bacterial pathogens as virulence factors, therapeutic targets and diagnostic markers. Int J Med Microbiol. 2019;309(8):151354.
66. Stehr F, et al. Microbial lipases as virulence factors. J Mol Catal B: enzymatic. 2003;22(5–6):347–55.
67. Secor PR, et al. The depletion mechanism actuates bacterial aggregation by exopolysaccharides and determines species distribution & composition in bacterial aggregates. Front Cell Infect Microbiol. 2022;12:869736.
68. Sulja A, et al. Comparative genomics to examine the endophytic potential of Pantoea agglomerans DAPP-PG 734. BMC Genomics. 2022;23(1):742.
69. Forte E, et al. The terminal oxidase cytochrome bd promotes sulfide-resistant bacterial respiration and growth. Sci Rep. 2016;6(1):23788.
70. Borisov VB, et al. Cytochrome bd oxidase from Escherichia coli displays high catalase activity: an additional defense against oxidative stress. FEBS Lett. 2013;587(14):2214–8.
71. Bouffartigues E, et al. Sucrose favors Pseudomonas aeruginosa pellicle production through the extracytoplasmic function sigma factor SigX. FEMS Microbiol Lett. 2014;356(2):193–200.
72. Costa Oliveira BE, et al. The route of sucrose utilization by Streptococcus mutans affects intracellular polysaccharide metabolism. Front Microbiol. 2021;12:636684.
73. Bowen W, Koo H. Biology of Streptococcus mutans-derived glucosyltransferases: role in extracellular matrix formation of cariogenic biofilms. Caries Res. 2011;45(1):69–86.
74. Suresh S, Alva PP, Premanath R. Modulation of quorum sensing-associated virulence in bacteria: carbohydrate as a key factor. Arch Microbiol. 2021;203(5):1881–90.
75. Bogs J, Geider K. Molecular analysis of sucrose metabolism of Erwinia amylovora and influence on bacterial virulence. J Bacteriol. 2000;182(19):5351–8.
76. Saticioglu IB, Duman M, Altun S. Antimicrobial resistance and molecular characterization of Pantoea agglomerans isolated from rainbow trout (Oncorhynchus mykiss) fry. Microb Pathog. 2018;119:131–6.
77. Kühn J, et al. Glucose-but not rice-based oral rehydration therapy enhances the production of virulence determinants in the human pathogen Vibrio cholerae. PLoS Negl Trop Dis. 2014;8(12):e3347.
78. Delétoile A, et al. Phylogeny and identification of Pantoea species and typing of Pantoea agglomerans strains by multilocus gene sequencing. J Clin Microbiol. 2009;47(2):300–10.
79. Meena M, et al. Virulence factors and their associated genes in microbes, in New and future developments in microbial biotechnology and bioengineering. Elsevier. 2019:181–208.
80. Wilson DM, Ajl S. Metabolism of L-rhamnose by Escherichia coli: II. The phosphorylation of L-rhamnulose. J Bacteriol. 1957;73(3):415–20.
81. Ye S, Kim J-w, Kim SR. Metabolic engineering for improved fermentation of L-arabinose. J Microbiol Biotechnol. 2019;29(3):339–46.
82. Dahiya P, et al. Comparative Genomic and Functional Analyses for Insights into Pantoea agglomerans Strains Adaptability in Diverse Ecological Niches. Curr Microbiol. 2024;81(8):254.
83. Yakubu I, Kong HG. The relationship between the sugar preference of bacterial pathogens and virulence on plants. plant Pathol J. 2023;39(6):529.
84. Winter H, Huber SC. Regulation of sucrose metabolism in higher plants: localization and regulation of activity of key enzymes. CRC Crit Rev Plant Sci. 2000;19(1):31–67.
85. Jung H-I, et al. Isolation and characteristics of bacteria showing biocontrol and biofertilizing activities. J Life Sci. 2007;17(12):1682–8.
86. Bankevich A, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455–77.

87. Taghavi S, et al. Genome sequence of the plant growth promoting endophytic bacterium Enterobacter sp. 638. PLoS Genet. 2010;6(5):e1000943.
88. Waack S, et al. Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models. BMC Bioinformatics. 2006;7(1):142.
89. Li F, et al. Patatin-like phospholipase CapV in Escherichia coli-morphological and physiological effects of one amino acid substitution. NPJ biofilms microbiomes. 2022;8(1):39.
90. Tettelin H, et al. Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial pan-genome. Proc Natl Acad Sci. 2005;102(39):13950–5.
91. Rouli L, et al. The bacterial pangenome as a new tool for analysing pathogenic bacteria. New microbes new infections. 2015;7:72–85.
92. Duval M, Cossart P. Small bacterial and phagic proteins: an updated view on a rapidly moving field. Curr Opin Microbiol. 2017;39:81–8.
93. Juhas M, et al. Genomic islands: tools of bacterial horizontal gene transfer and evolution. FEMS Microbiol Rev. 2009;33(2):376–93.
94. Andrews SJ, Rothnagel JA. Emerging evidence for functional peptides encoded by short open reading frames. Nat Rev Genet. 2014;15(3):193–204.
95. Storz G, Wolf YI, Ramamurthi KS. Small proteins can no longer be ignored. Annu Rev Biochem. 2014;83(1):753–77.
96. Orr MW, et al. Alternative ORFs and small ORFs: shedding light on the dark proteome. Nucleic Acids Res. 2020;48(3):1029–42.
97. Mott ML, Berger JM. DNA replication initiation: mechanisms and regulation in bacteria. Nat Rev Microbiol. 2007;5(5):343–54.
98. Katayama T, et al. Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC. Nat Rev Microbiol. 2010;8(3):163–70.
99. Garai P, Blanc-Potard A. Uncovering small membrane proteins in pathogenic bacteria: regulatory functions and therapeutic potential. Mol Microbiol. 2020;114(5):710–20.
100. Wolański M, et al. oriC-encoded instructions for the initiation of bacterial chromosome replication. Front Microbiol. 2015;5:735.
101. Nisa M-U, et al. The plant DNA damage response: signaling pathways leading to growth inhibition and putative role in response to stress conditions. Front Plant Sci. 2019;10:653.
102. Kastenmayer JP, et al. Functional genomics of genes with small open reading frames (sORFs) in S. cerevisiae. Genome Res. 2006;16(3):365–73.
103. Guo J, et al. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. Microbiome. 2021;9(1):37.
104. Vernikos GS, Parkhill J. Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the Salmonella pathogenicity islands. Bioinformatics. 2006;22(18):2196–203.
105. Brown CL, et al. mobileOG-db: a manually curated database of protein families mediating the life cycle of bacterial mobile genetic elements. Appl Environ Microbiol. 2022;88(18):e00991–22.
106. Koskella B, Hernandez CA, Wheatley RM. Understanding the impacts of bacteriophage viruses: from laboratory evolution to natural ecosystems. Annual Rev Virol. 2022;9(1):57–78.
107. Abdelsattar A, et al. Bacteriophages: From isolation to application. Curr Pharm Biotechnol. 2022;23(3):337–60.
108. Villa TG, et al. Horizontal gene transfer between bacteriophages and bacteria: antibiotic resistances and toxin production. Horizontal Gene Transfer: Breaking Borders Between Living Kingdoms. Springer. 2019:97–142.
109. Linares R, et al. Structure, function and assembly of the long, flexible tail of siphophages. Curr Opin Virol. 2020;45:34–42.
110. Martinez J, Baquero F. Mutation frequencies and antibiotic resistance. Antimicrob Agents Chemother. 2000;44(7):1771–7.
111. Lerminiaux NA, Cameron AD. Horizontal transfer of antibiotic resistance genes in clinical environments. Can J Microbiol. 2019;65(1):34–44.
112. Benveniste R, Davies J. Aminoglycoside antibiotic-inactivating enzymes in actinomycetes similar to those present in clinical isolates of antibiotic-resistant bacteria. Proc Natl Acad Sci. 1973;70(8):2276–80.
113. Martinez JL, et al. A global view of antibiotic resistance. FEMS Microbiol Rev. 2008;33(1):44–65.
114. Martínez JL, Baquero F, Andersson DI. Beyond serial passages: new methods for predicting the emergence of resistance to novel antibiotics. Curr Opin Pharmacol. 2011;11(5):439–45.
115. Lin C-S, et al. RssAB-FlhDC-ShlBA as a major pathogenesis pathway in Serratia marcescens. Infect Immun. 2010;78(11):4870–81.
116. Roychowdhury R, Tah J. Mutagenesis—a potential approach for crop improvement. Crop improvement: new approaches and modern techniques. 2013:149–187.
117. Holman DE et al. Ethyl methanesulfonate mutagenesis: advancing bacterial genetics for sustainable agriculture. Folia Microbiologica. 2025:1–11.

Publisher's note