Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genomic Characterization and Environmental Adaptability of Burkholderia cenocepacia Isolated from Commercial Liquid Soap in Bangladesh

Anika Tasnim

anikabcsir3324@bcsir.gov.bd

Industrial Microbiology Research Division, Bangladesh Council of Scientific and Industrial Research

Nizam Uddin

Genomic Research Laboratory, BCSIR

Saiful Islam

Industrial Microbiology Research Division, Bangladesh Council of Scientific and Industrial Research

Altaz Jashim

Department of Biochemistry and Biotechnology, University of Science and Technology

Muhammad Irfanul Islam

Industrial Microbiology Research Division, Bangladesh Council of Scientific and Industrial Research

Article

Keywords: Burkholderia cenocepacia, draft genome sequence, drug resistance, alkaliphile, liquid soap, whole-genome sequencing, antibiotic resistance, Bangladesh

Posted Date: November 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11344362/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Commercial liquid soaps may harbor opportunistic pathogens such as *Burkholderia cepacia* complex (BCC), yet genome-scale characterization of soap-associated BCC isolates remains limited in South Asia.

In this study, a bacterial isolate, designated DG4, was recovered from a commercial liquid detergent and characterized using phenotypic, biochemical, whole-genome, and comparative genomic approaches. The isolate exhibited substantial growth under alkaline conditions, with optimum growth observed at pH 8 and detectable growth up to pH 10, while growth was severely restricted at pH 11–12. The isolate was catalase and lipase-positive and its crude supernatant demonstrated surface-active properties, producing a 1.08 ± 0.10 cm oil-displacement zone and an emulsification index (E24) of 53.07 ± 0.11%. The strain also exhibited strong biofilm-forming ability, with an OD570 of 0.508 ± 0.00. Whole-genome sequencing revealed a 7.89-Mb draft genome comprising 459 contigs, with 67.16% GC content and 40× sequencing coverage. Genome-based taxonomic analyses, including ANI and phylogenetic analysis, identified DG4 as *Burkholderia cenocepacia*, with an ANI of 98.28% against the reference strain *B. cenocepacia* J2315 genome. Multilocus sequence typing revealed a potential novel allelic profile not assigned to a previously recognized sequence type. Genome annotation identified 7,964 coding sequences, including genes associated with stress adaptation, membrane transport, biofilm formation, antimicrobial resistance, virulence, and diverse metabolic pathways. The genome contained multiple resistance determinants, including *adeF*, *amrA*, *qacG*, *OXA-1142*, *PEN-B1*, and a *vanF*-associated *vanH* gene. PathogenFinder predicted a human pathogenicity probability of 0.9819. In addition, numerous genes involved in hydrocarbon degradation, aromatic compound metabolism, desulfurization, and biosurfactant-associated functions were identified, supporting considerable bioremediation potential. AntiSMASH analysis detected 21 biosynthetic gene clusters, while prophage, mobile genetic element, and CRISPR-Cas analyses demonstrated substantial genomic plasticity and defense capacity. Comparative orthologous analysis revealed a conserved core genome of 5,243 gene clusters and only 14 DG4-specific clusters among six *B. cenocepacia* genomes. Commercial liquid detergent can harbor environmentally adapted *B. cenocepacia* with clinically relevant resistance and virulence determinants. These findings emphasize the importance of microbiological quality control and AMR surveillance of personal care products while highlighting the strain's potential for bioremediation research.

1. Introduction

Burkholderia cepacia complex (Bcc) represents a group of environmentally versatile bacteria that occupy an unusual position at the interface of environmental microbiology and human health. Originally recognized as a plant-associated organism, Bcc members are now known to inhabit diverse environments, including soil, plant rhizospheres, freshwater, industrial settings, pharmaceutical products, and personal care products [1, 2]. Today, Bcc contains at least 24 species that are evolutionarily related but genetically distinct [3,4]. Bcc bacteria are widespread in pharmaceutical solutions, hospital equipment, and industrial products such as shampoo, oil, and fuel, and can engage in complex interactions [3,5]. Several

Bcc members are capable of biodegrading xenobiotics and organic pollutants such as crude oil, pesticides, phthalates, and solvents [3]. For instance, *B. vietnamiensis* strain G4 degrades trichloroethylene[6]. Bcc members can colonize the rhizosphere and exhibit PGPR potential in multiple crops, such as pea (*B. ambifaria*), maize (*B. cenocepacia*), rice (*B. vietnamiensis*), and wheat (*B. cepacia* and *B. cenocepacia*), and certain species are additionally able to fix atmospheric nitrogen (N_2)[3,7].

This environmental versatility, however, also creates an important product-safety concern. Bcc organisms are able to persist and multiply in nutrient-poor aquatic environments, including drinking water, rivers, lakes, and industrial water systems [8–11].

This ability allows them to persist in pharmaceutical formulations and consumer products, including antiseptics, cosmetics, shampoos, hand sanitizers, mouthwashes, and liquid soaps [12,13]. Their persistence is particularly concerning because many Bcc strains can tolerate or adapt to antimicrobial agents and disinfectants used in product formulations [14–16]. Benzalkonium chloride, for example, has frequently been associated with Bcc contamination [17]. Mechanisms contributing to this persistence include multidrug efflux systems, enzymatic inactivation of biocides, metabolic utilization of organic compounds, and biofilm formation [18,19]. Thus, products designed to provide antimicrobial or cleansing effects may paradoxically provide a selective environment in which highly adaptable Bcc populations can persist.

The significance of such contamination extends beyond product quality because Bcc is also an important opportunistic pathogen. Bcc infections pose a particular clinical challenge in patients with cystic fibrosis (CF), chronic granulomatous disease, and other conditions characterized by impaired host defense [20]. Although the frequency of Bcc infection is relatively low, infection outcomes can range from asymptomatic colonization to severe and rapidly progressive *cepacia* syndrome, which is associated with high mortality [16–18]. Treatment is further complicated by the intrinsic resistance of many Bcc organisms to multiple antimicrobial classes, including aminoglycosides, quinolones, and β-lactams, as well as their ability to withstand host antimicrobial defenses [21–23]. Consequently, the presence of Bcc in pharmaceutical and consumer products represents not only a microbiological quality-control issue but also a potential reservoir of clinically relevant resistance and virulence determinants.

Owing to their ability to persist under nutrient-deprived conditions, Bcc bacteria most frequently contaminate water systems, posing a considerable risk when pharmaceutical water supplies are affected [24]. Bcc bacteria are inherently difficult to detect and identify because of their specific metabolic requirements and slow growth, and most phenotypic and genotypic methods also have significant limitations in detecting them.

Bcc genomes are sizeable, spanning roughly 6 to 10 Mbp, and are usually partitioned into three replicons. Chromosome 1 is highly conserved and harbors most essential genes, chromosome 2 encodes functions linked to ecological adaptation, and the highly variable megaplasmid pC3 contributes to virulence, resistance, and the biosynthesis of specialized metabolites such as antifungal compounds

[25–27]. An estimated 10% of the Bcc genome has been acquired through horizontal gene transfer, giving rise to a diverse repertoire of mobile genetic elements [27,28]. *B. cenocepacia* harbors an extensive repertoire of virulence determinants, such as adhesins, invasins, toxins, extracellular enzymes, siderophores, and secretion systems, that act together to mediate host tissue attachment, cellular invasion, and evasion of host immune defenses[27,28].

Despite extensive investigation of Bcc in clinical and pharmaceutical settings, genome-scale information on Bcc isolates recovered from commercial consumer products remains comparatively limited, particularly in South Asia. This knowledge gap is important because consumer products can provide a route through which environmentally adapted bacteria carrying clinically relevant traits may persist outside healthcare environments. In this study, we performed comprehensive phenotypic and genomic characterization of *Burkholderia cenocepacia* DG4, an alkaliphilic isolate recovered during routine sterility testing of a commercially available liquid detergent in Bangladesh. We investigated its antimicrobial susceptibility, biofilm and biosurfactant production, genome sequencing, antimicrobial resistance and virulence determinants, mobile genetic elements, prophages, secondary metabolite potential, hydrocarbon-degradation capacity, and phylogenomic relationships. Comparative genomic analysis was further used to examine the evolutionary and functional characteristics of DG4 relative to ecologically diverse *B. cenocepacia* strains. To our knowledge, this represents the first genome-sequenced Bcc isolate recovered from a commercial liquid detergent in Bangladesh.

2. Materials and Methods

2.1 Sample Collection and Bacterial Isolation

A commercially available liquid soap was subjected to microbiological sterility testing following standard protocols. Initially, the soap sample was serially diluted in 0.85% (w/v) saline to a dilution of 10^{-3}. Three Nutrient Agar and three MacConkey Agar plates were spread-plated and incubated at 37°C for 24 h, after which a single colony was observed on both media in only 10^{-1} dilution. The isolate was preserved in 50% (v/v) glycerol at -80°C for long-term storage.

2.2 Phenotypic Characterization

2.2.1 Alkaline Tolerance Assessment

To evaluate alkaline tolerance, the isolate was inoculated into nutrient broth adjusted to pH values ranging from 7.0 to 12.0 and incubated at 37°C for 24 h with shaking (160 rpm). Bacterial growth was quantified by measuring the optical density at 600 nm (OD_{600}) [29].

2.2.2 Catalase Test

For the catalase test, 100 µL of a fresh overnight culture was transferred onto a sterile glass slide and mixed with 3–4 drops of 3% (v/v) H_2O_2. Immediate bubble due to oxygen release was scored as

catalase-positive, and no bubble formation as catalase-negative [30].

2.2.3 Lipid hydrolysis test

Lipolytic activity was evaluated on nutrient agar amended with 0.3 mL of emulsified vegetable oil. The isolate was streaked onto the medium with a sterile loop and incubated at 30°C for 48 h. Plates were flooded with $CuSO_4$ solution for 15 min, decanted, and gently rinsed with distilled water. The development of a bluish-green zone around the bacterial growth was scored as positive for lipid hydrolysis[31].

2.3 Preliminary screening tests for presumptive biosurfactant production

The isolate was recovered from a commercial liquid soap formulation and was capable of surviving and growing under soap-associated conditions, preliminary screening was performed to determine whether the bacterium produced extracellular surface-active compounds. Three complementary qualitative/semi-quantitative assays oil displacement, drop collapse, and emulsification index (E24) were employed to assess presumptive biosurfactant production.

2.3.1 Oil displacement assay

Biosurfactant production was screened by the oil displacement assay. Cell-free supernatant (20 µL) was gently dropped onto a thin layer of crude oil (25 µL) floating on 25 mL of Milli-Q water, and the diameter of the clear zone was measured immediately. Milli-Q water (20 µL) served as the negative control [32].

2.3.2 Drop Collapse Assay

The drop collapse assay was used to screen for biosurfactant production. Wells of a 96-well microtiter plate were coated with 2 µL of light paraffin oil and equilibrated at 37°C for 1 h, after which 5 µL of cell-free supernatant was applied to the oil surface. Distilled water (5 µL) served as the negative control, and drop collapse within 1 min was considered a positive result [33].

2.2.3 Emulsification test (E_{24})

The emulsifying capacity of the biosurfactant was determined by the emulsification index (E24) assay using diesel oil. Cell-free supernatant and diesel oil were mixed in equal volumes (e.g., 2 mL each), vortexed for 2 min, and left undisturbed at room temperature for 24 h. E24 was calculated as [34]:

$$\mathrm{E}_{24}\ (\%) = \frac{\text{Height of emulsification layer}\ (\mathrm{cm})}{\text{Total height of liquid column}\ (\mathrm{cm})} \times 100$$

2.4 Antibiotic Susceptibility Testing and Biofilm Formation Evaluation

Antibiotic susceptibility testing was carried out using the Kirby-Bauer disk diffusion method. The strain was grown overnight in Nutrient Broth, standardized to a 0.5 McFarland turbidity standard, and lawned onto Mueller-Hinton Agar plates. A panel of 11 commercially available antibiotic disks like vancomycin (VA, 30 µg), imipenem (IPM, 10 µg), amoxicillin/clavulanic acid (AMC, 20/10 µg), piperacillin/tazobactam (TZP, 100/10 µg), chloramphenicol (C, 30 µg), levofloxacin (LEV, 5 µg), aztreonam (ATM, 30 µg), trimethoprim/sulfamethoxazole (SXT, 1.25/23.75 µg), azithromycin (AZM, 15 µg), ceftriaxone (CRO, 30 µg), and co-trimoxazole (SXT, 1.25/23.75 µg) were used. The discs were inputted on MHA plates embedded with bacterial lawn and incubated at 37°C for 18–24 hours. The inhibition zones were measured in millimeters after incubation under the conditions described above. Because organism-specific interpretive criteria for disk diffusion of *B. cepacia* complex are limited and current CLSI guidance does not provide applicable disk-diffusion breakpoints for routine categorical interpretation of these results, the observed inhibition-zone diameters were reported as phenotypic measurements rather than being assigned definitive susceptible, intermediate, or resistant categories solely on the basis of CLSI breakpoints.

Biofilm formation was quantified by the crystal violet microtiter assay. LB-grown overnight cultures (adjusted to 1 McFarland) were inoculated (200 µL) into 96-well flat-bottom plates and incubated at 37°C for 24 h. Wells were washed to remove non-adherent cells, stained with 0.1% (w/v) crystal violet, rinsed, and the bound dye was dissolved in 200 µL of 96% ethanol. Absorbance (492 nm) was measured with a microplate reader, using sterile medium as the negative control, and isolates were categorized as non-, weak, moderate, or strong biofilm producers following Stepanović et al. (2007)[35].

2.5 Whole-Genome Sequencing and Bioinformatic Analysis

2.5.1 Genomic DNA Extraction, Library Preparation, and Sequencing

Genomic DNA was extracted from an overnight Nutrient Agar culture (37°C) using the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA) following the manufacturer's protocol. Libraries were prepared with the Illumina DNA Prep kit (Illumina, San Diego, CA, USA) and sequenced on an Illumina NextSeq 2000 platform (2 × 300 bp paired end) at the icddr,b Genome Center, Dhaka, Bangladesh.

2.5.2 Whole Genome Assembly, Annotation, and Identification

Raw sequencing reads were quality-checked with FastQC (v0.11) [36] and adapters and low-quality regions were removed with Trimmomatic (v0.39) [37]. The trimmed reads were assembled de novo using SPAdes (v3.15.4) [38]. The assembly was annotated with Prokka [39], the RAST Server [40] and the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) [41,42], and a circular genome map was generated with the Proksee server [43]. Species-level identification was based on whole-genome k-mer analysis using LINbase server

[44] and KmerFinder (Center for Genomic Epidemiology), and was further confirmed by calculating the average nucleotide identity (ANI) [45].

Whole-genome-based taxonomic classification was performed on the Type (Strain) Genome Server (TYGS; https://tygs.dsmz.de) [46], where dDDH values were calculated by the Genome BLAST Distance Phylogeny (GBDP) method and used to infer minimum evolution trees [47]. Trees were visualized in MEGA-X [48], and sequence types were determined with MLST 2.0 (Center for Genomic Epidemiology) [49].

2.5.3 Metabolic Pathways, and Secondary Metabolic Gene Clusters Investigation

Functional subsystems were assigned using the RAST server[40], and COG and KEGG annotations were obtained from the Type Strain Genome Database (https://gctype.wdcm.org/)[50,51] to infer metabolic pathways. Biosynthetic gene clusters were predicted with antiSMASH v8.0[52], and CAZymes were identified with dbCAN3[53].

2.5.4 Antimicrobial Resistance Genes, Virulence Factor Genes, and Pathogenicity Investigation

ARGs were predicted using the Resistance Gene Identifier (RGI) in CARD with default settings, considering only Perfect and Strict hits. VFGs were screened against the VFDB[54], and human pathogenic potential was assessed using PathogenFinder v1.1[55].

2.5.5 Genomic island, plasmid, prophage and CRISPR Cas system identification

Genomic islands were predicted with IslandViewer [56], and plasmid-related sequences were screened using PlasmidFinder [57] and MOB-suite[58]. Prophages were identified with VirSorter[59] and PHASTEST[60], MGEs were annotated using mobileOG-db[61], and CRISPR arrays were detected with CRISPRDetect[62].

2.6 Comparative Orthologous Cluster Analysis

Comparative orthologous cluster analysis was performed using OrthoVenn[363]. Due to the limitation of OrthoVenn3 to analyze a maximum of six genomes per analysis, six Burkholderia cenocepacia genomes were randomly selected based on their reported sample sources to represent different ecological origins. The selected genomes included the study isolate BCSIR_DG4 and five reference strains: TY921_sputum, 2008yc657_Clinical, PS27_Soil, toggle4_Clinical, and NML110041_Environment. The predicted protein sequences of these six genomes were uploaded to the OrthoVenn3 server using its default analysis and visualization parameters. Orthologous protein clusters were identified through an all-against-all sequence similarity search followed by Markov Cluster Algorithm (MCL) analysis using the default settings. The resulting orthologous clusters were used to characterize shared and unique protein clusters among the selected genomes, including core, accessory, and strain-specific gene sets. The

comparative clustering results were visualized using Venn/flower plots, UpSet plots, and pairwise cluster-sharing heatmaps.

3 Results

3.1 Isolation and Phenotypic Properties

The isolate produced creamy-white, circular colonies with smooth margins and a slightly raised surface on nutrient agar (Fig. 1A) and Pure sub-culture on nutrient agar (Fig. 1B). Furthermore, the strain exhibited pronounced alkaliphilic characteristics, maintaining substantial growth in nutrient broth at pH values up to 10, highlighting its remarkable adaptability to highly alkaline environments. The isolate showed good growth between pH 7–9, with the highest 24-h OD_{600} at pH 8 (0.364) (Supp Fig. 1). Growth decreased markedly at pH 10 and was almost completely inhibited at pH 11–12. Thus, the organism appears to tolerate moderately alkaline conditions but has substantially reduced growth under strongly alkaline conditions.

The isolate exhibited positive lipase activity, evidenced by a characteristic color change from yellow to whitish-blue following $CuSO_4$ treatment, indicating the secretion of extracellular lipase (Fig. 1C). It also tested positive for catalase, as demonstrated by immediate oxygen bubble formation upon exposure to hydrogen peroxide, confirming the enzymatic decomposition of H_2O_2 into water and oxygen (Fig. 1D). These biochemical traits, together with the observed morphological characteristics, are consistent with those of environmentally adapted Gram-negative bacteria.

3.2 Antibiotic Susceptibility Profile and Biofilm forming Ability

The antibiotic susceptibility profile of *B. cenocepacia* DG4 was evaluated using the Kirby-Bauer disk diffusion method against twelve antibiotics representing different antimicrobial classes (Fig. 2). The inhibition-zone diameters obtained for the tested antimicrobial agents are presented in Table 1. The isolate exhibited variable inhibition responses across the antimicrobial panel. No visible inhibition zone was observed for amikacin, ampicillin, vancomicin and imipenem, while a relatively small inhibition zone was observed for amoxicillin/clavulanate (7.5 mm). In contrast, measurable inhibition zones were observed for azithromycin (16.0 mm), levofloxacin (15.5 mm), chloramphenicol (23.5 mm), ciprofloxacin (18.0 mm), aztreonam (24.0 mm), cotrimoxazole (22.5 mm), and piperacillin/tazobactam (33.0 mm).

Because organism-specific interpretive criteria for disk diffusion of *B. cepacia* complex are limited, the measured inhibition zones were reported as phenotypic observations rather than being categorically classified as susceptible, intermediate, or resistant solely on the basis of current CLSI breakpoints. Thus, the absence or reduction of inhibition observed for particular antimicrobial agents indicates a reduced phenotypic response under the experimental conditions but should not be interpreted as a definitive clinical resistance category without an appropriate validated interpretive framework.

The isolate exhibited an average optical density (OD) of 0.508 ± 0.00 at 570 nm following the crystal violet biofilm assay, which was significantly higher than the calculated cut-off value (ODcut = 0.0917). According to the classification criteria described by Stepanović et al. (2000), this isolate was designated as a strong biofilm former ($OD > 4 \times OD_{(}cut_{)}$). The ability of bacteria to form biofilm is one of the mechanisms of resistance to survive in the presence of selective environmental pressure, like surfactants, organic solvents etc.

Table 1
Phenotypic antimicrobial response of *B. cenocepacia DG4* isolate based on disk-diffusion inhibition-zone measurements.

Antibiotic	Abbreviation	Zone of inhibition (mm)	Phenotypic observation
Azithromycin 30 µg	AZM	16.0	Inhibition observed
Levofloxacin 5 µg	LEV	15.5	Inhibition observed
Chloramphenicol 30 µg	C	23.5	Inhibition observed
Ciprofloxacin 30 µg	CIP	18.0	Inhibition observed
Amikacin 30 µg	AK	0.0	No inhibition observed
Ampicillin 30 µg	AMP	0.0	No inhibition observed
Aztreonam 30 µg	AT	24.0	Inhibition observed
Cotrimoxazole 25 µg	SXT	22.5	Inhibition observed
Imipenem 10 µg	IMP	0.0	No inhibition observed
Amoxicillin/clavulanate 30 µg	AMC	7.5	Limited inhibition observed
Piperacillin/tazobactam 100/10 µg	PIT	33.0	Inhibition observed
Vancomicin 30 µg	VA	0.00	No inhibition observed

3.3 Surface-active properties of the Isolate

The biosurfactant-producing ability of the isolate was confirmed through oil displacement and emulsification assays. In the oil displacement test, the cell-free supernatant produced a clear zone with a mean diameter of 1.08 ± 0.10 cm, whereas the Milli-Q water control produced a diameter of 0.50 cm, indicating significant surface activity (Fig. 3). Similarly, the isolate exhibited an E24 value of 53.07 ± 0.11% (mean ± SD, n = 3), demonstrating a strong capacity to emulsify hydrophobic substrates. The observed E24 value, which exceeds 50%, is considered indicative of good emulsifying potential and is comparable to those reported for several biosurfactant-producing bacteria.

3.4 Genome Assembly

To elucidate the genetic basis underlying these phenotypic traits, whole-genome sequencing and comprehensive genomic analyses were subsequently performed. *B. cenocepacia* DG4 possesses a draft genome of 7,888,607 bp (7.89 Mb), 459 contigs with 40× sequencing coverage, 92.1% completeness, and a high GC content of 67.16% (Table 2).

Table 2
Assembly details.

Parameters	***Burkholderia cenocepacia* DG4**
Contigs	459
Genome Size (bp)	7,888,607
GC content (%)	67.16
Contig L50	64
Contig N50	34,760
Coverage (×)	40
Completeness (%)	92.1

3.5 Phylogeny and Average Nucleotide Identity (ANI) Analysis

The whole-genome-based phylogenetic analysis demonstrates that the detergent-derived DG4 is unequivocally positioned within the *B. cepacia* complex (Bcc) and exhibits the closest evolutionary relationship with the reference strain *B. cenocepacia* J2315 (Fig. 4). The DG4-J2315 cluster is supported by a 100% bootstrap value, indicating exceptionally strong statistical confidence in their common ancestry. Furthermore, the near-zero branches length (0.00 substitutions per site) separating the two genomes suggests minimal evolutionary divergence, reflecting a highly conserved core genome. Beyond the DG4-J2315 clade, the phylogenetic topology clearly distinguishes *B. cenocepacia* from other members of the *B. cepacia* complex, including *B. orbicola*, *B. sola*, *B. semiarida*, *B. seminalis*, *B. arboris*, *B. theae*, *B. catarinensis*, *B. contaminans*, *B. lata*, *B. cepacia*, *B. reimsis*, *B. metallica*, *B. anthina*, and *B. territorii*. These taxa occupy progressively more distant branches with longer evolutionary distances (approximately 0.01–0.05 substitutions per site), indicating greater genomic divergence from DG4. The high bootstrap support observed throughout the internal nodes (predominantly 99–100%, with only a few moderately supported nodes demonstrates the robustness and reliability of the inferred phylogenetic relationships.

MLST analysis of *B. cenocepacia* DG4 showed exact matches for all seven housekeeping alleles: *atpD* (15), *gltB* (11), *gyrB* (10), *lepA* (11), *phaC* (6), *recA* (14), and *trpB* (79). However, their combined allelic profile was not assigned to any known sequence type, resulting in a probable novel MLST profile **fo**r DG4 strain.

The detergent-derived isolate DG4 was further validated using whole-genome Average Nucleotide Identity (ANI) analysis implemented in the Proksee platform. The genome exhibited an ANI value of 98.28% relative to the reference genome *B. cenocepacia* strain MSMB384WGS chromosome 2, based on 892 orthologous genomic matches distributed across 2,337 query fragments. This ANI value is substantially higher than the widely accepted bacterial species delineation threshold of 95–96% ANI, unequivocally confirming that DG4 belongs to the species *B. cenocepacia*.

3.6 Genome annotation

Genome annotation predicted 7,964 protein-coding sequences, 55 tRNA genes, and 22 rRNA genes. Functional annotation assigned putative functions to 6,103 proteins (76%), including 1,656 Enzyme Commission (EC) assignments, 1,451 Gene Ontology (GO) annotations, and 1,277 pathway-associated proteins, indicating extensive enzymatic and metabolic capabilities. Notably, 1,861 proteins (23%) remained hypothetical, representing a potential reservoir of novel uncharacterized genes may involve in environmental adaptation and specialized metabolism. Furthermore, the genome harbored 11 antimicrobial resistance genes and 36 virulence-associated factors (Table 3).

Table 3
Genomic features of Burkholderia cenocepacia_DG4 strain.

Parameters	***Burkholderia cenocepacia* DG4**
CDS	7,964
tRNAs	55
rRNAs	22
Hypothetical proteins	1,861
Proteins with functional assignments	6,103
Proteins with EC number assignments	1,656
Proteins with GO assignments	1,451
Proteins with Pathway assignments	1,277
Antibiotic resistance genes	11
Bacterial virulence factors	36

3.7 Subsystem analysis

RAST subsystem annotation revealed a metabolically versatile genome with broad biosynthetic, catabolic, and stress-adaptive capabilities (Fig. 5). Amino acid and derivative metabolism constituted the largest functional category (572 genes), followed by carbohydrate metabolism (374 genes), underscoring the strain's capacity for nutrient assimilation and central carbon metabolism. Substantial numbers of genes were also assigned to cofactors, vitamins, prosthetic groups, and pigments (237

genes), protein metabolism (217 genes), and fatty acid, lipid, and isoprenoid metabolism (164 genes), reflecting extensive biosynthetic and lipid-transforming potential. The presence of numerous genes involved in respiration (150 genes) and aromatic compound metabolism (139 genes) indicates a strong oxidative metabolism and the capacity to degrade structurally complex hydrocarbons. In addition, genes associated with membrane transport (120 genes), stress response (97 genes), iron acquisition and metabolism (84 genes), DNA metabolism (76 genes), and virulence, disease, and defense (71 genes) highlight the organism's adaptability to diverse environmental conditions and its ecological competitiveness. Other functional categories, including RNA metabolism, regulation and cell signaling, cell wall biosynthesis, phosphorus, sulfur and nitrogen metabolism, motility, and phage-associated elements, further support the remarkable metabolic plasticity and environmental resilience of *B. cenocepacia* DG4.

3.8 COG and KEGG functional Annotation

COG analysis revealed broad functional diversity in *B. cenocepacia* DG4 (Fig. 6A). Within Cellular Processes and Signaling, genes related to cell wall/membrane/envelope biogenesis, signal transduction, post-translational modification, defense mechanisms, secretion, and cell motility were prominent, suggesting capacity for environmental sensing and adaptation. In Information Storage and Processing, transcription was the largest category, followed by translation, ribosomal structure and biogenesis, and replication, recombination, and repair. The Metabolism category was dominated by amino acid transport and metabolism, followed by carbohydrate and lipid metabolism, energy production and conversion, inorganic ion and coenzyme metabolism, secondary metabolite biosynthesis, and nucleotide metabolism.

KEGG pathway annotation was consistent with the COG profile, with most genes assigned to Metabolism, particularly global and overview maps, carbohydrate metabolism, amino acid metabolism, xenobiotics biodegradation and metabolism, and cofactor and vitamin metabolism (Fig. 6B). The representation of xenobiotic-related pathways suggests metabolic potential for transforming diverse anthropogenic compounds, while membrane transport and signal transduction pathways may contribute to environmental adaptation. Cellular-process pathways related to motility, cellular community, and transport and catabolism further support the potential of DG4 for persistence in chemically complex, detergent-associated environments.

Figure 6. Functional annotation and metabolic profiling (A) Chart of the Clusters of Orthologous Groups (COG) functional classification, showing strong representation in Metabolism, Cellular Processes and Signaling, and Information Storage and Processing. (B) Histogram of KEGG pathway mapping, which shows the counts of genes in major pathway categories and enrichment in metabolic pathways.

3.9 Biosynthetic Gene Clusters and CAZyme Repertoire Analysis of *B. cenocepacia* DG4

AntiSMASH analysis of the DG4 genome (relaxed strictness) identified 21 putative biosynthetic gene clusters (BGCs) representing 10 distinct biosynthetic classes. Terpene-related clusters were the most abundant (n = 7; six terpene and one terpene-precursor), followed by homoserine lactone (n = 3) and NRPS (n = 3) clusters. The homoserine lactone clusters suggest potential for quorum sensing-based cell-to-cell communication, which may regulate virulence, biofilm formation, and other collective behaviors. The NRPS clusters (Regions 170.1, 306.1, 398.1), together with the NRPS-like/betalactone hybrid (Region 122.1) and the T1PKS (Region 72.1), point to the capacity to produce peptide- and polyketide-derived metabolites, including siderophores and other bioactive compounds. Single clusters were also detected for arylpolyene (Region 440.1), phosphonate (Region 79.1), hydrogen cyanide (Region 107.1), redox-cofactor (Region 71.1), and two RiPP-like clusters (Regions 106.1, 161.1). Arylpolyenes are membrane-associated pigments that may protect against oxidative stress, and hydrogen cyanide is a secondary metabolite with antimicrobial and toxic activity. Only Region 170.1 showed a match to a known cluster, with low similarity to ornibactin (a NRPS-type siderophore involved in iron acquisition). The other 20 regions had no significant homology to characterize BGCs in the MIBiG database, which suggests that the strain harbors largely uncharacterized biosynthetic potential and warrants further investigations for novel bioactive compounds (supp Fig. 2).

CAZyme annotation of *B. cenocepacia* DG4 revealed a carbohydrate-active enzyme repertoire dominated by glycosyltransferases (GTs; 55.56%), followed by glycoside hydrolases (GHs; 36.51%), carbohydrate-binding modules (CBMs; 4.76%), and carbohydrate esterases (CEs; 3.17%) (supp Fig. 3). The predominance of GTs suggests a strong capacity for the biosynthesis of extracellular polysaccharides, cell envelop components, and other glycoconjugates essential for biofilm formation and environmental adaptation. The abundance of GHs indicates considerable metabolic flexibility for carbohydrate utilization, while CBMs and CEs likely enhance substrate recognition and polysaccharide deacetylation, improving carbohydrate degradation efficiency. Notably, the absence of auxiliary activity (AA) and polysaccharide lyase (PL) families suggests that DG4 primarily relies on carbohydrate biosynthesis and hydrolytic metabolism rather than oxidative lignocellulose degradation or lyase-mediated polysaccharide cleavage, consistent with its adaptation to detergent-rich environments.

3.10 Genome Mapping, Antibiotic Resistance, and Pathogenicity

The circular genome map of *B. cenocepacia* DG4 revealed a 7.89 Mb draft genome with a dense and relatively uniform distribution of coding sequences (CDSs), together with tRNA, tmRNA, and rRNA features (Fig. 7). Essential genes associated with DNA replication (*dnaE, recB, recD*), transcription (*rpoB, rpoC*), amino acid biosynthesis (*gltB*), energy metabolism (*gdh*), and nucleotide biosynthesis (*purL*) were identified. Genes associated with environmental adaptation, including efflux-associated proteins (*acrB, mdtC, mexY, tmexD2*), heavy-metal transport, oxidoreductive processes, and DNA repair, were also detected. Metabolic genes including *glgE, glgB, xdhA, fadN*, and *gcvP* indicated diverse carbon and nutrient utilization capabilities. Genes associated with exopolysaccharide biosynthesis and capsule formation were additionally identified, along with numerous hypothetical proteins.

CARD analysis identified several antimicrobial resistance-associated determinants with different matching region in DG4 (Table 4). These included the RND-family efflux-associated genes *adeF* and *amrA*, the SMR-family efflux gene *qacG*, and the beta-lactamase genes *OXA-1142* and *PEN-B1*. A *vanH* gene associated with a *vanF*-type glycopeptide resistance cluster was also detected. The identified determinants were associated with resistance mechanisms including antibiotic efflux, antibiotic inactivation, and target alteration. PathogenFinder v2.0 predicted a human pathogenicity probability of **0.9819** for DG4.

Table 4
Antimicrobial resistance-associated determinants of *B. cenocepacia* DG4.

Antimicrobial Resistance Gene	AMR Gene Family	Drug Class	Resistance Mechanism	% identity of matching region	% length of reference sequence
adeF	resistance-nodulation-cell division (RND) antibiotic efflux pump	fluoroquinolone antibiotic, tetracycline antibiotic	antibiotic efflux	79.77	100.66
qacG	small multidrug resistance (SMR) antibiotic efflux pump	disinfecting agents and antiseptics	antibiotic efflux	40.78	104.67
vanH gene in *vanF* cluster	vanH, glycopeptide resistance gene cluster	glycopeptide antibiotic	antibiotic target alteration	38.02	103.11
adeF	resistance-nodulation-cell division (RND) antibiotic efflux pump	fluoroquinolone antibiotic, tetracycline antibiotic	antibiotic efflux	49.39	100.38
amrA	resistance-nodulation-cell division (RND) antibiotic efflux pump	macrolide antibiotic, aminoglycoside antibiotic	antibiotic efflux	75.38	100.25
OXA-1142	OXA beta-lactamase	penicillin beta-lactam	antibiotic inactivation	98.88	100.00
PEN-B1	PEN-B beta-lactamase	penicillin beta-lactam	antibiotic inactivation	99.62	84.94
adeF	resistance-nodulation-cell division (RND) antibiotic efflux pump	fluoroquinolone antibiotic, tetracycline antibiotic	antibiotic efflux	60.1	99.81
adeF	resistance-nodulation-cell division (RND) antibiotic efflux pump	fluoroquinolone antibiotic, tetracycline antibiotic	antibiotic efflux	42.8	100.38

3.11 Consortium of Bioremediation Potential Genes

Genome annotation revealed an extensive repertoire of catabolic genes associated with the degradation of aliphatic hydrocarbons, aromatic hydrocarbons, polycyclic aromatic hydrocarbons (PAHs), sulfur-

containing petroleum compounds, and halogenated xenobiotics. The genome encodes key hydrocarbon-degrading enzymes, including the alkane hydroxylases *alkB* and multiple copies of *ladA*, phenol hydroxylase (*dmpLMNOP*), the catechol ortho-cleavage pathway (*catABC*), and the protocatechuate branch of the β-ketoadipate pathway (*pcaGH*), suggesting genomic potential for utilization and transformation of aliphatic and aromatic compounds. Genes involved in PAH degradation (*ndoA, antA*), styrene oxidation (*styA*), and alternative aromatic ring-cleavage pathways (*hmgA, hpcB*, and *xlnE*) further broaden its catabolic capacity. The presence of multiple haloalkane dehalogenases (*dhaA*) and desulfurization genes (*dszB*, *dszC*) indicates the ability to detoxify halogenated pollutants and selectively remove sulfur from petroleum-derived compounds. In addition, biosurfactant-associated genes, including *rhlA*, together with *plcN* and *wbbL*, suggest the potential to produce surface-active molecules that enhance hydrocarbon bioavailability and degradation. These genomic features reveal a coordinated metabolic network for the uptake, transformation, and mineralization of structurally diverse organic pollutants, underscoring the promise of *B. cenocepacia* DG4 as a candidate for environmental bioremediation.

Table 5
Bioremediation-associated genes identified in the genome of *B. cenocepacia* DG4.

Locus tag	Gene	Function
Bioremediation-related genes		
DHKPAFHO_00055	*alkB*	Alkane 1-monooxygenase; initiates alkane degradation
DHKPAFHO_00451	*ladA_1*	Long-chain alkane monooxygenase
DHKPAFHO_03371	*ladA_2*	Long-chain alkane monooxygenase
DHKPAFHO_00444	*dhaA_1*	Haloalkane dehalogenase; degradation of halogenated compounds
DHKPAFHO_02688	*dhmA_1*	Haloalkane dehalogenase
DHKPAFHO_01202–01206	*dmpL/M/N/O/P*	Phenol 2-monooxygenase complex
DHKPAFHO_01355	*pcaG*	Protocatechuate 3,4-dioxygenase α-subunit
DHKPAFHO_01356	*pcaH*	Protocatechuate 3,4-dioxygenase β-subunit
DHKPAFHO_01534	*ndoA*	Naphthalene dioxygenase component
DHKPAFHO_02192	*antA*	Anthranilate 1,2-dioxygenase
DHKPAFHO_03575–03578	*andAa/Ab/Ac/Ad*	Anthranilate degradation pathway
DHKPAFHO_03584	*catA2*	Catechol 1,2-dioxygenase
DHKPAFHO_03583	*catB*	Muconate cycloisomerase
DHKPAFHO_03585	*catC*	Muconolactone isomerase
DHKPAFHO_02012	*xlnE*	Gentisate 1,2-dioxygenase
DHKPAFHO_02090	*hmgA*	Homogentisate 1,2-dioxygenase
DHKPAFHO_03058	*hpcB*	3,4-Dihydroxyphenylacetate 2,3-dioxygenase
DHKPAFHO_03132	*styA*	Styrene monooxygenase
DHKPAFHO_03142	*tsaM1*	Toluene sulfonate monooxygenase
DHKPAFHO_01153–01154	*dszC*	Dibenzothiophene monooxygenase; biodesulfurization
DHKPAFHO_01402	*dszB*	Desulfinase; biodesulfurization
DHKPAFHO_00971	*tcpC*	Chlorohydroxyquinol dioxygenase; chlorinated aromatic degradation

Locus tag	Gene	Function
Bioremediation-related genes		
Biosurfactant-related genes		
DHKPAFHO_03282	*rhlA*	3-(3-Hydroxydecanoyloxy)decanoate synthase
DHKPAFHO_00144	*wbbL*	Rhamnosyltransferase; involved in rhamnose-containing glycolipid/LPS synthesis
DHKPAFHO_03150	*plcN_1*	Phospholipase C
DHKPAFHO_03226	*plcN_2*	Phospholipase C
DHKPAFHO_03791	*plcN_3*	Phospholipase C
DHKPAFHO_03833	*plcN_4*	Phospholipase C
DHKPAFHO_03870	*plcN_5*	Phospholipase C
DHKPAFHO_00482	—	Monoacylglycerol lipase
DHKPAFHO_03391	—	Monoacylglycerol lipase

3.12 Genomic Island Analysis

IslandViewer analysis identified multiple predicted genomic island (GI) regions distributed across the *B. cenocepacia* BCSIR_DG4 genome. The predicted regions contained diverse functional genes associated with mobile genetic elements, secretion systems, transport, metabolism, and regulation. A phage-associated region (~ 1.7–34.1 kb) contained genes encoding phage structural, terminase, and lysis-associated proteins, suggesting a prophage-associated region. A larger region (~ 2.66–2.71 Mb) contained integrase, excisionase, partitioning, and conjugative-transfer-associated genes, indicating a candidate mobile genomic region. Another region (~ 4.26–4.28 Mb) contained multiple VirB/Trb proteins and a relaxase, supporting its association with conjugative DNA transfer. Additional regions contained type VI and type III secretion system components, transposases, efflux-associated proteins, and quorum-sensing genes. Overall, these predicted GI regions indicate substantial genomic plasticity and the acquisition of accessory functions potentially contributing to environmental adaptation of DG4.

3.13 Mobile Genetic Element, and Prophage Region Analysis of *B. cenocepacia* DG4

The genome analysis identified 232 mobile genetic element-associated features, predominantly related to phage functions (67), integration/excision (58), and replication, recombination, and repair (53), with additional stability/defense and transfer-related functions (Table 6). The results are illustrated by circular map (Fig. 9).

Table 6
Mobile Genetic Element-Associated Genes Identified by mobileOG-db

Category	Genes/features
Phage	67
Integration/excision	58
Replication/recombination/repair	53
Stability/transfer/defense	30
Transfer	24
Total	**232**

The PHASTEST analysis identified three intact prophage regions, comprising 124 prophage-associated genes and a total of 127 prophage-related features (Table 7). The three prophage regions ranged from 28.1 kb to 41.1 kb in length and all achieved scores exceeding 90, classifying them as intact prophages. The predicted regions showed similarity to phages infecting *Vibrio, Escherichia*, and *Burkholderia* hosts.

Table 7
Predicted Prophage Regions Identified in the Genome

Region	Length	Completeness	Most common phage
1	28.1 kb	Intact	PHAGE_Vibrio_VvAW1_NC_020488
2	41.1 kb	Intact	PHAGE_Escher_ec_EcM_CC0078_NC_041926
3	35.4 kb	Intact	PHAGE_Burkho_ST79_NC_021343

3.14 CRISPR–Cas system analysis of *B. cenocepacia* DG4.

The circular genome map illustrates the distribution of coding sequences (CDSs), CRISPR regions, and associated Cas clusters across the genome of *B. cenocepacia* DG4. As the visualization tool rendered a subset of assembled contigs meeting its internal size/quality thresholds, the displayed genome map corresponds to an estimated ~ 7.2-Mb visualized fraction of the complete 7,888,607 bp (7.89 Mb) draft genome (Table 2). Multiple **Type I CRISPR-associated Cas3 loci** were identified at distinct genomic positions, with several corresponding CRISPR/CAS clusters distributed throughout the chromosome (Fig. 10). Two CRISPR-associated regions, designated **DG4_contig_442_1** and **DG4_contig_425_1**, were also detected. The presence of multiple Cas3-associated loci indicates that the DG4 genome contains an extensive CRISPR-associated defense system, potentially contributing to protection against invading mobile genetic elements such as bacteriophages and plasmids.

3.15 Comparative orthologous gene-content analysis of *B. cenocepacia* DG4

To investigate the genomic diversity and ecological adaptation of the detergent-derived isolate, a comparative genome analysis was performed using six *B. cenocepacia* genomes representing diverse ecological niches, including alkaline detergen t (DG4), sputum (TY921), soil (PS27), clinical (toggle4 and 2008yc657), and environmental (NML110041) isolates.

The six genomes contained between 6,036 and 6,890 orthologous gene clusters, with DG4 encoding 6,694 clusters (Fig. 11A).The Up-Set analysis identified 5,243 gene clusters conserved across all six genomes, accounting for approximately 76–87% of the total gene repertoire of each strain. These universally conserved clusters responsible for fundamental cellular processes, including DNA replication, transcription, translation, and central carbon metabolism. Beyond the conserved core, the remaining genes constituted a dynamic accessory genome. Specifically, 966 gene clusters were shared among five genomes (Fig. 11B), while 319, 274, and 548 clusters were shared among four, three, and two genomes, respectively. This gradual reduction in shared gene content is characteristic of an open pangenome, reflecting ongoing gene acquisition and loss through horizontal gene transfer, homologous recombination, prophage integration, and other mobile genetic elements. Importantly, DG4 contained only 14 strain-specific gene clusters, a relatively small number compared with its total genome content. The intersection plots further revealed DG4 shares substantial accessory gene content with NML110041 and TY921, while fewer unique intersections were observed with the soil isolate PS27. The cluster occupancy heatmap (Fig. 11C) further illustrates that most large protein families are present across all six genomes. Notably, the largest cluster comprises 3,702 proteins distributed among all genomes, confirming that highly conserved functional modules remain evolutionarily stable. Overall, these findings indicate that *B. cenocepacia* DG4 possesses a highly conserved core genome together with a modest accessory genome that likely underpins adaptation to alkaline detergent environments.

Discussion

This study presents the isolation, phenotypic characterization, and genomic analysis of *B. cenocepacia DG4*, recovered from a liquid detergent product in Bangladesh.

Liquid detergents constitute highly selective environments characterized by alkaline pH, surfactants, preservatives, limited nutrient availability, and antimicrobial additives that collectively inhibit the growth of most microorganisms [64,65]. Detergent-associated survivors are typically alkali-tolerant, surfactant-tolerant, or metabolically specialized to exploit the limited resources available in such environments [64]. The recovery of DG4 from a sealed commercial product therefore demonstrates that certain BCC members possess exceptional physiological adaptations permitting long-term persistence under multiple concurrent chemical stresses. Consistent with this niche, DG4 tolerated moderately alkaline conditions P^H 7 to 10 and displayed positive catalase and lipase activity, traits that likely contributed synergistically to survival-catalase by neutralizing reactive oxygen species generated under aerobic, chemically stressful conditions [66], and lipase by enabling the hydrolysis and utilization of lipid-derived substrates present within the detergent formulation as an alternative carbon source [66].

A major strength of this study is the concordance between phenotype and genotype: strong biofilm formation, positive surface-active assays, and alkaline tolerance were corroborated by a genome enriched in genes for exopolysaccharide biosynthesis, membrane transport, stress response, lipid metabolism, and carbohydrate-active enzymes, strengthening confidence that these traits reflect genuine ecological strategies rather than isolated laboratory artifacts [69,70].

The predominance of glycosyl-transferases within the CAZyme repertoire further suggests enhanced capacity for extracellular polysaccharide biosynthesis, an essential component of mature biofilms that protects bacterial cells against surfactants, osmotic stress, and disinfectants. Biofilm formation is particularly important in industrial settings because microbial communities embedded within extracellular polymeric matrices exhibit markedly greater tolerance to preservatives and sanitation procedures than planktonic cells, thereby increasing the likelihood of persistent product contamination [71].

The preliminary surface-active properties observed in DG4 provide additional insight into its adaptation to detergent-rich environments. Positive drop-collapse and oil displacement assays together with an emulsification index exceeding 50% indicate the production of extracellular compounds capable of reducing interfacial tension [72]. Although chemical purification and structural characterization were not performed, genomic analysis identified several biosurfactant-associated genes, including *rhlA*, *plcN*, and *wbbL*, supporting the hypothesis that DG4 synthesizes molecules involved in emulsification and interactions with hydrophobic substrates[73]. Interestingly, canonical *rhlB* and *rhlC* genes were absent, suggesting that DG4 may produce alternative glycolipids or hydroxyalkanoyloxy alkanoates rather than classical rhamnolipids [74].

Beyond detergent survival, DG4's genome encodes complete pathways for degrading aliphatic hydrocarbons, monoaromatic compounds, polycyclic aromatic hydrocarbons, sulfur-containing petroleum compounds, and halogenated xenobiotics [75].

The coexistence of hydrocarbon-degrading enzymes (*alkB*, *ladA*, *dmpLMNOP*, *catABC*, *pcaGH*, *ndoA*, *styA*, *dhaA*, *dszBC*) with biosurfactant-associated genes suggest DG4 may simultaneously enhance pollutant bioavailability and catalyze its degradation, a combination favorable for bioremediation applications. The genome's 21 predicted biosynthetic gene clusters (NRPS, terpene, homoserine lactone, RiPP-like, phosphonate, β-lactone, and polyketide pathways) further indicate capacity for specialized metabolite production [76]; only one cluster showed similarity to a characterized siderophore, consistent with growing evidence that environmental *Burkholderia* genomes represent reservoirs of undiscovered natural products[77]. Multiple homoserine lactone clusters additionally suggest quorum-sensing regulation that may coordinate biofilm formation and stress adaptation in nutrient-limited environments [78,79].

The genome's resistom analysis identified multiple, mechanistically distinct resistance determinants. The detection of *adeF* and *amrA* suggests efflux-mediated reduced susceptibility across several antimicrobial classes, consistent with the established role of RND-family efflux systems in the intrinsic

multidrug resistance of *B. cenocepacia* [80].OXA-1142 and PEN-B1 provide plausible mechanisms for β-lactam inactivation, aligning with reports of extensive β-lactamase-mediated resistance in BCC [81]. Notably, *qacG*, associated with quaternary ammonium compound tolerance, may confer an adaptive advantage under disinfectant/preservative exposure characteristic of this niche. The additional *vanH*-associated glycopeptide resistance cluster broadens the resistance profile, though its phenotypic relevance in a Gram-negative background warrants cautious interpretation.

To contextualize DG4's susceptibility profile, we compiled literature-derived resistance data from four additional *B. cenocepacia* genomes (**Supplementary Fig. 4**) (Bioproject- PRJEB18658). Resistance to amikacin, ampicillin, gentamicin, cephalothin, netilmicin, and amoxicillin/clavulanic acid was universal across all four reference genomes, consistent with DG4's own resistant phenotype for these agents, supporting them as conserved, species-level intrinsic resistance traits rather than isolate-specific findings. Conversely, ceftazidime showed universal susceptibility, suggesting broader clinical utility against *B. cenocepacia*, though this agent was not tested in DG4.

Agents such as piperacillin/tazobactam, trimethoprim/sulfamethoxazole, meropenem, and ceftriaxone showed mixed resistant/susceptible distributions, indicating strain-dependent susceptibility. DG4 aligned with the majority-susceptible phenotype for piperacillin/tazobactam and cotrimoxazole, consistent with its larger inhibition zones (33.0 mm and 22.5 mm). For imipenem, the reference panel was evenly split (2/2), and DG4's complete resistance places it within this variable subset, reflecting strain-dependent carbapenem susceptibility in contrast to the more conserved resistance seen for aminoglycosides and earlier-generation β-lactams.

The coexistence of these determinants indicates a substantial genomic repertoire for antimicrobial and chemical-stress tolerance; however, gene presence alone does not confirm phenotypic resistance and should be interpreted alongside susceptibility testing results. The PathogenFinder-predicted pathogenicity probability of 0.9819 suggests DG4 carries genomic features characteristic of human-associated pathogens; however, such in silico predictions reflect probabilistic classification based on comparative genomic signatures rather than direct pathogenicity evidence, and should be regarded as hypothesis-generating rather than definitive [82].

Genomic plasticity further characterizes DG4: 232 mobileOG-associated genes and three intact prophages (related to *Burkholderia*, *Vibrio*, and *Escherichia* phages) indicate an active mobilome, consistent with ongoing horizontal gene transfer contributing to trait acquisition [83,84]. Such findings reinforce concerns that environmental reservoirs, including contaminated consumer products, may facilitate the emergence and persistence of opportunistic pathogens capable of future resistance/virulence gene dissemination [85].The relatedness of one prophage to *Burkholderia* phage ST79 suggests long-term phage–host co-evolution, whereas similarity to *Vibrio*/*Escherichia* phages likely reflects the mosaic architecture typical of temperate phage genomes.

Interestingly, auxiliary activity (AA) enzymes and polysaccharide lyases (PLs) were absent from the predicted CAZyme repertoire. AA enzymes are commonly linked to oxidative breakdown of lignin and other recalcitrant plant polymers, whereas PLs catalyze the cleavage of acidic polysaccharides such as pectin and alginate. This pattern is consistent with the ecological lifestyle of *Burkholderia* species, which are widely distributed in soil and water environments and are generally known for metabolic versatility rather than strong lignocellulose-degrading capacity[86].

Comparative genomic analyses demonstrated that DG4 shares a highly conserved core genome with clinical and environmental *B. cenocepacia* strains while maintaining only a small accessory genome associated with ecological specialization. The high ANI value(98.28%) is consistent with species-level identity, since ANI values above about 95–96% are widely used to define bacterial species boundaries[87]. Phylogenomic clustering and genome comparisons of clinical and environmental *B. cenocepacia* strains also show that strain-level differences are often driven by variable accessory regions, genomic islands, and niche-associated loci[88].

Overall, DG4 illustrates the dual ecological nature of environmental *B. cenocepacia*. On one hand, its extensive catabolic pathways, biosynthetic diversity, and presumptive surface-active traits indicate considerable promise for hydrocarbon biodegradation and environmental biotechnology. On the other hand, its drug resistance, disinfectant tolerance, biofilm-forming ability, and predicted pathogenicity necessitate caution when considering any environmental or industrial application. Future investigations should prioritize functional validation of biosurfactant biosynthesis, transcriptomic analysis under detergent stress, and experimental assessment of virulence and biodegradation efficiency before any practical exploitation is considered.

Conclusions

B. cenocepacia DG4 was recovered from a commercial liquid soap and showed phenotypic traits consistent with persistence under alkaline and surface-active chemical stress. Genome analysis revealed multiple resistance, virulence-associated, stress-response, metabolic and mobile-element determinants. Together, these findings identify commercial personal-care products as a potential environmental reservoir requiring continued microbiological surveillance, while functional studies are needed to validate the predicted resistance, biodegradation and virulence-associated functions.

Data Availability

The present whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under accession number JBXHAV000000000. The associated BioSample and BioProject identifiers are SAMN57363271 and PRJNA1455362. Raw sequencing reads are available in the Sequence Read Archive under accession number SRR38199625.

Abbreviations

BCC
Burkholderia cepacia complex
WGS
Whole-genome sequencing
CDS
Coding sequences
QAC
Quaternary ammonium compound
CLSI
Clinical and Laboratory Standards Institute
EC
Enzyme commission
GO
Gene ontology
CARD
Comprehensive Antibiotic Resistance Database
VFDB
Virulence Factor Database
BV-BRC
Bacterial and Viral Bioinformatics Resource Center
iGC
icddr,b Genome Center
GC
Guanine-cytosine
bp
Base pairs
BTRI
Biomedical and Toxicological Research Institute
BCSIR
Bangladesh Council of Scientific and Industrial Research.

Declarations

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The work was conducted as part of the institutional research activities of the Bangladesh Council of Scientific and Industrial Research (BCSIR).

Author Contribution

AT conceived and designed the study, performed the bacterial isolation and identification, and drafted the manuscript. NU performed bioinformatics analysis and genome annotation. SI contributed to methodology validation. AJ contributed to phenotypic characterization and antibiotic susceptibility testing. MII performed genomic DNA extraction and quality assessment. All authors reviewed, edited, and approved the final manuscript.

Acknowledgements

The authors gratefully acknowledge the icddr,b Genome Center (iGC) for whole-genome sequencing services. The authors also thank the Bangladesh Council of Scientific and Industrial Research (BCSIR) for institutional support and resources.

Data Availability

The present whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under accession number JBXHAV000000000. The associated BioSample and BioProject identifiers are SAMN57363271 and PRJNA1455362. Raw sequencing reads are available in the Sequence Read Archive under accession number SRR38199625.

References

1. Vandamme, P. & Dawyndt, P. Classification and identification of the *Burkholderia cepacia* complex: Past, present and future. *Syst. Appl. Microbiol.* **34**, 87–95 (2011).
2. Vandamme, P. et al. Occurrence of Multiple Genomovars of Burkholderia cepacia in Cystic Fibrosis Patients and Proposal of Burkholderia multivorans sp. nov. *Int. J. Syst. Evol. Microbiol.* **47**, 1188–1200 (1997).
3. Tavares, M., Kozak, M. & Balola, A. Sá-Correia, I. Burkholderia cepacia Complex Bacteria: a Feared Contamination Risk in Water-Based Pharmaceutical Products. *Clin. Microbiol. Rev.* **33** 10.1128/cmr.00139 - 19 (2020).
4. Wang, G., Zarodkiewicz, P. & Valvano, M. A. Current Advances in Burkholderia Vaccines Development. *Cells* **9**, 2671 (2020).
5. Mahenthiralingam, E., Baldwin, A. & Dowson, C. G. Burkholderia cepacia complex bacteria: opportunistic pathogens with important natural biology. *J. Appl. Microbiol.* **104**, 1539–1551 (2008).
6. O'Sullivan, L. A. & Mahenthiralingam, E. Biotechnological potential within the genus Burkholderia. *Lett. Appl. Microbiol.* **41**, 8–11 (2005).
7. Parke, J. L., Gurian-Sherman, D., DIVERSITY OF THEBURKHOLDERIA CEPACIA COMPLEX & AND IMPLICATIONS FOR RISK ASSESSMENT OF BIOLOGICAL CONTROL STRAINS. *Annu. Rev. Phytopathol.* **39**, 225–258 (2001).

8. Ahn, Y. et al. Evaluation of liquid and solid culture media for the recovery and enrichment of Burkholderia cenocepacia from distilled water. *J. Ind. Microbiol. Biotechnol.* **41**, 1109–1118 (2014).
9. Fang, Y., Xie, G., Lou, M., Li, B. & Muhammad, I. Diversity analysis of Burkholderia cepacia complex in the water bodies of West Lake, Hangzhou, China. *J. Microbiol.* **49**, 309–314 (2011).
10. Vermis, K., Brachkova, M., Vandamme, P. & Nelis, H. Isolation of *Burkholderia cepacia* Complex Genomovars from Waters. *Syst. Appl. Microbiol.* **26**, 595–600 (2003).
11. Zanetti, F., De Luca, G. & Stampi, S. Recovery of *Burkholderia pseudomallei* and *B. cepacia* from drinking water. *Int. J. Food Microbiol.* **59**, 67–72 (2000).
12. Eissa, M. E. Distribution of bacterial contamination in non-sterile pharmaceutical materials and assessment of its risk to the health of the final consumers quantitatively. *Beni-Suef Univ. J. Basic. Appl. Sci.* **5**, 217–230 (2016).
13. Jimenez, L. Microbial Diversity in Pharmaceutical Product Recalls and Environments. *PDA J. Pharm. Sci. Technol.* **61**, 383–399 (2007).
14. Gnanadhas, D. P., Marathe, S. A. & Chakravortty, D. Biocides – resistance, cross-resistance mechanisms and assessment. *Expert Opin. Investig Drugs*. **22**, 191–206 (2013).
15. Gnanadhas, D. P., Marathe, S. A. & Chakravortty, D. Biocides – resistance, cross-resistance mechanisms and assessment. *Expert Opin. Investig Drugs*. **22**, 191–206 (2013).
16. McDonnell, G. & Russell, A. D. Antiseptics and Disinfectants: Activity, Action, and Resistance. *Clin. Microbiol. Rev.* **12**, 147–179 (1999).
17. Weber, D. J., Rutala, W. A. & Sickbert-Bennett, E. E. Outbreaks Associated with Contaminated Antiseptics and Disinfectants. *Antimicrob. Agents Chemother.* **51**, 4217–4224 (2007).
18. Ahn, Y. et al. Intrinsic Resistance of Burkholderia cepacia Complex to Benzalkonium Chloride. *mBio* 7, (2016). 10.1128/mbio.01716-16
19. Kampf, G. Adaptive microbial response to low-level benzalkonium chloride exposure. *J. Hosp. Infect.* **100**, e1–e22 (2018).
20. Greenberg, D. E. et al. Recurrent Burkholderia Infection in Patients with Chronic Granulomatous Disease: 11-Year Experience at a Large Referral Center. *Clin. Infect. Dis.* **48**, 1577–1579 (2009).
21. Ganesh, P. S. et al. Intracellular survival and innate immune evasion of Burkholderia cepacia: Improved understanding of quorum sensing-controlled virulence factors, biofilm, and inhibitors. *Microbiol. Immunol.* **64**, 87–98 (2020).
22. Rhodes, K. A. & Schweizer, H. P. Antibiotic resistance in *Burkholderia* species. *Drug Resist. Updat*. **28**, 82–90 (2016).
23. Scoffone, V. C. et al. Burkholderia cenocepacia Infections in Cystic Fibrosis Patients: Drug Resistance and Therapeutic Approaches. *Front Microbiol* **8**, (2017).
24. Torbeck, L., Raccasi, D., Guilfoyle, D. E., Friedman, R. L. & Hussong, D. Burkholderia cepacia: This Decision Is Overdue. *PDA J. Pharm. Sci. Technol.* **65**, 535–543 (2011).

25. Agnoli, K. et al. Use of Synthetic Hybrid Strains To Determine the Role of Replicon 3 in Virulence of the Burkholderia cepacia Complex. *Appl. Environ. Microbiol.* **83**, e00461–e00417 (2017).
26. Gutiérrez Santana, J. C. & Coria Jiménez, V. R. Burkholderia cepacia complex in cystic fibrosis: critical gaps in diagnosis and therapy. *Ann Med* **56**, 2307503 .
27. Xavier, T. S. et al. Genomic Insights into a New Burkholderia cenocepacia Sequence Type Linked to Cepacia Syndrome in Cystic Fibrosis. *Curr. Microbiol.* **83**, 420 (2026).
28. Drevinek, P. & Mahenthiralingam, E. Burkholderia cenocepacia in cystic fibrosis: epidemiology and molecular mechanisms of virulence. *Clin. Microbiol. Infect.* **16**, 821–830 (2010).
29. Uddin, N. et al. Genomic analysis of Acinetobacter baumannii DUEMBL6 reveals diesel bioremediation potential and biosafety concerns. *PLOS ONE*. **21**, e0339456 (2026).
30. Iwase, T. et al. A Simple Assay for Measuring Catalase Activity: A Visual Approach. *Sci. Rep.* **3**, 3081 (2013).
31. Lee, L. P. et al. Lipase-Secreting Bacillus Species in an Oil-Contaminated Habitat: Promising Strains to Alleviate Oil Pollution. *BioMed Res. Int.* 820575 (2015). (2015).
32. Morikawa, M. et al. A new lipopeptide biosurfactant produced by Arthrobacter sp. strain MIS38. *J. Bacteriol.* **175**, 6459–6466 (1993).
33. Youssef, N. H. et al. Comparison of methods to detect biosurfactant production by diverse microorganisms. *J. Microbiol. Methods*. **56**, 339–347 (2004).
34. Abouseoud, M., Maachi, R., Amrane, A., Boudergua, S. & Nabi, A. Evaluation of different carbon and nitrogen sources in production of biosurfactant by Pseudomonas fluorescens. *Desalination* **223**, 143–151 (2008).
35. Stepanović, S. et al. Quantification of biofilm in microtiter plates: overview of testing conditions and practical recommendations for assessment of biofilm production by staphylococci. *APMIS* **115**, 891–899 (2007).
36. Ward, C. M., To, T. H. & Pederson, S. M. ngsReports: a Bioconductor package for managing FastQC reports and other NGS related log files. *Bioinforma Oxf. Engl.* **36**, 2587–2588 (2020).
37. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
38. Bankevich, A. et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol. J. Comput. Mol. Cell. Biol.* **19**, 455–477 (2012).
39. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinforma Oxf. Engl.* **30**, 2068–2069 (2014).
40. Aziz, R. K. et al. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genom.* **9**, 75 (2008).
41. Li, W. et al. RefSeq: expanding the Prokaryotic Genome Annotation Pipeline reach with protein family model curation. *Nucleic Acids Res.* **49**, D1020–D1028 (2021).

42. Tatusova, T. et al. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* **44**, 6614–6624 (2016).
43. Grant, J. R. et al. Proksee: in-depth characterization and visualization of bacterial genomes. *Nucleic Acids Res.* **51**, W484–W492 (2023).
44. Tian, L., Huang, C., Mazloom, R., Heath, L. S. & Vinatzer, B. A. LINbase: a web server for genome-based identification of prokaryotes as members of crowdsourced taxa. *Nucleic Acids Res.* **48**, W529–W537 (2020).
45. Kirstahler, P. et al. Genomics-Based Identification of Microorganisms in Human Ocular Body Fluid. *Sci. Rep.* **8**, 4126 (2018).
46. Meier-Kolthoff, J. P. & Göker, M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun.* **10**, 2182 (2019).
47. Auch, A. F., von Jan, M., Klenk, H. P. & Göker, M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand. Genomic Sci.* **2**, 117–134 (2010).
48. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
49. Larsen, M. V. et al. Multilocus sequence typing of total-genome-sequenced bacteria. *J. Clin. Microbiol.* **50**, 1355–1361 (2012).
50. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
51. Sun, Y. et al. gcMeta 2025: a global repository of metagenome-assembled genomes enabling cross-ecosystem microbial discovery and function research. *Nucleic Acids Res.* **54**, D724–D733 (2026).
52. Blin, K. et al. antiSMASH 8.0: extended gene cluster detection capabilities and analyses of chemistry, enzymology, and regulation. *Nucleic Acids Res.* **53**, W32–W38 (2025).
53. Zheng, J. et al. dbCAN3: automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Res.* **51**, W115–W121 (2023).
54. Chen, L. et al. VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res.* **33**, D325–328 (2005).
55. Cosentino, S., Voldby Larsen, M., Møller Aarestrup, F. & Lund, O. PathogenFinder - Distinguishing Friend from Foe Using Bacterial Whole Genome Sequence Data. *PLoS ONE*. **8**, e77302 (2013).
56. Dhillon, B. K. et al. IslandViewer 3: more flexible, interactive genomic island discovery, visualization and analysis: Fig. 1. *Nucleic Acids Res.* **43**, W104–W108 (2015).
57. Carattoli, A. et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* **58**, 3895–3903 (2014).
58. Robertson, J. & Nash, J. H. E. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microb. Genomics*. **4**, e000206 (2018).

59. Roux, S., Enault, F., Hurwitz, B. L. & Sullivan, M. B. VirSorter: mining viral signal from microbial genomic data. *PeerJ* **3**, e985 (2015).
60. Wishart, D. S. et al. PHASTEST: faster than PHASTER, better than PHAST. *Nucleic Acids Res.* **51**, W443–W450 (2023).
61. Brown, C. L. et al. mobileOG-db: a Manually Curated Database of Protein Families Mediating the Life Cycle of Bacterial Mobile Genetic Elements. *Appl. Environ. Microbiol.* **88**, e0099122 (2022).
62. Biswas, A., Staals, R. H. J., Morales, S. E., Fineran, P. C. & Brown, C. M. CRISPRDetect: A flexible algorithm to define CRISPR arrays. *BMC Genom.* **17**, 356 (2016).
63. Sun, J. et al. OrthoVenn3: an integrated platform for exploring and visualizing orthologous data across genomes. *Nucleic Acids Res.* **51**, W397–W403 (2023).
64. Cheng, K., Khoo, Z. S., Lo, N. W., Tan, W. J. & Chemmangattuvalappil, N. Design and performance optimisation of detergent product containing binary mixture of anionic-nonionic surfactants. *Heliyon* **6**, (2020).
65. Zarinviarsagh, M., Ebrahimipour, G. & Sadeghi, H. Lipase and biosurfactant from Ochrobactrum intermedium strain MZV101 isolated by washing powder for detergent application. *Lipids Health Dis* **16**, (2017).
66. Corral, J. et al. The Ralstonia solanacearum catalase-peroxidase KatG is crucial to survive environmental stresses and also plays a role in plant infection. *Sci. Rep.* **15**, 35168 (2025).
67. Gupta, S. & Pandey, S. A. C. C. Deaminase Producing Bacteria With Multifarious Plant Growth Promoting Traits Alleviates Salinity Stress in French Bean (Phaseolus vulgaris) Plants. *Front Microbiol* **10**, (2019).
68. Kuan, J., Tsai, C. H., Chou, C., Wu, C. & Wu, W. F. Enzymatic Characterization of a Novel HSL Family IV Esterase EstD04 from Pseudomonas sp. D01 in Mealworm Gut Microbiota. *Molecules* 28, (2023).
69. Lu, K., Wang, X., Zhou, Y. & Zhu, Q. Genomic characterization and probiotic potential assessment of an exopolysaccharide-producing strain Pediococcus pentosaceus LL-07 isolated from fermented meat. *BMC Microbiol* **24**, (2024).
70. Meza, C. et al. Genomic basis and functional characterization of the exopolysaccharide production by a thermotolerant Bacillus isolated from Tolhuaca hot spring. *Front Microbiol* **16**, (2025).
71. Flemming, H. C. & Wingender, J. The biofilm matrix. *Nat. Rev. Microbiol.* **8**, 623–633 (2010).
72. Nayarisseri, A., Singh, P. & Singh, S. K. Screening, isolation and characterization of biosurfactant producing Bacillus subtilis strain ANSKLAB03. *Bioinformation* 14, 304–314 (2018).
73. Guzmán, E., Ortega, F. & Rubio, R. G. Exploring the world of rhamnolipids: A critical review of their production, interfacial properties, and potential application. *Curr. Opin. Colloid Interface Sci.* **69**, 101780 (2024).
74. Thakur, P. et al. Rhamnolipid the Glycolipid Biosurfactant: Emerging trends and promising strategies in the field of biotechnology and biomedicine. *Microb. Cell. Factories*. **20**, 1 (2021).

75. Dardouri, M., Mendes, R. M., Frenzel, J., Costa, J. & Ribeiro, I. A. C. Seeking faster, alternative methods for glycolipid biosurfactant characterization and purification. *Anal. Bioanal Chem.* **413**, 4311–4320 (2021).
76. Kunakom, S. & Eustáquio, A. S. Burkholderia as a Source of Natural Products. *J. Nat. Prod.* **82**, 2018–2037 (2019).
77. Alam, K. et al. In silico genome mining of potential novel biosynthetic gene clusters for drug discovery from Burkholderia bacteria. *Comput. Biol. Med.* **140**, 105046 (2022).
78. Ma, H. et al. The diversity, distribution and function of N-acyl-homoserine lactone (AHL) in industrial anaerobic granular sludge. *Bioresour Technol.* **247**, 116–124 (2018).
79. Wang, N. et al. Realizing the role of *N*-acyl-homoserine lactone-mediated quorum sensing in nitrification and denitrification: A review. *Chemosphere* **274**, 129970 (2021).
80. Buroni, S. et al. Assessment of three Resistance-Nodulation-Cell Division drug efflux transporters of Burkholderia cenocepacia in intrinsic antibiotic resistance. *BMC Microbiol.* **9**, 200 (2009).
81. Abbott, F. K., Milne, K. E. N., Stead, D. A. & Gould, I. M. Combination antimicrobial susceptibility testing of Burkholderia cepacia complex: significance of species. *Int. J. Antimicrob. Agents*. **48**, 521–527 (2016).
82. Ferrer Florensa, A. et al. Whole-genome prediction of bacterial pathogenic capacity on novel bacteria using protein language models with PathogenFinder2. *Bioinformatics* 42, btag129 (2026).
83. Dobrindt, U., Hochhut, B., Hentschel, U. & Hacker, J. Genomic islands in pathogenic and environmental microorganisms. *Nat. Rev. Microbiol.* **2**, 414–424 (2004).
84. Tokuda, M. & Shintani, M. Microbial evolution through horizontal gene transfer by mobile genetic elements. *Microb. Biotechnol.* **17**, e14408 (2024).
85. Han, B. et al. Mobile genetic elements encoding antibiotic resistance genes and virulence genes in Klebsiella pneumoniae: important pathways for the acquisition of virulence and resistance. *Front. Microbiol.* **16**, 1529157 (2025).
86. Sumranwanich, T. et al. Evaluating lignin degradation under limited oxygen conditions by bacterial isolates from forest soil. *Sci. Rep.* **14**, 13350 (2024).
87. Kim, M., Oh, H. S., Park, S. C. & Chun, J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int. J. Syst. Evol. Microbiol.* **64**, 346–351 (2014).
88. Wallner, A., King, E., Ngonkeu, E. L. M., Moulin, L. & Béna, G. Genomic analyses of Burkholderia cenocepacia reveal multiple species with differential host-adaptation to plants and humans. *BMC Genom.* **20**, 803 (2019).

Figures

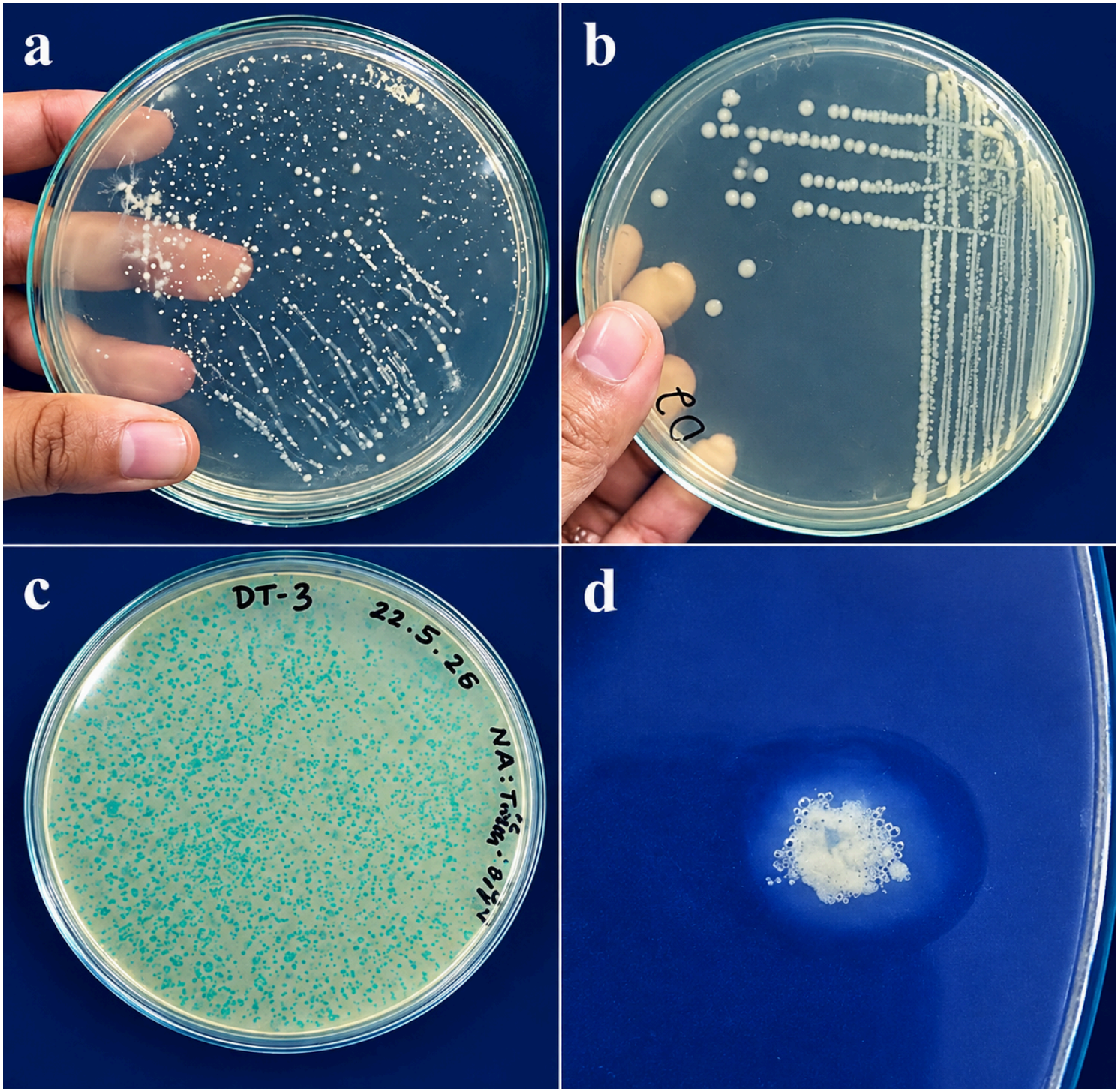


Figure 1

Cultural morphology and enzymatic characterization of the bacterial isolate. (A) colony morphology on nutrient agar (10^{-3}), (B) pure colonies growth on Nutrient agar, (C) positive lipase activity demonstrating extracellular lipid hydrolysis, and (D) positive catalase reaction indicated by oxygen bubble formation following hydrogen peroxide exposure.

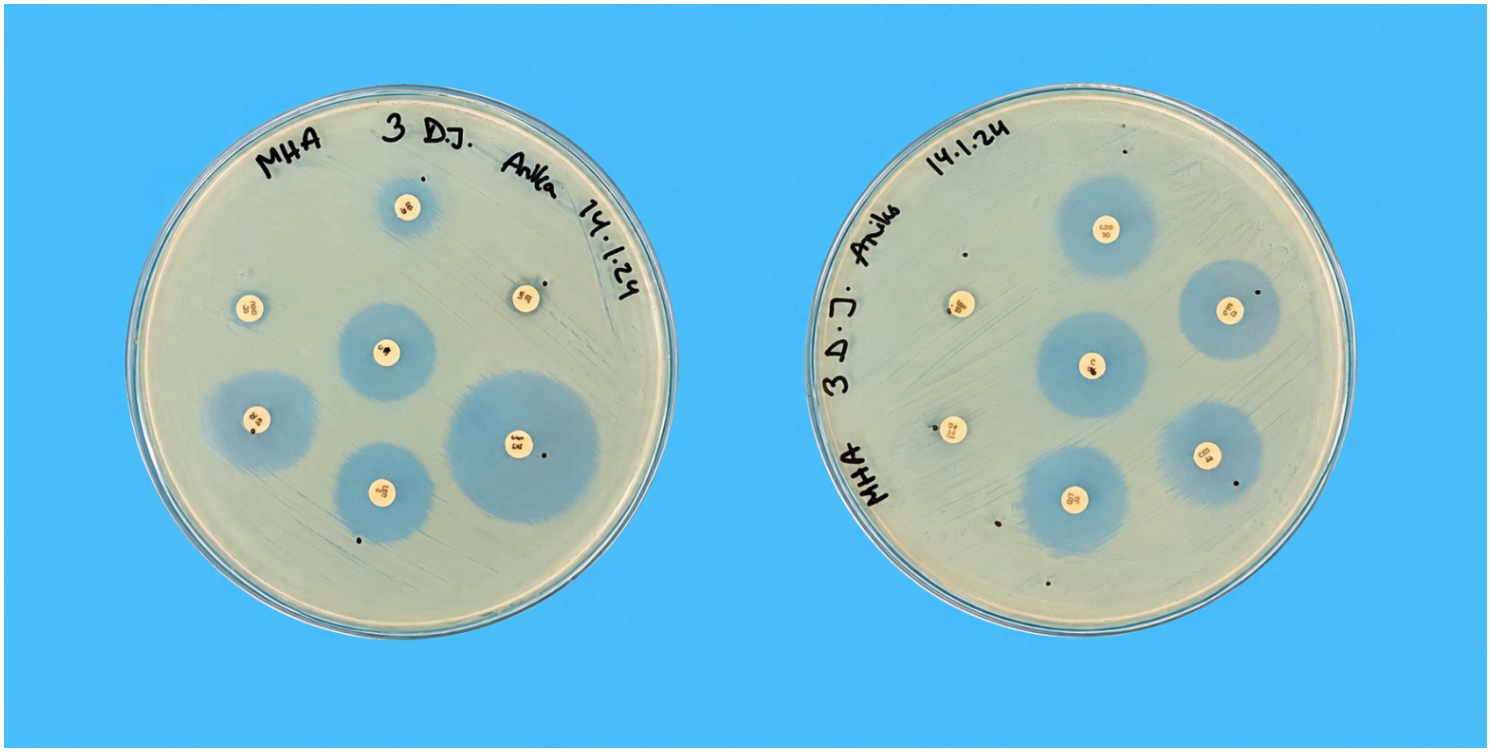


Figure 2

Antibiotic susceptibility profile of *B. cenocepacia* DG4 determined by the Kirby-Bauer disk diffusion assay.

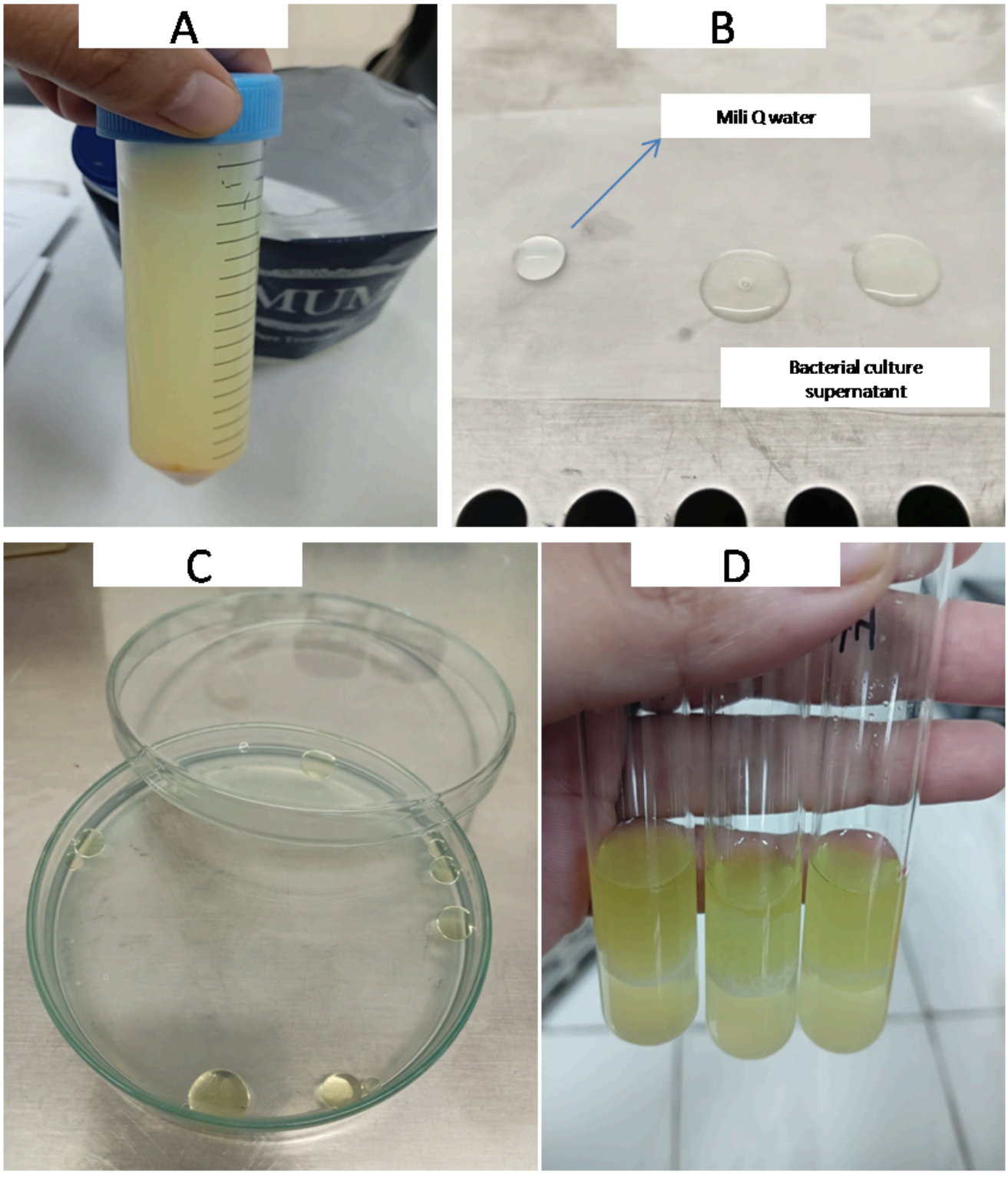


Figure 3

Evaluation of surface-active properties of the bacterial isolate. (A) bacterial culture grown in liquid medium with oil supplementation. (B) Positive drop-collapse assay of the cell-free supernatant compared with Milli-Q water control. (C) Oil displacement assay demonstrating biosurfactant-mediated spreading activity on an oil surface. (D) Emulsification index (E24) assay showing stable emulsion formation after 24 hours incubation.

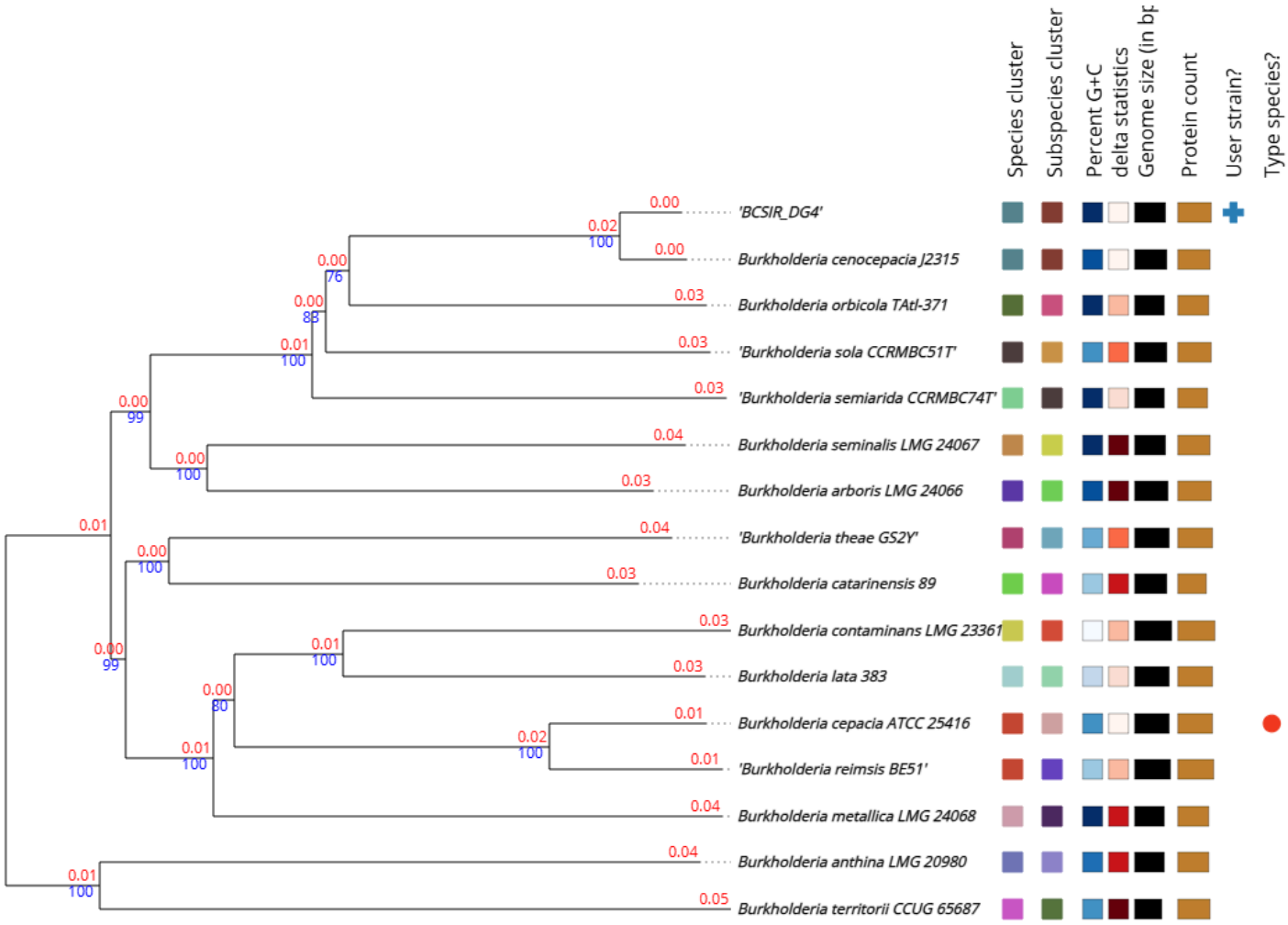


Figure 4

TYGS based phylogenetic tree indicating the evolutionary position of BCSIR_DG4, highlighted in “+”, within the genus *Burkholderia*. It is clustered together with the reference strain, namely, *B. cenocepacia* J2315.

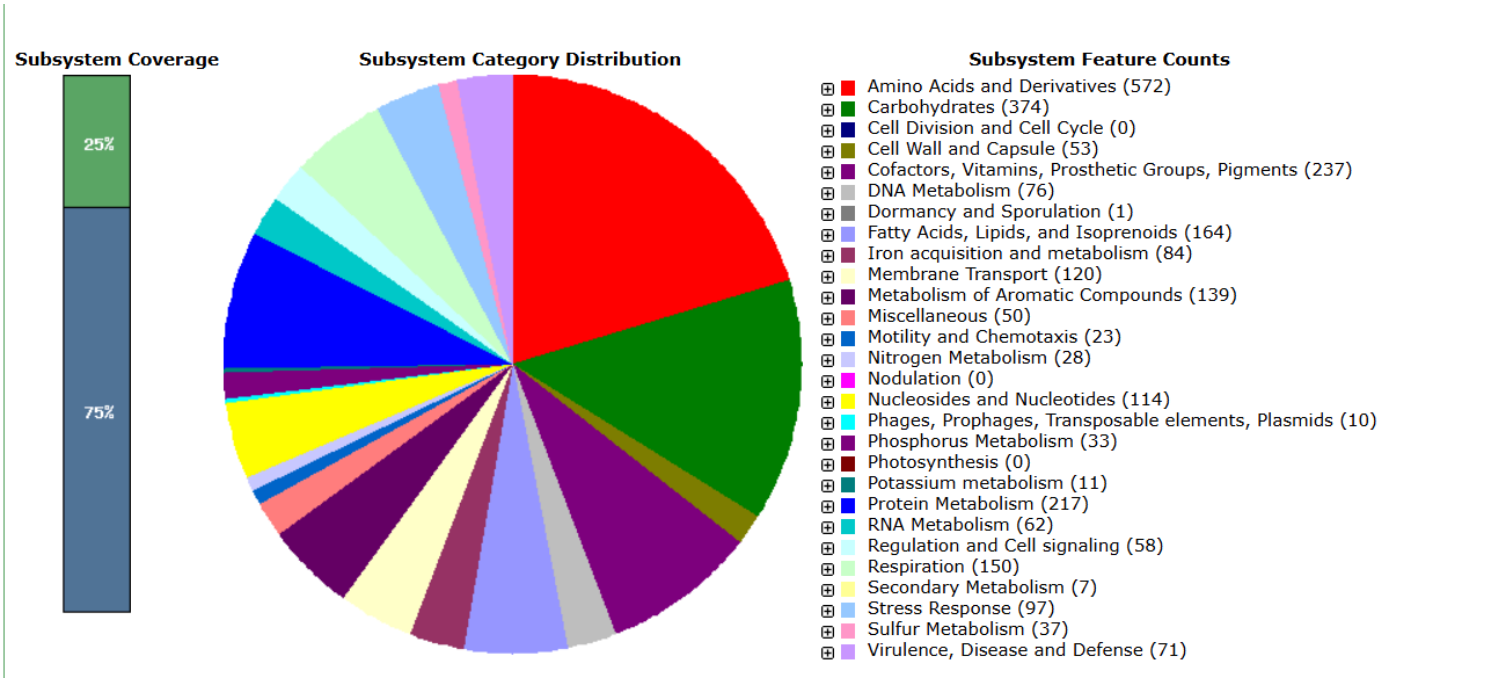


Figure 5

Genome annotation of *B. cenocepacia* DG4 using RAST server.

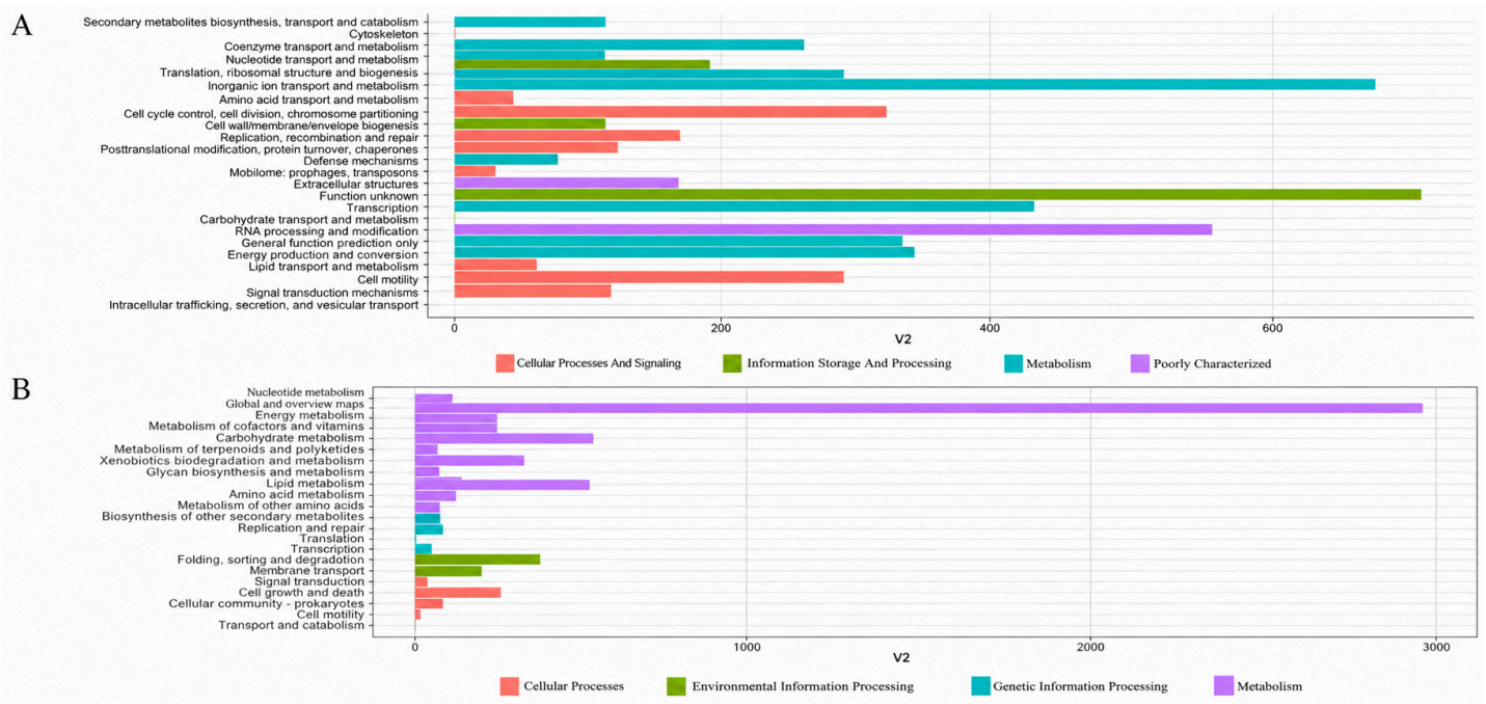


Figure 6

Functional annotation and metabolic profiling (A) Chart of the Clusters of Orthologous Groups (COG) functional classification, showing strong representation in Metabolism, Cellular Processes and Signaling, and Information Storage and Processing. (B) Histogram of KEGG pathway mapping, which shows the counts of genes in major pathway categories and enrichment in metabolic pathways.

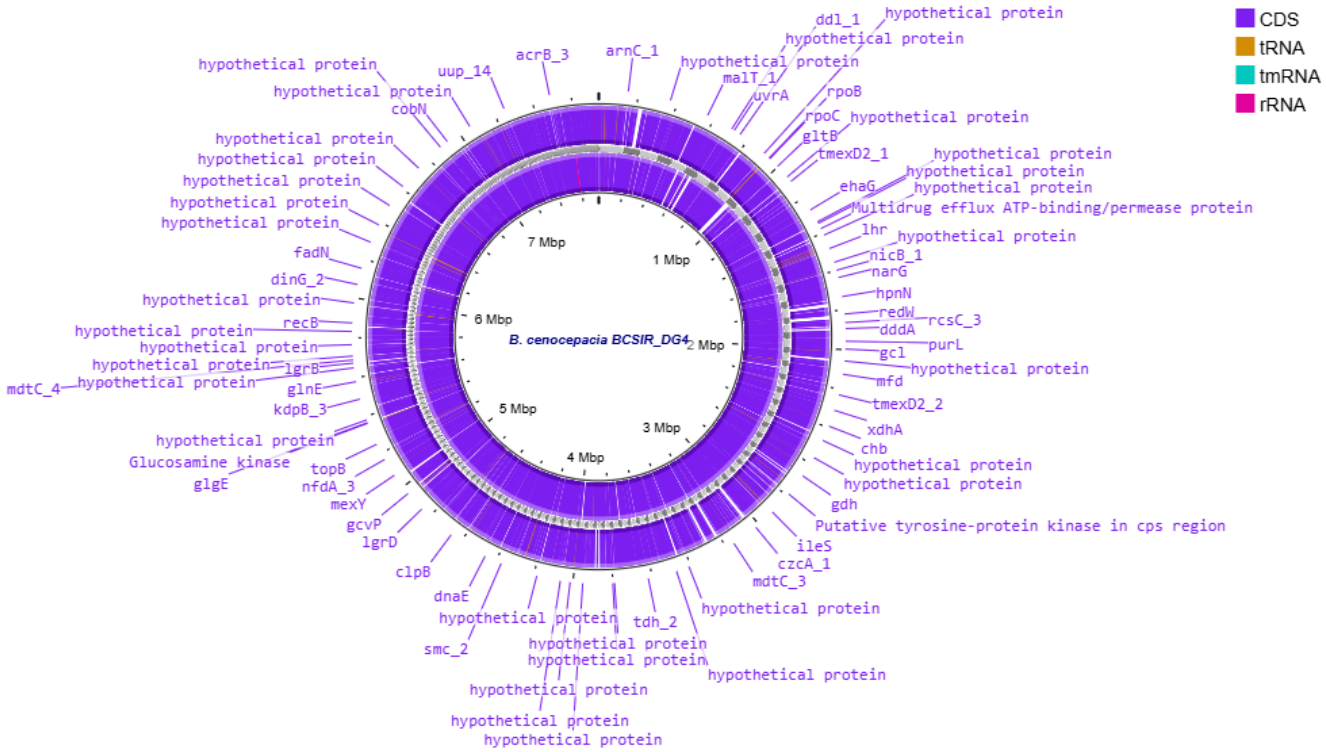


Figure 7

Genome Architecture of *Burkholderia cenocepacia* DG4.

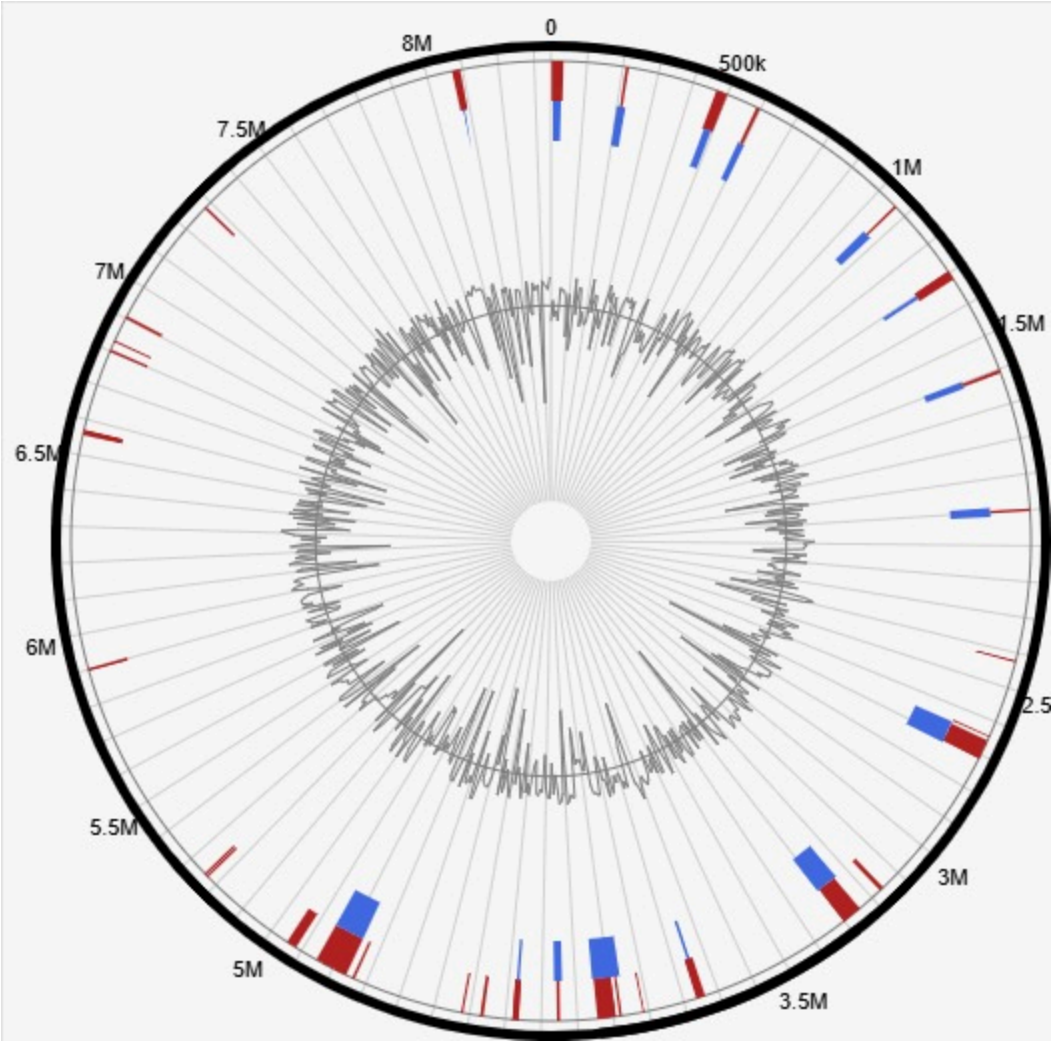


Figure 8

Genomic islands in Genomic islands in *B. cenocepacia* DG4. The red, blue and orange colors show different prediction methods such as Integrated, IslandPath-DIMOB, and SIGI-HMM respectively. The zigzag lines demonstrate the contig boundary in the genome.

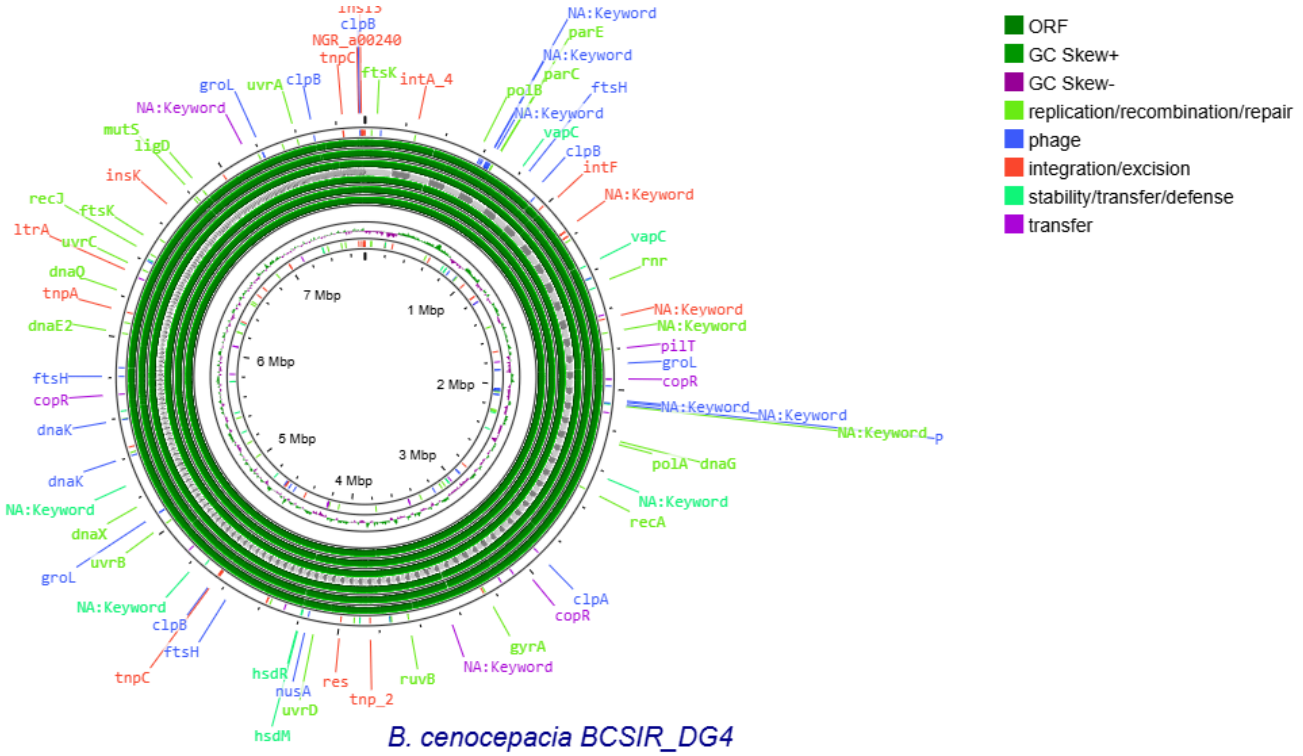


Figure 9

Circular genome map of *B. cenocepacia*BCSIR_DG4 showing ORFs, GC skew, and major functional genomic features, including replication/repair, phage, integration/excision, defense, stability, and transfer-related genes.

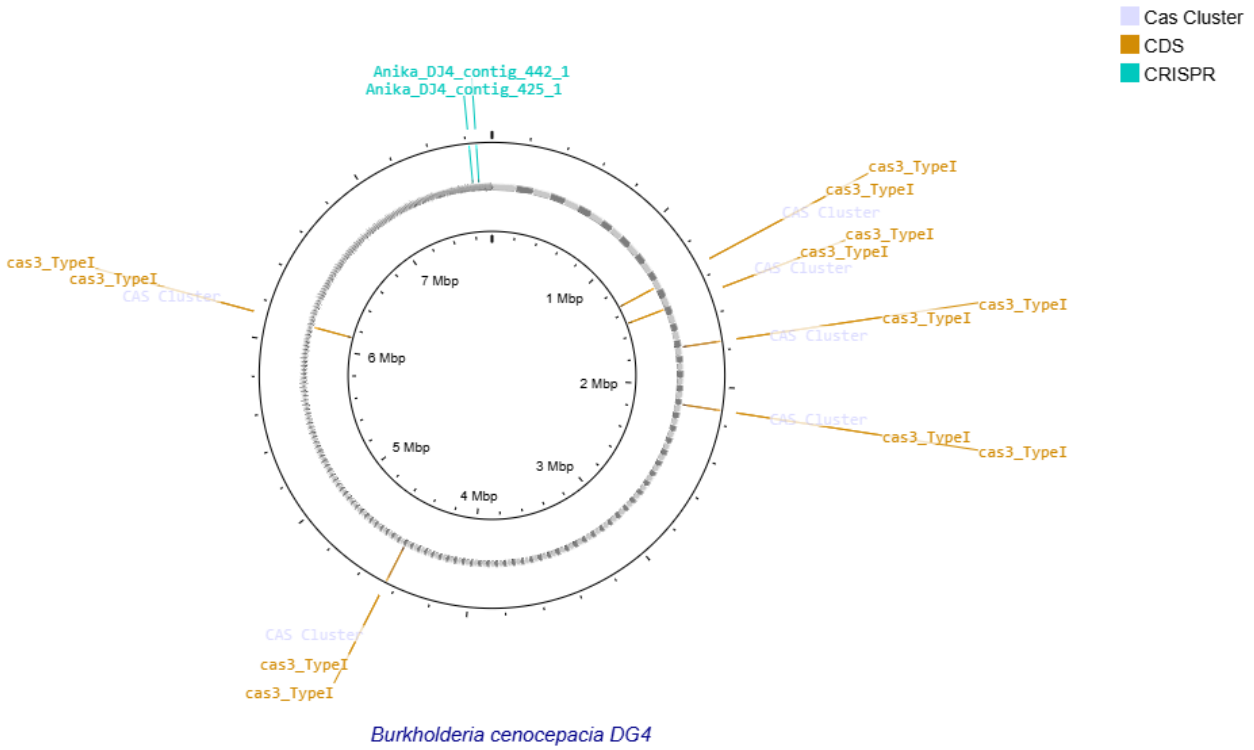


Figure 10

Genome-wide distribution of CRISPR–Cas-associated loci in *Burkholderia cenocepacia* DG4.

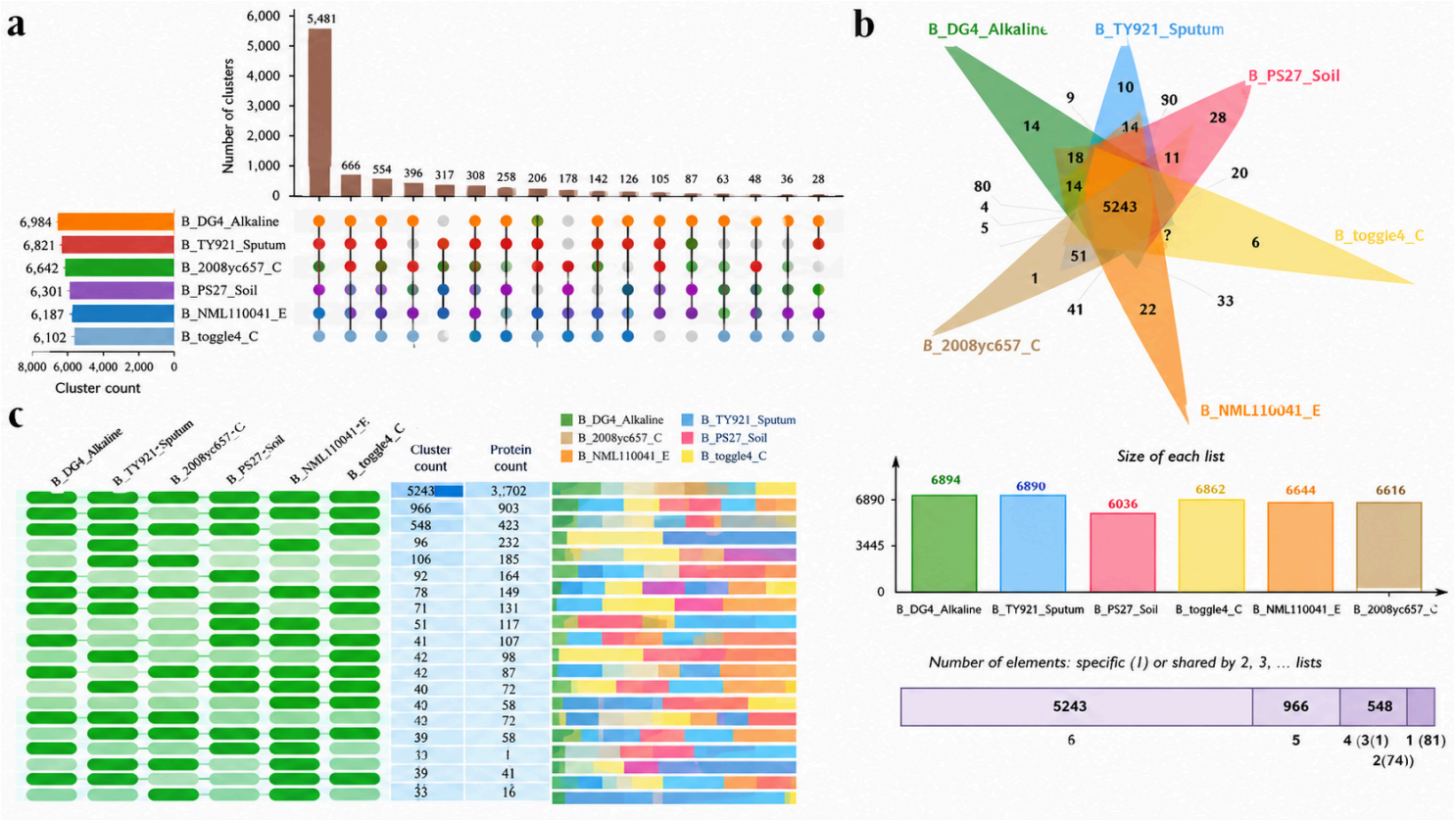


Figure 11

Comparative genome analysis. (a) Upset gene cluster distribution. (b) Intersection Plot- based gene distribution. (c) The cluster occupancy heatmap.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFileIslandViewer.xls
- COGandKEGGfigureusedpermission.pdf