Research Article

BIOMEDICAL Journal of Scientific & Technical Research

ISSN: 2574 -1241

DOI: 10.26717/BJSTR.2025.63.009834

Genome Sequence of a Beta-Hemolytic *Pseudomonas aeruginosa* GXDRC_02 Isolated from a Clinical Sample, Gujarat, India

Vivek N Upasani[1], Zuber Saiyed[1], Jitendra Nayak[1], Anjali D Prajapati[2], Hetvi P Patel[1], Alpesh R Patel[1]*, Shiva S Chettiar[1], Sushma R Shah[3] and Devendrasinh D Jhala[4]

[1]*Genexplore Diagnostics and Research Centre Pvt. Ltd., India*

[2]*Department of Microbiology, Gujarat Vidyapith, India*

[3]*Department of Obstetrics and Gynaecology, Smt. NHL Municipal Medical College and SVPIMSR, India*

[4]*Department of Zoology, School of Sciences, Gujarat University, India*

***Corresponding author:** Alpesh R Patel, Genexplore Diagnostics and Research Centre Pvt. Ltd., 1201-1210, Iconic Shyamal, Shyamal Cross Road, Ahmedabad 380015, Gujarat, India

ARTICLE INFO

Received: August 12, 2025

Published: August 25, 2025

Citation: Vivek N Upasani, Zuber Saiyed, Jitendra Nayak, Anjali D Prajapati, Hetvi P Patel, Alpesh R Patel, Shiva S Chettiar, Sushma R Shah and Devendrasinh D Jhala. Genome Sequence of a Beta-Hemolytic *Pseudomonas aeruginosa* GXDRC_02 Isolated from a Clinical Sample, Gujarat, India. Biomed J Sci & Tech Res 63(1)-2025. BJSTR. MS.ID.009834.

ABSTRACT

Objective: To carryout Whole Genome Sequence analysis of a beta-hemolytic *P. aeruginosa* GXDRC_02 isolated from an enriched LIM broth inoculated with a vaginal swab from a pregnant woman in Gujarat, India to serve as a reference for future research.

Methods: The strain was isolated on Blood agar plate and pure culture was obtained. Whole genome sequence analysis of the DNA extracted from this isolate was carried out using the MGI DNBSEQ-G400 platform to investigate the genomic characteristics. Various bioinformatics tools were used to analyse the NGS data such as BV-BRC genome analysis pipeline, PGAP, Phylogenetic analysis, RAST toolkit (RASTtk), eggNOG-mapper v2, KEGG Automatic Annotation Server, CheckM, CARD, ResFinder v4.1, VFDB, antiSMASH v7.0, PlasmidFinder, PHASTEST, etc.

Results: The genome consisted of 60 contigs, with a total length of 6,233,771bp. The assembly displayed a GC content of 66.38%, metrics included an N50 of 306,098bp and an L50 of 8, indicating a relatively contiguous assembly. No plasmids or chromosomal circularization were reported. Phylogenetic analyses, including codon tree-based, further established a close relationship with *P. aeruginosa* PAO1. The annotated genome comprised 5,913 protein-coding sequences (CDS), 55 transfer RNA (tRNA) genes, and 3 ribosomal RNA (rRNA) genes. Among the CDSs, 4,813 proteins were assigned functional annotations, while 1,100 proteins were identified as hypothetical. The genome contained one intact prophage (32.3kb). The probability score for it as a human pathogen was 0.752. BioProject No. PRJNA1231733; Sample ID: SAMN47218841

Conclusion: These findings provide important scientific insights into the genome of this opportunistic pathogen and its impact on nosocomial infections during pregnancy.

Keywords: *Pseudomonas aeruginosa*; beta-hemolytic; WGS; Vaginal Swab; Pregnancy; LIM Broth; MGI DNBSEQ-G400

Abbreviations: UTIs: Urinary Tract Infections; ICUs: Intensive Care Units; ACOG: American College of Obstetricians and Gynaecologists; LOGBS: Late-Onset GBS Disease; ACNM: American College of Nurse-Midwives; EOD: Early Onset Disease; EF: Enzymatic Fragmentation; RCA: Rolling Circle Amplification; PGAP: Prokaryotic Genome Annotation Pipeline; BGC: Biosynthetic Gene Cluster; LB: Luria-Bertani; BP: Base Pairs; ANI: Average Nucleotide Identity; GO: Gene Ontology; EC: Enzyme Commission; COG: Clusters of Orthologous Groups; GO: Gene Ontology; KO: KEGG Orthology; WHO: World Health Organization

Introduction

P. aeruginosa can colonize the vagina, however does not commonly cause vaginal infections. It is a primary pathogen that causes infections in respiratory tract, burn infections, bacteremia, urinary tract infections (UTIs), vaginal infections, diverse sepsis syndromes in medically fragile populations and intensive care units (ICUs). It produces biofilm, which increases its resistance to environmental stress, antibiotics, and host immune systems. This transient behaviour is important for establishing chronic infections, specifically in the chronic lung infection of cystic fibrosis patient due to adherence to medical devices [1]. The epidemiology, risk factors, maternal and neonatal outcomes of nosocomial acquisition of *P. aeruginosa* and its role in preterm premature rupture of membranes has been studied. Out of 63 women who received co-amoxiclav as a prophylactic antibiotic, 11 acquired *P. aeruginosa* vaginal carriage. All the five neonates born to these positive mothers were colonized or infected, out of which three died of fulminant sepsis. Vertical transmission was suggested between mother and neonates [2]. Group B streptococcus (GBS, *Streptococcus agalactiae*) colonizes in recto-vaginal tract of normal as well as pregnant women, and leads to serious complications such as urinary tract infections, chorioamnionitis and postpartum endometritis in the later.

Neonates with early-onset GBS disease develop sepsis, pneumonia and meningitis leading to severe neurological outcomes that may end in death. The bacterium's dangerous nature becomes evident through late-onset GBS disease (LOGBS) whose symptoms emerge after the first week of birth [3]. American College of Obstetricians and Gynaecologists' (ACOG) Committee in collaboration with American College of Nurse-Midwives (ACNM) published an interim update in 2019 on GBS screening for the prevention of GBS Early Onset Disease (EOD) in new-borns. According to this guideline it is recommended that all woman should undergo screening for GBS at 36 0/7–37 6/7 weeks of gestation, exception: intrapartum antibiotic prophylaxis for GBS due to GBS bacteriuria during the pregnancy or history of a previous neonatal GBS-infection [4-6]. LIM broth (Todd -Hewitt broth supplemented with selective antibiotics such as colistin and nalidixic acid is primarily used as a selective enrichment medium for GBS detection [7]. However, bacteria such as *Streptococcus faecalis*, *Lactobacillus, Gardnerella vaginalis*, *Mycoplasma hominis*, or other anaerobic bacteria resistant to the inhibitory agents can also grow in the LIM broth [8-10]. *P. aeruginosa* is a gram-negative, aerobic, rod-shaped, motile, chemoorganotrophic, bacterium present in diverse habitats including soil, water, animals, plants, and humans [11]. Due its resistance to several drugs and ability to degrade phenolic disinfectants it has been reported commonly in nosocomial infections [12,13].

Its virulence systems include: cell-associated factors (lipopolysaccharides, flagella, pili) and an array of extracellular secretions such as proteases, exotoxins, elastases, siderophores, extracellular polysaccharides, etc. [14,15]. Hemolysins or cytolysins are extracellular toxic proteins produced by many gram-positive and gram-negative bacteria that form pores in membranes host (RBCs, epithelial cells, leukocytes leading to cell lysis and ultimately cell death [16,17]. The first whole genome of *P. aeruginosa* was reported in 2000 for the strain PAO1, with a genome size of 6.3Mb (the largest reported at that time) [18]. The hemolytic trait in *P. aeruginosa* is reported to occur in environmental isolates as much as in clinical isolates. *P. aeruginosa* PAO1 causes plant damage when present in the rhizosphere and disruption of the hemolysin gene lowers virulence towards poplar and barley [19,20]. This is a first report on the WGS of a beta-hemolytic *P. aeruginosa* strain isolated from enriched LIM broth inoculated with a vaginal swab of pregnant woman from Gujarat, India.

Materials and Methods

Sample Collection and Isolation

A vaginal swab sample was collected with informed written consent from a pregnant woman with premature membrane rupture from a tertiary hospital in Ahmedabad for detection of GBS (group B *Streptococcus*). It was inoculated into LIM broth and incubated at 37^{0}C for 24-48 hours for enrichment. After incubation it was streaked onto Blood agar plate, incubated at 37ºC for 24-48 hours, and observed for colonies with clear zone of hemolysis (beta-hemolysis) as described earlier [21,22]. A beta-hemolytic Pseudomonas strain designated as GXDRC_02 (earlier designated as 830) was isolated on Luria-Bertani agar plate. Primary identification was done by routine microbiological tests such as gram staining, catalase, oxidase test, motility, pigmentation, etc. Glycerol stocks of the pure culture was prepared and stored at -20^{o}C. for further studies.

Whole Genome Sequencing (WGS)

WGS of the bacterial isolate was carried out by an optimized workflow using the high-quality genomic DNA (gDNA). The extraction procedure included cell lysis using DNA Extraction Buffer (DEB), phase separation with phenol: chloroform: isoamyl alcohol (25:24:1) for the removal of proteins and other contaminants, followed by ethanol based precipitation of the aqueous phase to obtain gDNA. It was then purified through silica-based spin column, ensuring intact and high-purity gDNA acceptable for the library preparation. Quantification of DNA was measured by using Qubit 3.0 Fluorometer (Thermo Fisher Scientific) and DNA integrity was confirmed by 0.8% agarose gel electrophoresis. Library preparation was carried out step-by-step as per the manual using the Twist Library Preparation Enzymatic Fragmentation (EF) Kit 2.0 (Twist Bioscience, USA) for the enzymatic fragmentation, end-repair, A-tailing, adapter ligation with dual indexes, and enrichment PCR. Fluorometric Quantification of final library was done with Qubit 3.0 Fluorometer (Thermo Fisher Scientific), and the library quality based on fragment distribution was assessed by Agilent Bioanalyzer 2100. This was an Illumina-compatible library, so a conversion step was performed to render the library compatible with the MGI DNBSEQ-G400RS platform. The constructed library was

processed for step-by-step conversion and sequencing of the library by using MGI's proprietary reagent kits.

Library conversion was done by AC-PCR kit and circularized to form single-stranded DNA nanoballs (DNBs) using Rolling Circle Amplification (RCA) chemistry. These DNBs were loaded onto the patterned flowcell using the MGI DNB loader followed by placing the flowcell into the MGI DNBSEQ-G400RS sequencer. High-throughput sequencing was done using paired-end 150bp read chemistry (2 x 150bp). The sequencing run generated high-quality reads with minimal duplication and error rates, which is required for the downstream bacterial genome assembly, annotation, comparative genome analysis and for other additional bioinformatics analysis parameters.

Assembly Quality Assessment and Annotation

The quality assessment of the genome completeness and contamination levels of the assembled genome was evaluated by using BUSCO and CheckM based on conserved single-copy orthologs and lineage-specific marker genes, respectively. The filtered high quality assemblies with contig lengths >500 bp were further annotated for downstream analyses. Multiple platforms were employed for genome annotation, including Bakta (version 1.9.4), the RASTtk pipeline (https://rast.nmpdr.org/rast.cgi), and the NCBI Prokaryotic Genome Annotation Pipeline v 6.6 (PGAP), to ensure comprehensive prediction of the gene and its functional assignment. For further exploration of the genomic content for orthologous pan-genome analysis and generation of a gene presence-absence matrix, which aids in the identification of conserved and additional genomic features across isolate using the Roary pipeline [23].

Genome Assembly and Annotation

The sequencing data was analysed using the BV-BRC genome analysis pipeline, which assists for quality filtering, adapter trimming, assembly, and annotation [24]. Raw reads were processed using Trim Galore v0.6.5, a wrapper around Cutadapt v2.2 [25], and FastQC. de novo assembly was conducted using Spades v4.0.0 with a criterion of minimum contig length threshold of 500bp [26]. Genome annotation was processed using the RAST toolkit (RASTtk), which provides annotated subsystem classification and gene function prediction [27]. Further functional annotation was carried out using eggNOG-mapper for orthology and GO term assignments [28], and the KEGG Automatic Annotation Server (KAAS) (https://www.genome.jp/tools/kaas/) for metabolic and pathway annotation [29,30].

Pangenome and Phylogenetic Analysis

The classification of core, soft-core, shell, and cloud genes among multiple isolates was performed by Pan-genome analysis using Roary v3.12.0 [23,31]. A core gene-based phylogenetic tree was built using the Codon Tree method within BV-BRC utilizing RAxML. Additionally, a reference-based phylogenetic analysis was performed along with a total of 52 related genomes from BV-BRC's Similar Genome Finder, followed by codon-alignment and reconstruction of phylogeny. The phylogenetic tree results were visualized using the Phylogenetic Tree Viewer tool within BV-BRC (https://www.bv-brc.org).

Genome Characterization and Species Identification

A combination of sequence-based and genome-wide comparative approaches were used for genome characterization and species identification. The sequences of 16S rRNA gene and housekeeping gene sequences were aligned against the NCBI database for initial taxonomic assignment using the BLAST algorithm. Further detailed resolution on sequence typing of the assembled genomes using allelic profiles of core genes to accurately assign sequence types (STs) was obtained through Ribosomal Multilocus Sequence Typing (rMLST) (https://pubmlst.org), cgMLST-based typing was done via the BacWGSTdb (http://bacdb.cn/BacWGSTdb/analysis_single.php) and MLST 2.0 tools (https://cge.food.dtu.dk/services/MLST-2.0/). The similarity searches for the whole genome was performed using the BV-BRC Similar Genome Finder and confirmed using Mash v2.3 [32]. The visualization of circular genome and CRISPR-Cas identification were conducted by using Proksee (https://proksee.ca).

Antimicrobial Resistance and Virulence Genes

The genes responsible for Antimicrobial Resistance (AMR) were ascertained using the BV-BRC (PATRIC) tools [33], CARD (https://card.mcmaster.ca0 [34], and ResFinder v4.1 (http://genepi.food.dtu.dk/resfinder) [35]. Virulence factors included genes associated with iron acquisition, toxin production, adhesion and colonization, secretion systems, immune evasion, invasion and intracellular survival, as well as regulatory systems controlling virulence gene expression were predicted and categorized using VFDB (http://www.mgc.ac.cn/cgi-bin/VFs/v5/main.cgi) [36]. These were also analysed by Victors, and PATRIC VF annotation tools and further confirmed with Abricate-VFDB for identified key virulence determinants.

Secondary Metabolites Biosynthetic Gene Cluster (BGC) Prediction

Secondary metabolite biosynthetic gene cluster (BGC) analysis was done by using antiSMASH v7.0 (https://antismash.secondarymetabolites.org) [37]. Additionally, analysis for bacteriocin-related gene clusters was done by BAGEL4 (http://bagel4.molgenrug.nl) specifically suggesting the strain's ability for the production of ribosomally synthesized antimicrobial peptides [38].

Mobile Genetic Elements and Pathogenicity

Mobile genetic elements were identified using multiple databases: PlasmidFinder (https://cge.food.dtu.dk/services/PlasmidFinder/) for plasmid-associated sequences; PHASTEST (https://phastest.ca) for prophages and its regions, and ISfinder (https://www-is.biotoul.fr/blast.php) for insertion sequences (IS elements). PathogenFinder v1.1 (Cosentino, et al. 2013) was employed to predict the likelihood of the isolate being a human pathogen [39].

Results

Enrichment, Isolation and Identification

Growth (after 24 hours of incubation) from LIM broth inoculated with a vaginal swab from pre-term pregnant woman was streaked on Blood agar plate, beta-hemolytic, pigmented colonies were observed. A pure culture was obtained and maintained as a glycerol stock at -20°C. The colonies were slightly raised, mucoid, translucent and brownish on Luria-Bertani (LB) agar plates (Figures 1 & 2). The gram-negative, catalase and oxidase positive isolate was designated as GXDRC_02 (earlier as strain 830) and used for further studies.

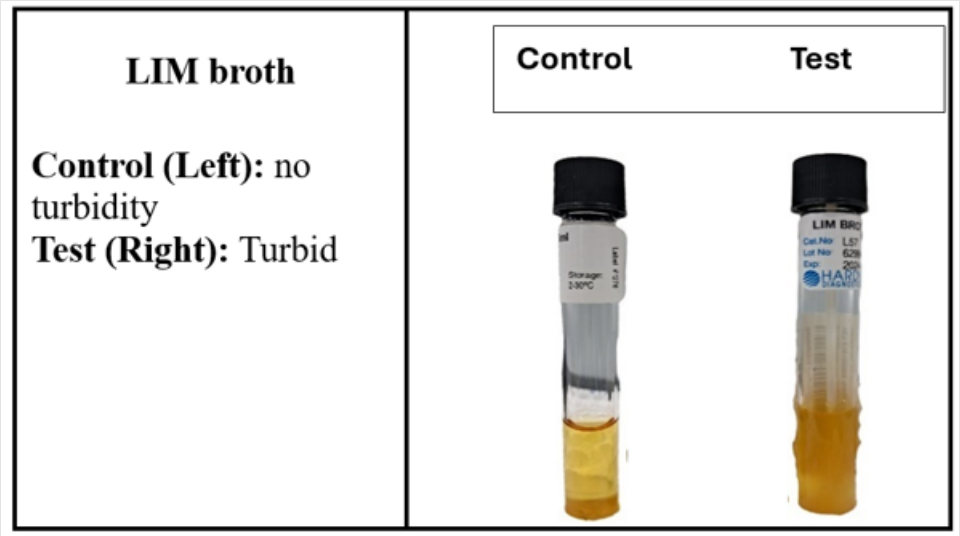

Figure 1: Enrichment in LIM broth inoculated with vaginal swab sample.

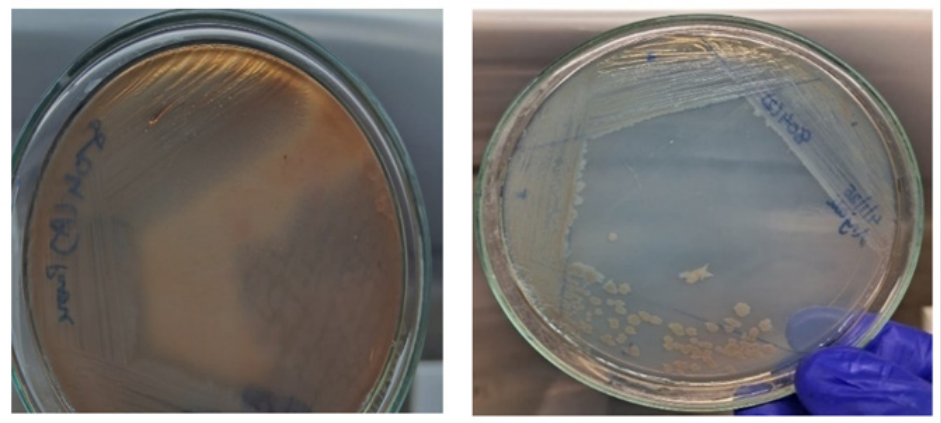

Figure 2: Left: beta-hemolysis (clear zone) on Blood agar plate; and Right: colony characteristics on LB agar plates.

Genome Assembly and Annotation

The assembled genome of *P. aeruginosa* GXDRC_02 by PATRIC consisted of 60 contigs, with a total length of 6,233,771bp. The assembly displayed a GC content of 66.38%, metrics, included an N50 of 306,098bp and L50 of 8, indicating a relatively contiguous assembly. No plasmids or chromosomal circularization were reported. Whereas, the genome was assembled using SPAdes genome assembler v4.0.0 contained 95 contigs, with 74 of these being at least 1000 base pairs (bp) long. The total genomic length across all contigs was 6465243bp (6.46Mb), of which 5,678,405bp accounted for contigs longer than 1000bp. The largest contig within the assembly stretches to 550357bp. The assembly had a GC content of 66.27%. The N50 statistic, a measure of assembly quality, is 266902bp, suggesting that

half of the total genome is contained in contigs of this length or longer. Similarly, the N90 is 63427bp, suggesting 90% of the genome is composed of contigs of at least this length. The average contig size weighted by the length of the contigs, known as auN, was 274162.9bp. The L50 and L90 values were 9 and 26, respectively, reflecting the number of contigs required to cover 50% and 90% of the genome. The assembly and annotation features are listed in Table 1.

Table 1: Assembly and annotation features of *P. aeruginosa* GXDRC_02 (BV-BRC).

Characteristics	Terms	
Taxonomy	cellular organisms > Bacteria > Pseudomonadati > Pseudomonadota > Gammaproteobacteria > Pseudomonadales > *Pseudomonadaceae* > *Pseudomonas* > *Pseudomonas aeruginosa* group > *Pseudomonas aeruginosa*	
Genome statistics	Number of contigs	60
	Size (bp)	6,233,771
	GC content (%)	66.38%
	Contig N50 value (bp)	306,098
	Contig L50 value	8
Genomic features	Genomic features CDS	5,913
	tRNA	55
	Repeat regions	9
	rRNA	3
Genome Quality	Completeness	97.9
	Contamination	1
	Overall remark	Good quality
Annotation features	Transporter (TCDB)	189
	Drug target (Drug Bank)	63
	Antibiotic resistance	50 (CARD), 5 (NDAR), 95 (PATRIC)
	Antibiotic sensitivity	Levofloxacin
Protein features	Proteins with functional assignments	4,183
	Proteins with EC	1,288
	Proteins with GO assignments	1.091
	Proteins with Pathway assignments	974
	Hypothetical proteins	1,100

Taxonomic Classification

The taxonomic placement of the genome was based on NCBI's bacterial taxonomy, confirmed through multiple layers of annotation and classified under the lineage (Table 1): The genome was annotated using genetic code 11, which is standard for bacterial genomes, and was assigned the unique genome identifier 287.42117.

Phylogenetic Analysis

Phylogenetic placement was achieved using reference and representative genomes selected by NCBI and aligned using PATRIC Global Protein Families (PGFams). Sequence alignments were performed with MUSCLE, and phylogenetic trees were constructed using RAxML, incorporating bootstrap support. The analysis employed Mash/MinHash for rapid similarity-based genome selection and confirmed that strain 830 clusters tightly with other *P. aeruginosa* strains. This robust phylogenomic evidence reinforces the strain's taxonomic classification. Whole-genome taxonomic identification using Similar genome finder (BV-BRC) confirmed the isolate as *P. aeruginosa*. Phylogenetic analyses, including codon tree-based, further established a close relationship with *P. aeruginosa* PAO1 (Figure 3). Additional sequence analysis using BIGSdb (targeting loci: BACT000001 to BACT000065) supported the species identification, with all alleles matching with known entries in the pubMLST database. No novel alleles were detected across the loci analysed. Species identification was consistently aligned with *P. aeruginosa*, with each locus showing highest similarity to this species across multiple matched alleles. To further validate the taxonomic identity, Average Nucleotide Identity (ANI) analysis was conducted using the assembled genome (501100830_Ps_830_contigs.fasta). The analysis revealed a highest ANI value of 99.429% with *P. aeruginosa* (GCA_008801675.1, ASM880167v1), along with high query and subject coverage values (96.5% and 93.7%, respectively).

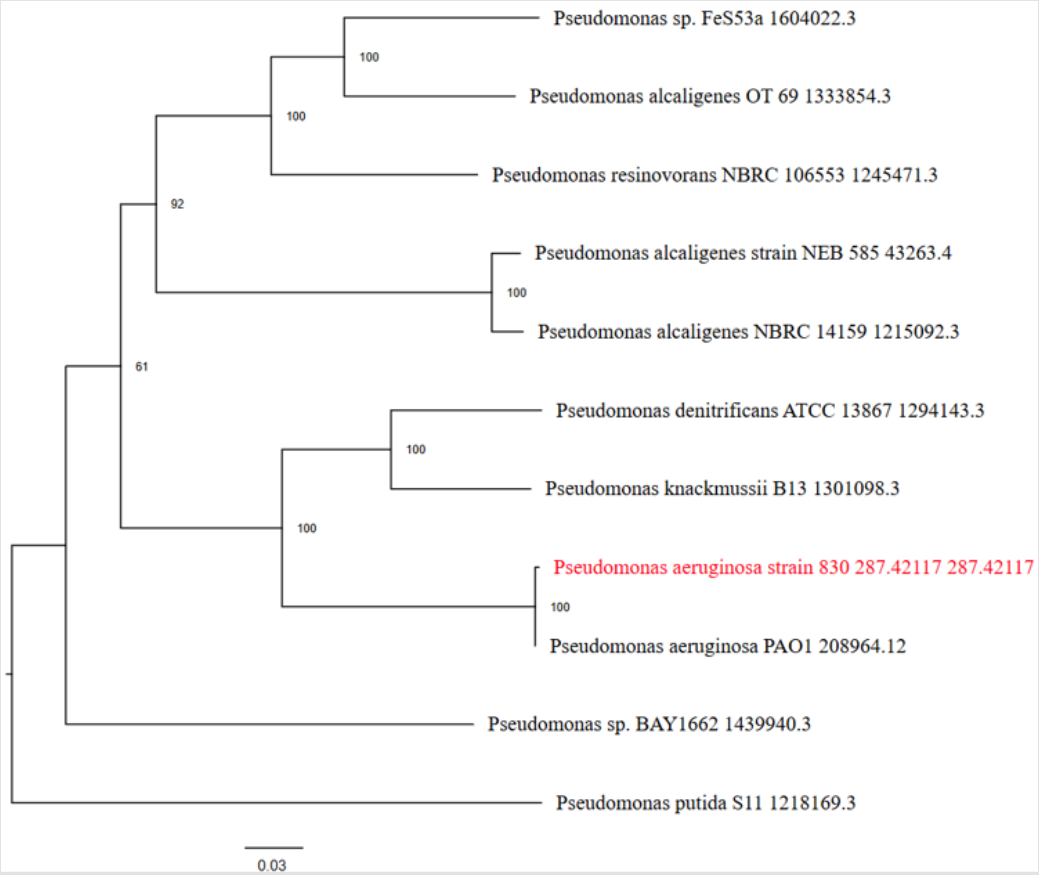

Figure 3: Phylogenetic tree of the isolate *P. aeruginosa* GXDRC_02.

Multiple reference genomes of *P. aeruginosa* also demonstrated ANI values exceeding 99.4%, with consistent high coverage, reinforcing the species-level assignment. Notably, the submitted genome showed markedly lower ANI values (<94%) with other Pseudomonas species, such as P. paraeruginosa, and significantly lower values (<84%) with more distantly related taxa. These findings, in conjunction with MLST-based allele typing, confidently confirm the organism as *P. aeruginosa*, with high confidence and no ambiguity in species delineation. The annotated genome using the RAST toolkit (RASTtk) integrated within the PATRIC framework comprised 5,913 protein-coding sequences (CDS), 55 transfer RNA (tRNA) genes, and 3 ribosomal RNA (rRNA) genes. Among the CDSs, 4,813 proteins were assigned functional annotations, while 1,100 proteins were identified as hypothetical. Functional annotation revealed that 1,288 proteins were associated with Enzyme Commission (EC) numbers, 1,091 with Gene Ontology (GO) terms, and 974 proteins were mapped to KEGG metabolic pathways (Table 1).

Protein Family Classification

PATRIC annotation includes two types of protein families; this data set contains 5,756 proteins belonging to the genus-specific protein families (PLFams), and 5,833 proteins belong to the cross-genus protein families (PGFams). This highlights the comprehensive functional coverage and supports evolutionary analysis across different bacterial taxa (Table 1).

Graphical Visualization

A circular genome map was generated depicting various genome elements including contigs, CDS (on both strands), RNA genes, virulence factors, antimicrobial resistance genes, GC content, and GC skew. The graphical representation also indicated subsystem affiliations, offering insight into the genome's structural and functional organization (Figure 4).

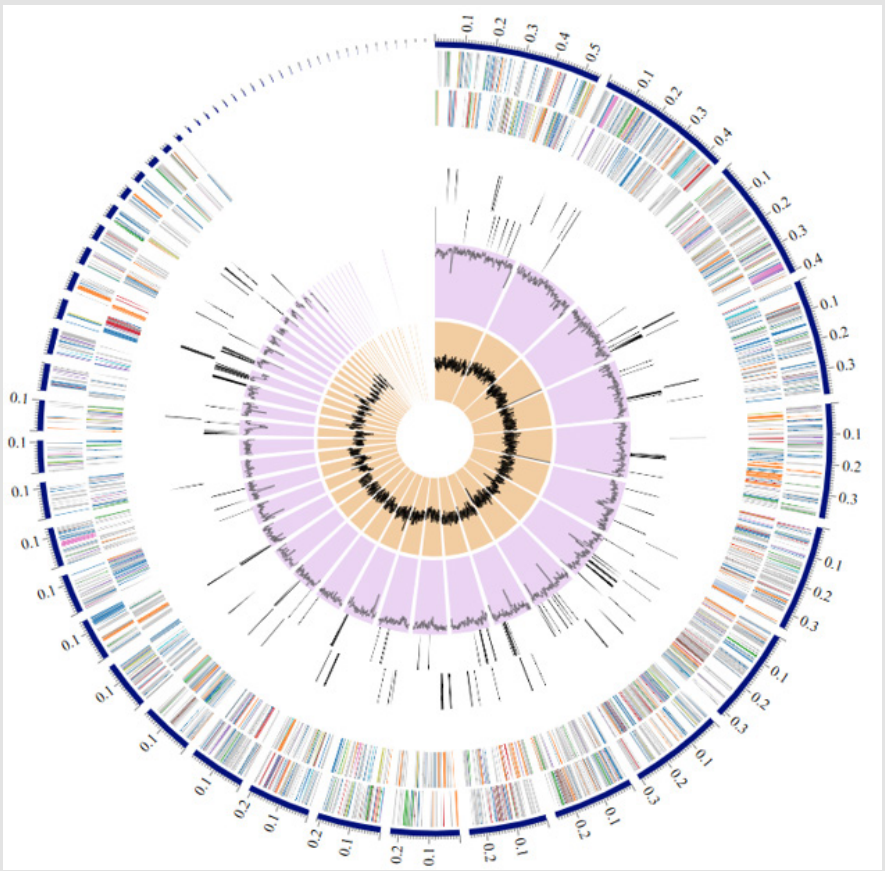

Figure 4: A graphical representation of the P. aeruginosa GXDRC_02 genome.

Subsystem Analysis

Subsystem annotation revealed a broad distribution of functional roles within the genome. Using the PATRIC RASTtk pipeline, genes were assigned to biological subsystems that represent coherent sets of proteins responsible for specific cellular processes or structural complexes. Major categories identified include Metabolism (124 subsystems, 1164 genes), Protein Processing (45, 263), Stress Response, Defense, and Virulence (36, 202), and Membrane Transport (28, 377). Additional assignments included functions in Energy Production, Cellular Processes, DNA/RNA Processing, and Cell Envelope structure and signalling, underscoring the organism's adaptability and pathogenic potential (Figure 5). The genome harbored numerous genes with known roles in virulence, antibiotic resistance, and host-pathogen interactions. Using curated databases, PATRIC identified: Virulence Factors from VFDB (230 genes), Victors (88 genes), and PATRIC_VF (1 gene); Antibiotic Resistance Genes from CARD (50), NDARO (5), and PATRIC (95); Drug Targets from DrugBank (63) and TTD (9); Transporters from the TCDB database (189 genes). These genes contribute to the bacterium's ability to survive hostile environments, evade host immunity, and resist treatment, making them important for clinical risk assessment (Table 1).

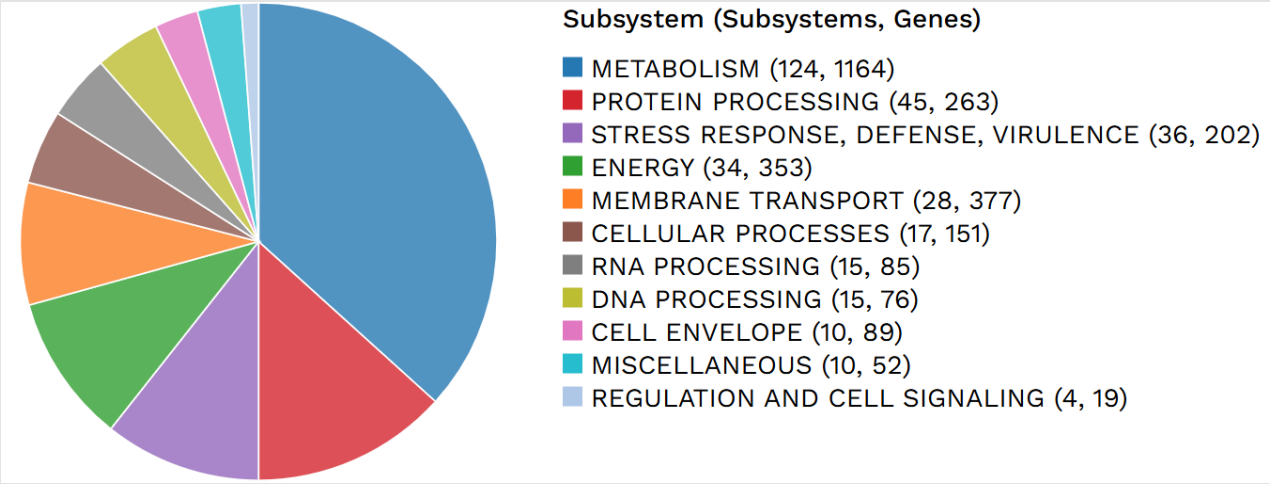

Figure 5: Subsystem analysis of the *P. aeruginosa* GXDRC_02 genome.

Antimicrobial Resistance Analysis

Phenotypic resistance prediction was performed using AdaBoost-based machine learning models. The analysis showed that *P. aeruginosa* GXDRC_02 is susceptible to levofloxacin. This prediction is based on genome sequence features and highlights the potential usefulness of levofloxacin for treating infections caused by this strain. A comprehensive AMR gene analysis identified multiple resistance mechanisms. Annotated AMR genes indicates a variety of mechanisms for antibiotic resistance. The majority of these genes are involved in utilizing efflux pumps for example: EmrAB-OMF, MacA, MacB, etc. to confer resistance, and proteins such as OccD1/OprD, OccK2/OpdF, OccK9/OpdG, etc. for modulating permeability to antibiotics. It also encodes enzymes involved in antibiotic inactivation (such as: APH(3')-II, OXA-50 family, PDC family, etc.), as well as protein (OxyR) modulating the expression of antibiotic resistance genes (Table 2). These findings underscore the strain's arsenal of resistance determinants and provide molecular insight into its potential drug resistance phenotype.

Table 2: Antimicrobial resistance gene.

AMR Mechanism	Genes
Antibiotic inactivation enzyme	APH(3')-II/APH(3')-XV, CatB family, OXA-50 family, PDC family
Antibiotic target in susceptible species	Alr, Ddl, dxr, EF-G, EF-Tu, folA, Dfr, folP, gyrA, gyrB, inhA, fabI, Iso-tRNA, kasA, MurA, rho, rpoB, rpoC, S10p, S12p
Antibiotic target replacement protein	FabG, fabV, HtdX
Efflux pump conferring antibiotic resistance	EmrAB-OMF, EmrAB-TolC, MacA, MacB, MdtABC-OMF, MdtABC-TolC, MexAB-OprM, MexCD-OprJ, MexCD-OprJ system, MexEF-OprN, MexEF-OprN system, MexHI-OpmD, MexHI-OpmD system, MexJK-OprM/OpmH, MexPQ-OpmE, MexPQ-OpmE system, MexVW-OprM, MexXY-OMP, TolC/OpmH, TriABC-OpmH
Gene conferring resistance via absence	gidB
Protein altering cell wall charge conferring antibiotic resistance	GdpD, PgsA
Protein modulating permeability to antibiotic	OccD1/OprD, OccD2/OpdC, OccD3/OpdP, OccD4/OpdT, OccD5/OpdI, OccD6/OprQ, OccD7/OpdB, OccK1/OpdK, OccK10/OpdN, OccK11/OpdR, OccK2/OpdF, OccK3/OpdO, OccK5/OpdH, OccK6/OpdQ, OccK7/OpdD, OccK8/OprE, OccK9/OpdG, OprB, OprB family, OprF
Regulator modulating expression of antibiotic resistance genes	OxyR

Functional Annotation using EggNOG-mapper

To enhance the functional annotation of protein-coding genes, the predicted proteome of *P. aeruginosa* GXDRC_02_was analyzed using EggNOG-mapper v2.1.12. A total of 5,481 proteins were subjected to orthology-based annotation using the DIAMOND search mode against the EggNOG v5.0 database. This annotation revealed that 14.62% of the proteins were associated with information storage and processing, which included functions such as DNA replication, recombination and repair, transcription regulation, translation, and ribosomal structure biogenesis. Approximately 9.47% of proteins were categorized under cellular processes and signalling, covering pathways related to cell wall/membrane/envelope biogenesis, signal transduction, post-translational modifications, intracellular trafficking, and secretion systems. Out of all proteins analysed, 5,478 proteins (99.9%) were assigned putative functional descriptions and Clusters of Orthologous Groups (COG) categories. COG assignments enabled broad classification of protein functions, contributing to the understanding of metabolic and cellular pathways in the strain. A total of 1,333 proteins were annotated with Gene Ontology (GO) terms, encompassing molecular functions, biological processes, and cellular components. Additionally, 3,390 proteins were mapped to KEGG Orthology (KO) identifiers, linking gene products to canonical metabolic and signalling pathways.

Notably, 5,051 proteins were also associated with conserved Pfam protein families, highlighting structurally or evolutionarily conserved domains. Further, 32.46% of annotated proteins were linked to metabolic processes, including those involved in amino acid transport and metabolism, carbohydrate metabolism, energy production and conversion, and inorganic ion transport. Also, 22.05% of the proteome fell under the poorly characterized category, highlighting genes with either general predictions or unknown functions. These findings reflect the functional diversity and metabolic versatility of this strain, consistent with its adaptability and environmental resilience (Figure 6).

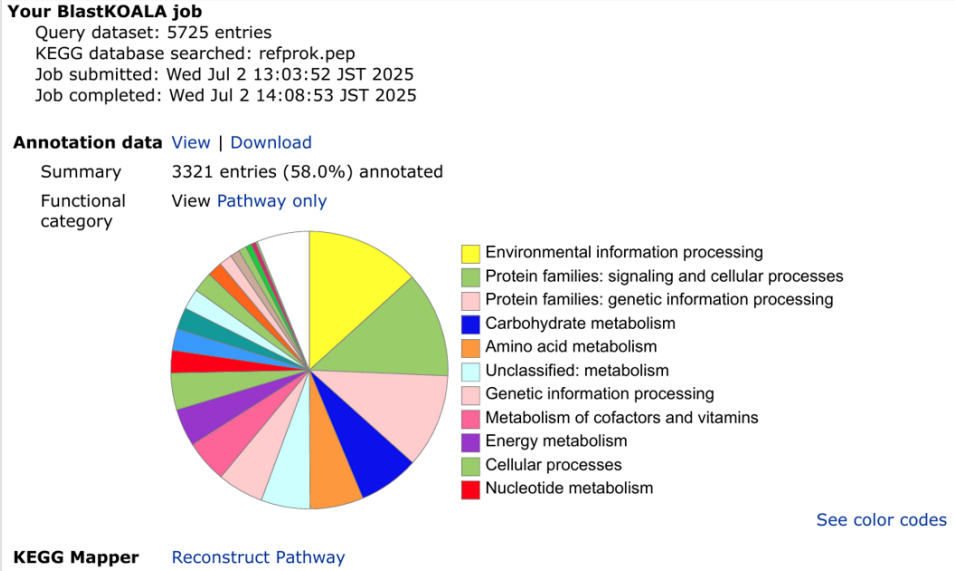

Figure 6: Functional Annotation using EggNOG-mapper of the *P. aeruginosa* GXDRC_02 genome.

Prophage Prediction and Characterization

Analysis of the genome using PHASTEST identified one intact prophage region within the genome, located between nucleotide positions 16,504 and 48,860 on contig_4, with a length of approximately 32.3 Kb. This region was classified as intact with a total score of 150, based on the presence of multiple structural phage proteins, including tail and plate proteins, and a high proportion of phage-related coding sequences. Specifically, 35 of the 41 predicted proteins (85.4%) in this region were phage-related, with an additional 1 hypothetical protein, resulting in a phage + hypothetical protein representation of 87.8%. These characteristics strongly support the presence of an actively integrated prophage.

Secondary Metabolite Biosynthetic Gene Cluster (BGC) Prediction

To explore the biosynthetic potential, the genome was analyzed using antiSMASH v8.0.1. A total of 17 putative biosynthetic gene clusters (BGCs) were identified under relaxed detection stringency, distributed across multiple genomic contigs. These BGCs repre-

sented diverse classes of secondary metabolites, suggesting a rich biosynthetic repertoire. Among the predicted clusters, several were identified with high similarity to known secondary metabolite pathways. Notably: An opine-like metallophore cluster (Region 1.1) was predicted with high confidence, closely resembling the pseudopaline biosynthetic pathway; a Type I NRPS cluster (Region 3.2) matched the azetidomonamide A/B pathway, indicating the potential for cyclic peptide production; Hydrogen cyanide biosynthesis genes (Region 19.1) were also detected with high similarity to known clusters, highlighting the strain's possible cyanogenic capability; a cluster in Region 25.1 showed high similarity to the pyochelin biosynthetic pathway, a well-characterized siderophore involved in iron acquisition; Additional clusters included: RiPP-like; NRPS-like; and hserlactone clusters, distributed across contigs 3, 9, 20, and 23; a phenazine biosynthetic cluster (Region 30.1), known for antimicrobial pigment production; Multiple NRP-metallophore and betalactone-associated clusters suggesting metal-chelation and potential antibiotic activity; One cluster (Region 24.1) was weakly similar to azotobactin D, while others were classified as novel with no close reference in the MiBIG database.

Pathogenic Potential

To assess the likelihood of human pathogenicity, the genome was analysed using PathogenFinder, the analysis predicted that the organism is a probable human pathogen, with a probability score of 0.752. A total of 682 protein sequences matched with known pathogenic families, whereas only 56 matched non-pathogenic families, covering 11.92% of the genome. The Z-score of 25.37 and a high prediction score of 1355.323 further support the confidence of classification it as a pathogenic strain. These matches include proteins with known virulence associations such as penicillin-binding proteins, type III secretion system components, siderophore receptors, and multiple hypothetical and conserved hypothetical proteins aligned with *P. aeruginosa* reference strains (e.g., PAO1 and LESB58).

Pangenome Composition and Genomic Plasticity

Pangenome analysis was performed to assess the genomic diversity and evolutionary dynamics among *P. aeruginosa* isolates. The analysis yielded a total of 6,592 gene families, partitioned into distinct genomic categories based on their prevalence across strains. The core genome, comprising genes present in 99–100% of strains, accounted for 5,269 genes, representing approximately 79.95% of the total pangenome. These genes are highly conserved and likely responsible for essential cellular functions and species-specific characteristics. No soft core genes (present in 95–99% of strains) or cloud genes (present in <15% of strains) were detected, suggesting a stable genomic backbone with minimal strain-specific outliers in this dataset. The remaining 1,323 genes (20.05%) were classified as shell genes, present in 15–95% of the strains (Figures 7 & 8). These genes represent the accessory genome and may encode functions related to environmental adaptability, stress response, host interactions, or antimicrobial resistance. The absence of cloud genes and soft core genes suggests a relatively closed pangenome, indicative of strong genomic conservation among the studied strains. However, the substantial shell genome component points to a degree of functional flexibility, potentially reflecting adaptations to clinical or ecological niches. Overall, these findings underscore the balance between genomic stability and adaptive potential in *P. aeruginosa*, contributing to its success as a ubiquitous opportunistic pathogen.

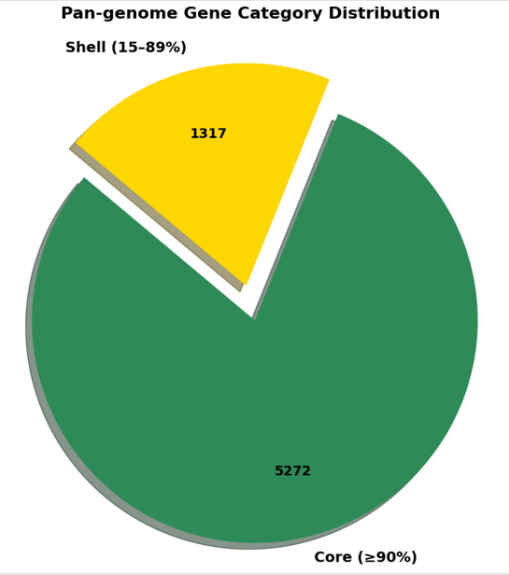

Figure 7: Pangenome gene category distribution of *P. aeruginosa* GXDRC_02 genome.

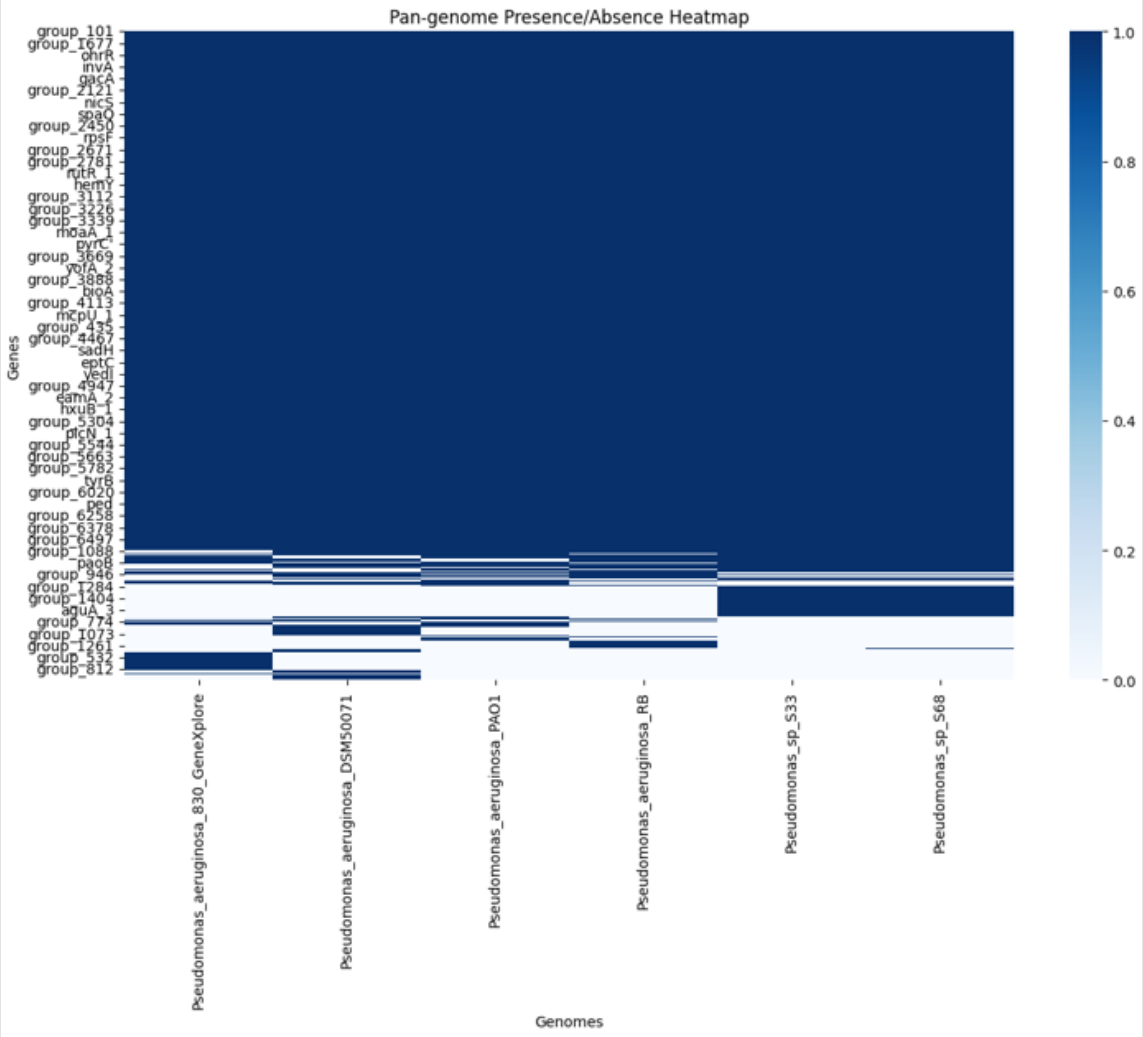

Figure 8: Pangenome Presence/Absence Heatmap of *P. aeruginosa* GXDRC_02 genome with related genomes.

Discussion

P. aeruginosa is one of the most widely studied opportunistic pathogen as per literature search in Google scholar as on date (4,20,000 results for opportunistic pathogen and 2,67,000 results for *P. aeruginosa* + opportunistic pathogen). The virulent strains of this bacterium cause a variety of acute and chronic infections which is of concern due its socio-economic burden to healthcare worldwide [40]. The World Health Organization (WHO) has classified *P. aeruginosa* among the ESKAPE (Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, *Pseudomonas aeruginosa*, and Enterobacter species) pathogens for discovery of novel therapeutics and of critical need [41]. GBS detection during 35-37 week of pregnancy is mandatory in the U. S. and other developed countries. As mentioned in the introduction, LIM broth is recommended as a selective and enrichment medium for this purpose, but it also allows the growth of other bacteria resistant to the inhibitory substances that may also cause complications during pregnancy solely or by interactions with other microflora. This concept of interactions between pathogens or microbial flora that have an impact on virulence in multi-strain and multi-species infections has been reviewed using *P. aeruginosa* as a case example [42].

Phylogenetic analyses of *P. aeruginosa* GXDRC_02 established a close relationship with the widely studied *P. aeruginosa* PAO1 a derivative (resistant to chloramphenicol) of the original strain POA isolated from a wound in Australia in 1954 [43]. Taxonomic identity based on ANI revealed a highest ANI value of 99.429% with *P. aeruginosa* (GCA_008801675.1 and ASM880167v1), along with high query and subject coverage values (96.5% and 93.7%, respectively). PAO1 and PA14 are the two most frequently used *P. aeruginosa* strains for research, the former being less virulent. The genome size of *P. aeru-*

ginosa GXDRC_02 was approx. 6.23 - 6.46Mb (66.38% G+C), containing 5913 CDS; whereas the genome of PAO1 is 6.3 Mbp (66.6% G + C content) in size, possesses 5,700 genes with 5,584 predicted open reading frames (ORFs) [44,45]. Among the bacterial genomes sequenced as on date Pseudomonas has one of the largest genome (7.0 Mb) which is at the higher end of the genome size [46]. This isolate possessed genes that encode proteins with known virulence associations such as penicillin-binding proteins, type III secretion system components, siderophore receptors, and multiple hypothetical and conserved hypothetical proteins aligned with *P. aeruginosa* reference strains (e.g., PAO1 and LESB58).

These results are consistent with the established clinical relevance of *P. aeruginosa* as an opportunistic pathogen, particularly in immunocompromised hosts. The presence of such virulence-related protein families reinforces the genome's pathogenic capacity and supports its prioritization for further investigation of virulence mechanisms and resistance traits. Various virulence factors that enable this bacterium to cause infection and regulatory systems controlling these factors is elaborated by Veetilvalappil, et al. [47], therefore it is important to understand/study the virulence of *P. aeruginosa* GXDRC_02 strain for better clinical management of infections in various settings. PHASTEST revealed the presence of an actively integrated prophage in the genome. The most common phage homologs identified within this region were associated with several Pseudomonas-specific phages, including PHAGE_Pseudo_YMC11/02/R656 (8 hits) and PHAGE_Pseudo_phiCTX (5 hits), among others. The GC content of the prophage region was 66.53%, closely matching that of the host genome, suggesting evolutionary adaptation. The role of plasmids, phages or phage-plasmids elements in bacterial evolution and horizontal gene transfer has been reviewed [48]. The identification of this intact prophage underscores the genomic plasticity of *P. aeruginosa* and indicates potential for horizontal gene transfer, which may contribute to strain-specific variation in virulence, stress response, or antimicrobial resistance.

Conclusion

The present study reports for the first time the isolation of a beta-hemolytic *P. aeruginosa* strain from LIM broth inoculated with a vaginal swab from a woman with preterm membrane rupture from Gujarat, India. Therefore, whole genomic sequencing and bioinformatics analysis of *P. aeruginosa* GXDRC_02 was carried out to reveal its identity with other related stains, metabolic and AMR profiling, virulence, pathogenicity, etc. It highlights the importance of surveillance for this opportunistic pathogens besides GBS during pregnancy, and other nosocomial infections. Collectively, these findings also underscore the biosynthetic versatility of *P. aeruginosa* GXDRC_02 and supports its potential for producing diverse natural products, including siderophores, antimicrobial compounds, and signalling molecules.

Acknowledgement

The authors thank Niraj Sheth, Dean, Gujarat Vidyapith for academic support to ADP, and Taral Patel for technical assistance.

Conflict of Interest

The authors declare no conflict of interest.

Contributions

AP, SSC, SRS, VNU DDJ have conceptualized, designed, provided resources and written the original draft paper. ADJ, HPP and VNU have carried out the microbiological experimental work. NGS and Bioinformatics analysis and methodology was written by JN and ZS. All the authors have reviewed and provided inputs for the final submission.

Funding Sources Declaration

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ethical Statement

Informed written consent has been obtained from the human subject for the clinical sample used in this research work. No human clinical trails or unethical practices have been involved.

References

1. Spagnolo AM, Sartini M, Cristina ML (2021) *Pseudomonas aeruginosa* in the healthcare facility setting. Rev Med Microbiol 32(3): 169-175.
2. Casetta A, Audibert F, Brivet F, Boutros N, Boithias C, et al. (2003) Emergence of nosocomial *Pseudomonas aeruginosa* colonization/infection in pregnant women with preterm premature rupture of membranes and in their neonates. Journal Hosp Infect 54(2):158-60
3. Moreira de Sá RA, Neto ARB, da Câmara França BE (2022) Group B Streptococcus and Pregnancy. In: Moreira de Sá RA, Fonseca, EBd (Eds.,), Perinatology. Springer, Cham.
4. (2019) American College of Obstetricians and Gynecologists. Prevention of group B streptococcal early-onset disease in newborns: ACOG Committee Opinion, Number 782. Obstet Gynecol 134(1): e19-e40.
5. Puopolo KM, Lynfield R, Cummings JJ (2019) Management of infants at risk for Group B Streptococcal disease. Pediatrics 144(2): e2019188.
6. Schrag SJ, Verani JR (2013) Intrapartum antibiotic prophylaxis for the prevention of perinatal group B streptococcal disease: Experience in the United States and implications for a potential group B streptococcal vaccine. Vaccine 31(4):D20-D26,
7. Rosa-Fraile M, Spellerberg B (2017) Reliable detection of group B Streptococcus in the clinical laboratory. Journal Clin Microbiol 55(9): 2590-2598.
8. Baud A, Hillion KH, Plainvert C, Tessier V, Tazi A, et al. (2023) Microbial diversity in the vaginal microbiota and its link to pregnancy outcomes. Sci Rep 13: 9061.
9. Bennett PR, Brown RG, MacIntyre DA (2020) Vaginal microbiome in preterm rupture of membranes. Obstet Gynecol Clin North Am 47(4): 503-521,

10. Wali IE, Sorour SE, Abdalla MAH (2007) Assessment of different methods for detection of group B streptococci carriage among pregnant females. Egyptian Journal of Med Microbiol 16(4): 593-597

11. Hardalo C, Edberg S (1997) *Pseudomonas aeruginosa*: assessment of risk from drinking water. Crit Rev Microbiol 23(1): 45-75.

12. Okino N, Ito M (2007) Ceramidase Enhances Phospholipase C-induced Hemolysis by *Pseudomonas aeruginosa*. J Biol Chem 282(9): 6021-6030.

13. Pang Z, Raudonis R, Glick BR, Lin TJ, Cheng Z (2018) Antibiotic resistance in *Pseudomonas aeruginosa*: Mechanisms and alternative therapeutic strategies. Biotechnol Adv 37(1): 177-192.

14. Balasubramanian D, Schneper L, Kumari H, Mathee KA (2013) Dynamic and intricate regulatory network determines *Pseudomonas aeruginosa* virulence. Nucleic Acids Res 41(1): 1-20.

15. Jurado-Martin I, Sainz-Mejias M, McClean S (2021) *Pseudomonas aeruginosa*: An audacious pathogen with an adaptable arsenal of virulence factors. Int J Mol Sci 22(6): 3128.

16. Fujita K, Akino T, Yoshioka H (1988) Characteristics of heat-stable extracellular hemolysin from *Pseudomonas aeruginosa*. Infect Immun 56(5):1385-1387.

17. Allam A, Hamza E, Morad E, Shafei E, Etriby E (2020) Biochemical and immunological characterization of haemolysin produced by *Pseudomonas aeruginosa* PAO1 isolated from burn wounds. Afri J Clin Exp Microbiol 21(2): 132-139.

18. Brinkman Fiona FSL, Brinkman SL, Winsor GL, Rachel E Done RE, Filloux A, et al. (2021) The *Pseudomonas aeruginosa* whole genome sequence: A 20th anniversary celebration. Adv Microb Physiol 79: 25-88.

19. Alonso A, Rojo F, Martínez JL (1999) Environmental and clinical isolates of *Pseudomonas aeruginosa* show pathogenic and biodegradative properties irrespective of their origin. Environ Microbiol 1(5): 421-430.

20. Das D, Baruah R, Roy AS, Singh AK, Deka Boruah HP, et al. (2015) Complete genome sequence analysis of *Pseudomonas aeruginosa* N002 reveals its genetic adaptation for crude oil degradation. Genomics 105(3):182-190.

21. Montague NS, Cleary TJ, Martinez OV, Procop GW (2008) Detection of group B streptococci in Lim broth by use of group B Streptococcus peptide nucleic acid fluorescent in situ hybridization and selective and nonselective agars. Journal Clin Microbiol 46(10): 3470-3472.

22. Upasani VN, Patel HP, Patel AR. Chettiar SS, Shah SS, et al. (2025) Whole Genome Sequence of *Pandoraea pnomenusa* GXDRC_01 isolated from a clinical sample, Gujarat, India". Asian J Biotechnol Bioresour Technol (AJB2T) 11(3): 45-56.

23. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, et al. (2015) Roary: Rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691-3693.

24. Olson RD, Assaf R, Brettin T, Conrad N, Cucinell C, et al. (2023) Introducing the Bacterial and Viral Bioinformatics Resource Center (BV-BRC): a resource combining PATRIC, IRD and ViPR. Nucleic Acids Res 6:51(D1): D678-D689.

25. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17: 10-12

26. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A (2020) Using SPAdes de novo Assembler. Curr Protoc Bioinformatics 70(1): e102.

27. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, et al. (2015) RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5: 8365.

28. Cantalapiedra CP, Herņandez-Plaza A, Letunic I, Bork P, Huerta-Cepas J (2021) eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol Biol Evol 38(12): 5825-5829.

29. Kanehisa M, Furumich M, Sato Y, Matsuura Y, Ishiguro-Watanabe M (2025) KEGG: biological systems database as a model of the real world. Nucleic Acids Res 53: D672-D677.

30. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007) KAAS: An automatic genome annotation and pathway reconstruction server. Nucleic Acids Res 35(suppl_2): W182-W185.

31. Jovčić B, Novović K, Filipić B, Velhner M, Todorović D, Matović K, et al. (2020) Genomic Characteristics of Colistin-Resistant Salmonella enterica subsp. enterica serovar infantis from Poultry Farms in the Republic of Serbia. Antibiotics 9(12): 886.

32. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, et al. (2016) Mash: fast genome and metagenome distance estimation using MinHash. Genome Biol 17(1):132.

33. Gillespie JJ, Wattam AR, Cammer SA, Gabbard JL, Shukla MP, et al. (2011) PATRIC: The comprehensive bacterial bioinformatics resource with a focus on human pathogenic species. Infect Immun 79 (11): 4286-4298.

34. Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya AR, et al. (2023) CARD 2023: Expanded Curation, Support for Machine Learning, and Resistome Prediction at the Comprehensive Antibiotic Resistance Database. Nucleic Acids Res 51: D690-D699.

35. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, et al. (2020) ResFinder 4.0 for predictions of phenotypes from genotypes. Journal Antimicrob Chemother 75: 3491-3500.

36. Liu B, Zheng D, Zhou S, Chen L, Yang J (2022) VFDB: a general classification scheme for bacterial virulence factors. Nucleic Acids Res. 50(D1): D912-D917.

37. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, et al. (2023) AntiSMASH 7.0: New and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Res 51(W1): W46-W50.

38. Van Heel AJ, de Jong A, Song C, Viel JH, Kok J, et al. (2018) BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. Nucleic Acids Res 46: W278-W281.

39. Cosentino S, Voldby LM, Møller AF, Lund O (2013) PathogenFinder – distinguishing friend from foe using bacterial whole genome sequence data. PLoS ONE 8: e77302.

40. Kolimi P, Narala S, Nyavanandi D, Youssef AAA, Dudhipala N (2022) Innovative treatment strategies to accelerate wound healing: Trajectory and recent advancements. Cells 11(15): 2439.

41. De Olivera DMP, Forde BM, Kidd TJ, Harris PNA, Schembri MA, et al. (2020) Antimicrobial resistance in ESKAPE pathogen. Clin Microbiol Rev 33(3): e00181-e0019.

42. Rezzoagli C, Granato ET, Kümmerli R, (2020) Harnessing bacterial interactions to manage infections: a review on the opportunistic pathogen *Pseudomonas aeruginosa* as a case example. Journal Med Microbiol 69: 147-167.

43. Chandler CE, Horspool AM, Hill PJ, Wozniak DJ, Schertzer JW, et al. (2019) Genomic and phenotypic diversity among ten laboratory isolates of *Pseudomonas aeruginosa* PAO1. Journal Bacteriol 201.

44. Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, et al. (2000) Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. Nature 406: 959-964.

45. de Sousa T, Hébraud M, Dapkevicius MLNE, Maltez L, Pereira JE, et al.

(2021) Genomic and metabolic characteristics of the pathogenicity in *Pseudomonas aeruginosa*. Int J Mol Sci 22(23): 12892.

46. Poulsen BE, Yang R, Clatworthy AE, White T, Osmulski SJ, et al. (2019) Defining the core essential genome of *Pseudomonas aeruginosa*. Proc Natl Acad Sci USA 116: 10072-10080.

47. Veetivalappil VV, Manuel A, Aranjani JM, Tawale R, Koteshwara A (2022) Pathogenic arsenal of *Pseudomonas aeruginosa*: An update on virulence factors. Future Microbiol 17(6): 465-481.

48. Pfeifer E, Moura de Sousa JA, Touchon M, Rocha EPC (2021) Bacteria have numerous distinctive groups of phage-plasmids with conserved phage and variable plasmid gene repertoires. Nucleic Acids Res 18;49(5): 2655-2673.

ISSN: 2574-1241

DOI: 10.26717/BJSTR.2025.63.009834

Alpesh R Patel. Biomed J Sci & Tech Res

Submission Link: https://biomedres.us/submit-manuscript.php