Indian J. Fish., 73 (3): 133-149, 2026 doi:10.21077/ijf.2026.73.3.182102-14

Available online at: epubs.icar.org.in/index.php/IJF

Whole genome sequencing and comparative genomic analysis of *Photobacterium ganghwense* CF1.2, bacterial host of the lytic marine bacteriophage PhCF1.2: Insights into its mobilome and metabolic potential

Keerthi R. Babu[1,2], K. S. Sobhana[1*], Aleena Alex[1,3], K. P. Neethu[1,2], Md Umar[4], V. S. Jayasree[1,2] and Grinson George[1]

[1]ICAR-Central Marine Fisheries Research Institute, Ernakulam North P. O., Kochi-682 018, Kerala, India
[2]Cochin University of Science and Technology, Cochin University P. O., Kochi-682 022, Kerala, India
[3]Mangalore University, Mangalagangotri, Mangaluru-574 199, Karnataka, India
[4]Amity University, Devanahalli, Doddaballapura, Bengaluru-562 110, Karnataka, India

Abstract

Photobacterium species are abundant in marine ecosystems, where they help break down organic matter and serve as hosts of lytic bacteriophages that influence the dynamics of bacterial populations. Here, we report the draft genome sequence and comparative genome analysis of *Photobacterium ganghwense* strain CF1.2, isolated from Cochin Fisheries Harbour, Kochi, on the south-west coast of India, and identified as the host bacterium of the marine lytic phage PhCF1.2. The isolate was identified based on phenotypic and biochemical traits as well as 16S rRNA gene sequence (GenBank accession MZ489219) and further characterised by Illumina 2×150 bp paired-end whole genome sequencing. The draft scaffolded genome had a total length of 5.42 Mb with a GC content of 50.59%. The assembly showed high contiguity (N50=3,527,779 bp; L50=1), with the two largest scaffolds mapped to Chromosome 1 (3.53 Mb) and Chromosome 2 (1.89 Mb). Genome annotation predicted 4,807 protein coding genes. The genome contained 101 CAZyme encoding genes corresponding to 115 CAZyme family assignments. Annotation based on the clusters of orthologous groups (COG) assigned 3,963 genes (82.44% of the predicted proteome) to 26 functional categories. AntiSMASH analysis identified 10 biosynthetic gene clusters (BGCs) distributed in both chromosomes. Pairwise comparison with the reference genome of *P. ganghwense* strain C2.2 gave FastANI value of 98.77%, which was well above the species delimitation threshold of 95–96% and confirmed the taxonomic classification of CF1.2. The genome revealed a range of metabolic pathways, mobile genetic elements and biosynthetic capabilities, suggesting the adaptive capacity of this marine bacterium. The genome of CF1.2, the bacterial host of a newly identified lytic phage, is a useful resource for studying phage–host interactions, bacterial adaptation, and the ecological roles of *Photobacterium* species in marine ecosystems. The raw sequencing reads and genome assembly have been deposited in GenBank under BioProject accession PRJNA1490486, BioSample accession SAMN61429489 and Whole Genome Shotgun (WGS) accession JCAQNK000000000.

*Correspondence e-mail:
sobhana.ks@icar.org.in
sobhanapradeep11@gmail.com

Keywords:
Average nucleotide identity, Biosynthetic gene clusters, CAZymes, Clusters of orthologous groups, Genome annotation, Taxonomic placement

Received : 29.07.2026
Accepted : 18.09.2026

Introduction

Photobacterium is a genus of Gram negative, facultatively anaerobic Gammaproteobacteria within the family Vibrionaceae whose members are widely distributed in marine environments, occurring in seawater, marine sediments, saline lakes, the surfaces and gastrointestinal tracts of marine animals, and the light organs of fish, where several species establish bioluminescent symbioses through the luxCDABEG operon (Urbanczyk *et al*., 2010). The genus is one of the nine genera constituting the family Vibrionaceae (Labella *et al*., 2018), a large family that may represent up to 50% of the prokaryotic community in marine waters. Polyphasic

taxonomic studies have recognised ~27 valid *Photobacterium* species, and this number is likely to grow as cultivation independent approaches and whole genome sequencing unveil previously unrecognised diversity (Labella *et al*., 2017; Machado and Gram, 2017). Comparative genomic studies have revealed that, similar to other members of Vibrionaceae, *Photobacterium* species characteristically possess bipartite genomes comprising two chromosomes of unequal size, multiple *rrn* operons (Machado and Gram, 2017), and harbour a flexible accessory gene pool rich in loci associated with energy production and carbohydrate metabolism that supports adaptation to heterogeneous marine habitats (Lau *et al*., 2022). Similar genomic approaches have been used to understand host adaptation and symbiotic association in other marine bacteria such as *Bacillus aryabhattai* associated with ascidians (Merlin and Umar, 2022). The genus encompasses free living, symbiotic and pathogenic species, with pathogenicity restricted to a comparatively small subset of species, notably *P. jeanii*, *P. rosenbergii*, *P. sanctipauli* and the two subspecies of *P. damselae* (Labella *et al*., 2017). In contrast, marine environmental isolates of species such as *P. ganghwense* are increasingly recognised for their halophilic nature and ecological importance, with growing interest in their biotechnological potential, such as polyhydroxyalkanoate production (Lascu *et al*., 2021) and pesticide degradation (Jayasree *et al*., 2020; Mishra *et al*., 2021).

P. ganghwense strain CF1.2, is the specific bacterial host of the lytic marine bacteriophage PhCF1.2 isolated in our laboratory. A detailed characterisation of the host genome provides an important basis to understand phage susceptibility, genome evolution and bacterial adaptation in marine environments. In this study, we combined phenotypic observations with whole genome sequencing and comparative genomics to clarify the taxonomic status of *P. ganghwense* strain CF1.2 and delineate a comprehensive overview of its genomic architecture, mobilome and metabolic potential. Such taxonomic clarification is of interest in view of the recent description of the distinct Ganghwense clade in the genus *Photobacterium* (Labella *et al*., 2018), which highlights the evolutionary uniqueness of these organisms. These genomic data give a holistic view of the metabolic pathways and adaptive mechanisms that allow the bacterium to survive and thrive in diverse marine environments.

Materials and methods

Sample collection, isolation and preliminary characterisation of *Photobacterium ganghwense* strain CF1.2

The bacterial strain CF1.2, the specific host of the lytic marine bacteriophage PhCF1.2, was isolated from seawater sample collected from the Cochin Fisheries Harbour (9.99°N; 76.22°E), Kochi, south-west coast of India, on 18 February 2018, adhering to standard ISO guidelines (ISO, 2017). The isolate was initially characterised based on colony morphology, biochemical tests, physiological attributes and 16S rRNA gene sequencing using universal primers 27F and 1492R, and identified as *P. ganghwense* (Jayasree *et al*., 2020). The 16S rRNA gene sequence of *P. ganghwense* strain CF1.2 was compared with representative *Photobacterium* species from the NCBI GenBank database. The sequences were aligned with the ClustalW algorithm in MEGA XI (Tamura *et al*., 2021). Branch support was assessed by constructing a Neighbor-joining (NJ) phylogenetic tree based on the Kimura two-parameter (K2P) nucleotide substitution model with 1001 bootstrap replications.

This identity was subsequently verified through whole genome sequencing. For standard propagation, the strain was grown in BD Difco Nutrient Broth supplemented with 2.5% (w/v) NaCl at 28±2°C under continuous agitation at 129 *g*. Long term preservation of the isolate was done by storing the culture at -80°C in double strength Nutrient Broth supplemented with an equal volume of sterile 50% glycerol, resulting in a final glycerol concentration of 25% (v/v), while short term maintenance was carried out at room temperature in semisolid agar overlaid with sterile mineral oil (Jayasree *et al*., 2020).

Genomic DNA isolation, library preparation and whole genome sequencing

Genomic DNA was extracted from an overnight culture of *P. ganghwense* strain CF1.2 using the standard phenol chloroform extraction method (Sambrook and Russell, 2006). DNA quality and quantity were assessed using agarose gel electrophoresis and spectrophotometric analysis prior to library preparation. Whole genome sequencing was performed on the Illumina platform using paired-end reads of 2×150 bp. Raw sequence reads were subjected to quality assessment using FastQC (Andrews, 2020), followed by quality trimming and adapter removal using Fastp (Chen *et al*., 2018). The raw sequencing reads generated in this study have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession PRJNA1490486 and BioSample accession SAMN61429489. High quality processed reads were *de novo* assembled using SPAdes (SPAdes-4.2.0) (Bankevich *et al*., 2012). The resulting contigs were subsequently scaffolded using MeDuSa (Multi-Draft based Scaffolder) (Bosi *et al*., 2015), with the reference genome of *P. ganghwense* strain C2.2 (GenBank accession no. GCA_017329545.1), downloaded from NCBI, serving as the guide genome. The assembled genome was screened for potential contamination using the NCBI Foreign Contamination Screen (FCS-GX) (Astashyn *et al*., 2024) prior to downstream analyses. The assembled genome was submitted to NCBI under the accession JCAQNK000000000.

Assembly evaluation and functional annotation

Structural contiguity and standard assembly metrics of the scaffolded genome were evaluated using QUAST (Gurevich *et al*., 2013). Structural annotation and preliminary functional prediction were performed using Prokka (v1.14.6) (Seemann, 2014) and Bakta (Schwengers *et al*., 2021). Circular visualisation of the draft genome assembly showing coding sequences, RNA genes, GC content, GC skew and annotated genomic features was generated using Proksee (Grant *et al*., 2023). Prophage regions were identified using PHASTEST (Wishart *et al*., 2023). Visualisation of the predicted prophage regions was performed using the graphical output of PHASTEST.

Taxonomic confirmation and genome comparison

Taxonomic placement and genomic similarity boundaries were verified by calculating the Average Nucleotide Identity with

FastANI (Jain *et al*., 2018) against the *P. ganghwense* strain C2.2 reference genome. To explore evolutionary and taxonomic relationships within the genus *Photobacterium*, eight complete or draft assemblies representing type strains and distinct species were retrieved from the NCBI GenBank database: *P. ganghwense* DSM 22954^T, *P. ganghwense* JCM 12487^T, *Photobacterium gaetbulicola* Gung47^T, *Photobacterium profundum* SS9, *Photobacterium damselae* subsp. *damselae*, *Photobacterium phosphoreum* MIP2473, *Photobacterium swingsii* DSM 22953^T, and *Photobacterium sanguinicancri* JCM 19183^T (Vezzi *et al.,* 2005; Park *et al*., 2006; Kim *et al*., 2010; Gomez-Gil *et al*., 2011; Labella and Borrego, 2023). To evaluate structural genome conservation and chromosomal alignment, a synteny map was constructed comparing *P. ganghwense* CF1.2 against the complete reference genome of strain *P. ganghwense* C2.2 using Minimap2 (v2.24) (Li, 2018), with the asm5 preset. Filtered alignments were visualised using an in-house Python script.

Functional genomic analysis

Complete functional classification of the translated proteome into clusters of orthologous groups (COGs) was conducted using COGclassifier (v2.0.0), driven by an RPS-BLAST search against the NCBI COG database (Galperin *et al*., 2015) with an E-value threshold of 0.01. The Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology database (Kanehisa *et al*., 2025) was used to annotate the predicted proteome of *P. ganghwense* strain CF1.2 to reconstruct the metabolic pathways. The completeness of the pathways was evaluated by mapping the assigned KEGG Orthology (KO) identifiers to the logical Boolean definitions of the curated KEGG MODULE database. The KEGG Level 2 functional categories were summarised and visualised by bar plots in Python.

Carbohydrate active enzymes were profiled using the automated annotation pipeline on the dbCAN3 platform (Zheng *et al*., 2023). Predicted CAZymes were categorised as glycoside hydrolases (GH), glycosyltransferases (GT), carbohydrate binding modules (CBM), auxiliary activities (AA), carbohydrate esterases (CE), and polysaccharide lyases (PL). The distribution of CAZyme families was analysed using Python, following the methodology described by Umar *et al.* (2024).

Specialised genomic features

Specialised genome traits of the *P. ganghwense* strain CF1.2 were revealed by genome mining and functional annotation strategies. Insertion sequence (IS) elements were detected and classified using pyISFinder by homology to the curated ISfinder database (Siguier *et al*., 2006). The abundance of IS families and their distribution in the genome across the assembled scaffolds were summarised and visualised using Python. Biosynthetic gene clusters (BGCs) involved in secondary metabolite production were identified using antiSMASH version 8.0 (Blin *et al*., 2025) with default parameters. The genome was screened for known classes of secondary metabolite biosynthetic pathways, and the predicted BGCs were visualised with the graphical output generated by antiSMASH. Bioremediation-associated genes were manually curated from Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology assignments (Kanehisa *et al*., 2025) and from functional annotations produced by Bakta and Prokka. The genomic biosafety profile was determined by screening for virulence factors and acquired antimicrobial resistance determinants using AMRFinderPlus v 4.2.7 (Feldgarden *et al.*, 2021) against the core NCBI database.

Results and discussion

Phenotypic and biochemical characteristics of CF1.2

The isolate CF1.2 formed cream coloured, circular, convex, shiny colonies with an opaque centre on Nutrient Agar plates supplemented with 2.5% NaCl when incubated at 28±2°C for 48 h (Fig. 1). The isolate was found to be a halophilic, Gram negative, rod shaped bacterium that tested positive for oxidase, catalase, methyl red, indole production, and glucose fermentation, while it tested negative for Voges Proskauer (Table 1). The taxonomic identity of the isolate was confirmed by 16S rRNA gene sequencing followed by sequence similarity search using BLAST against the NCBI GenBank and EZBioCloud databases. The isolate CF1.2 exhibited 99.46%

Fig. 1. Colony morphology of *P. ganghwense* CF1.2 on Nutrient Agar plate supplemented with 2.5% NaCl

Table 1. Comparative phenotypic and biochemical profiles of *P. ganghwense* strain CF1.2 and type strain *P. ganghwense* FR1311^T

Phenotypic and biochemical tests	*P. ganghwense* CF1.2	Type strain (*P. ganghwense* FR1311^T)
Gram reaction	-ve	-ve
Cell morphology	Rod	Rod
Swarming on solid media	-ve	-ve
Growth on TCBS agar	+ve (Green colony)	+ve (Green colony)
Motility	+ve	+ve
Voges-Proskauer (VP)	-ve	-ve
Methyl Red (MR)	+ve	+ve
Indole	+ve	+ve
ONPG	+ve	+ve
Esculin hydrolysis	-ve	-ve
Glycerol	+ve	-ve
Citrate utilisation	+ve	+ve
Nitrate reduction	+ve	+ve
Growth in % NaCl		
1	+ve	+ve
3	+ve	+ve
5	+ve	+ve
7	+ve	+ve
8	-ve	-ve
Amino acid utilisation		
Arginine	+ve	+ve
Ornithine	-ve	-ve
Lysine	-ve	-ve
Acid from sugar		
D-Arabinose	-ve	-ve
Dextrose	+ve	+ve
Fructose	+ve	+ve
Galactose	+ve	+ve
Inositol	+ve	+ve
Inulin	-ve	-ve

sequence similarity to the type strain *P. ganghwense* FR1311 and was therefore identified as *P. ganghwense*. The 16S rRNA gene sequence has been deposited in the NCBI GenBank database under the Accession no. MZ489219.

This species was first described by Park *et al*. (2006), who proposed *P. ganghwense* sp. nov. based on the type strain FR1311^T isolated from seawater collected off Ganghwa Island, South Korea. The species was characterised as a Gram negative, facultatively anaerobic, halophilic, rod-shaped bacterium with Q-8 as the predominant respiratory lipoquinone and a genomic DNA G+C content of 44 mol% (Park *et al*., 2006). Later, the phylogenetic position of *P. ganghwense* was clarified by the extensive taxonomic revision of the genus carried out by Labella *et al.* (2018), based on 16S rRNA gene and multi-locus sequence analyses. In this study, the authors established the Ganghwense clade, including *P. aphoticum, P. aquae, P. galatheae, P. ganghwense, P. halotolerans, P. panuliri,* and *P. proteolyticum*. According to this classification, 16S rRNA gene sequence analysis placed strain CF1.2 in the Ganghwense clade. (Fig. 2) Phylogenetic placement is an important foundation for comparative genomic analyses because members of the same clade typically share conserved core housekeeping and metabolic genes, but exhibit significant variation in their accessory genomes, which reflects adaptation to different ecological niches and evolutionary histories (Labella *et al.,* 2018).

Whole genome sequencing, assembly and annotation

Following quality filtering, more than 87% of the total pre-processed reads of the sample were successfully aligned to the reference genome. A summary of the read alignment statistics is presented in Table 2. The average mapping quality was 51.7, indicating a reliable alignment of the sequencing reads. Genome coverage analysis showed

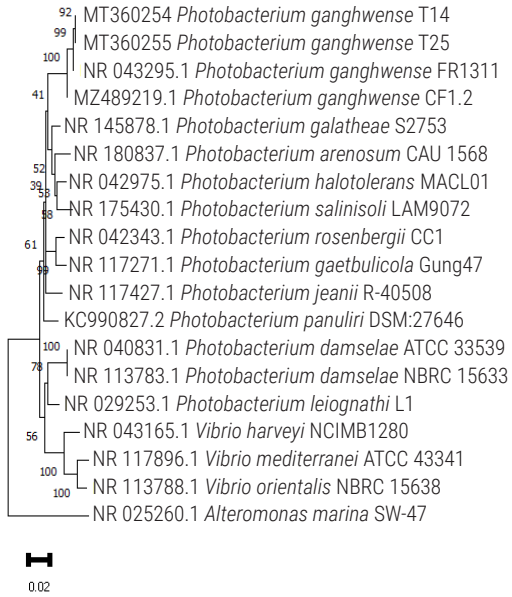

Fig. 2. Neighbor-joining phylogenetic tree based on 16S rRNA gene sequences showing the relationship of *P. ganghwense* strain CF1.2 with related species. The phylogenetic tree was constructed in MEGA XI using the Kimura two-parameter model with 1001 bootstrap replications. Bootstrap percentages are indicated on the nodes. Scale bar represents the number of substitutions per nucleotide position

Table 2. Summary of Illumina sequencing read quality and alignment statistics for *P. ganghwense* CF1.2

Total reads	Total paired reads	Raw read length	Total data	Total data >= Q30	Average GC%
5,229,044	2,614,522	151	789.59 (Mb)	75.02%	58.78%

that more than 89% of the reference contigs NZ_CP071325.1 and NZ_CP071326.1 were covered, with average sequencing depths of 88 and 64 fold, respectively. The average insert size of the paired-end reads was 284 bp.

De novo assembly using SPAdes yielded 13 contigs, which were subsequently scaffolded using MeDuSa with the complete genome of *P. ganghwense* C2.2 as the reference. The resulting draft scaffolded genome had a combined length of 5.42 Mb, a GC content of 50.59%, and an N50 value of 3.53 Mb (L50=1). The two largest scaffolds were assigned to Chromosome 1 and Chromosome 2 based on their alignment to the corresponding chromosomes of *P. ganghwense* C2.2. Genome annotation using Prokka predicted 4,807 protein coding genes, 13 rRNA genes, and a notably high tRNA gene count of 132. Relative to the reference C2.2 (5.75 Mbp), the CF1.2 genome is slightly smaller, which may reflect the loss of variable plasmids or genomic islands. A genus-wide comparative analysis of 35 genomes from 16 of the 28 currently described *Photobacterium* species revealed genome sizes of 4.2–6.4 Mb, with an average of ~4,750 protein coding genes per genome, and a pan-genome of 28,951 genes (Machado and Gram, 2017). In the genomic context, the genome size of *P. ganghwense* CF1.2 (5.42 Mb) and the number of predicted protein coding genes (4,807) fall within the reported size range of the genus. Comparative genomic studies show relatively conserved genome sizes, GC contents and core gene sets in *Photobacterium* species, with differences among strains arising mostly from modifications in accessory genomic regions responsible for ecological adaptation (Fuertes-Perez *et al*., 2021). The CF1.2 genome is marginally smaller than the closely related *P. ganghwense* C2.2 (5.75 Mb), but both strains maintain the characteristic bipartite genome organisation of the genus. This shows conservation of the core genomic architecture despite strain specific differences in mobile genetic elements, genomic islands, prophages and other accessory DNA that contribute to adaptation and niche specialisation.

Circular visualisation of the draft genome assembly of *P. ganghwense* CF1.2 (Fig. 3) depicts the distribution of genomic features across the genome including coding sequences (CDSs), regulatory regions, ncRNAs, tRNAs, rRNAs, tmRNA, replication origin, assembly gaps, GC content and GC skew. The distribution of CDSs throughout the genome indicates a conserved genomic organisation without pronounced clustering of coding regions. GC content and GC skew vary moderately and reflect nucleotide compositional asymmetry associated with replication related strand bias (Rocha, 2008). Similar genomic organisation has been reported for other *Photobacterium* species, with a conserved chromosomal architecture and a flexible accessory genome that contributes to ecological adaptation and metabolic diversification (Vezzi *et al*., 2005; Machado and Gram, 2017; Fuertes-Perez *et al*., 2021; Lau *et al*., 2022). The genome harbours a variety of genes involved in transport, metabolism, stress response, signal transduction and DNA repair including ABC transporters, phosphotransferase systems, histidine kinases and chemotaxis proteins, which indicates a large metabolic flexibility and adaptive potential of strain CF1.2 in the coastal marine environment.

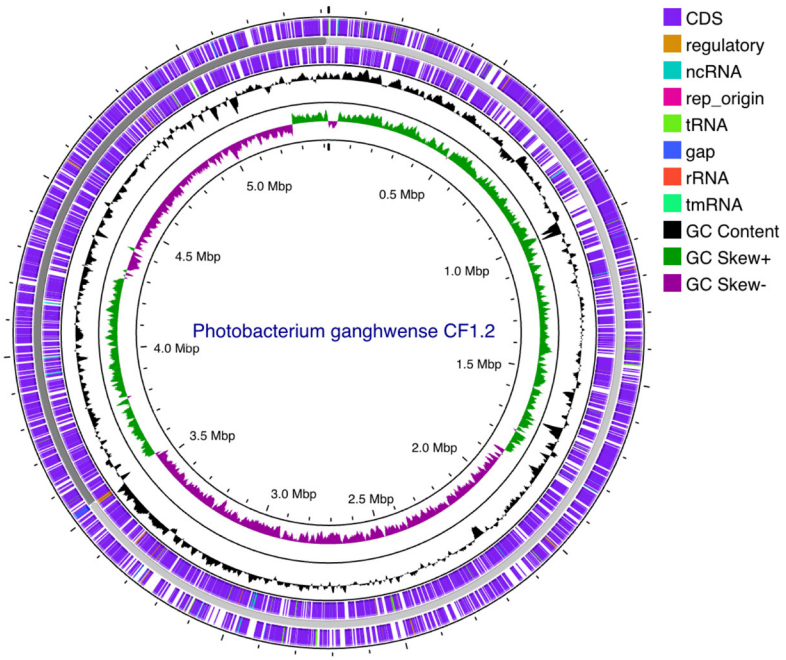

Fig. 3. Circular visualisation of the draft genome assembly of *P. ganghwense* CF1.2 generated using Proksee

Prophages constitute an important component of bacterial accessory genomes and play a significant role in genome evolution, horizontal gene transfer and ecological adaptation. Genome-wide prophage analysis identified four prophage regions on Chromosome 1 and one on Chromosome 2 of *P. ganghwense* CF1.2 (Fig. 4 a, b), comprising both complete and incomplete prophages. Similar prophage associated genomic elements have been reported in several *Photobacterium* species, where they contribute to strain diversification through the acquisition of accessory genes involved in environmental adaptation and host interactions (Machado and Gram, 2017). The identified regions contain genes encoding phage associated proteins, including integrases, terminases, portal, head, tail, tail fibre, baseplate, replication and regulatory proteins and several hypothetical proteins, suggesting that these prophages contain characteristic structural and functional modules of bacteriophages, although incomplete regions are likely remnants of ancestral phage integration events (Bobay *et al*., 2014). The dispersed location of the prophage regions along the chromosome indicates that several prophage integration events have taken place in the strain CF1.2 evolutionary history. These prophage regions might contribute to genome plasticity and might increase the adaptive potential of the bacterium in dynamic marine environments (Touchon *et al*., 2017). Prophage genetic diversification is a common feature of marine bacteria and a major mechanism to generate genome variability without affecting the conserved core genome (Canchaya *et al*., 2003; Fortier and Sekulovic, 2013).

Taxonomic confirmation and genome comparison

Average nucleotide identity (ANI)

The taxonomic assignment of strain CF1.2 to *P. ganghwense* was confirmed by pairwise ANI analysis based on the reference genome of *P. ganghwense* strain C2.2. The FastANI value between strains CF1.2 and C2.2 was 98.77%, which is significantly above the species delineation threshold of 95–96%. To further evaluate genomic relatedness, eight publicly available *Photobacterium* genome assemblies were retrieved; together with CF1.2 and C2.2, a total of 10 genome assemblies were included in the all-vs-all FastANI analysis (Fig. 5). The heatmap showed a cluster of *P. ganghwense* strains CF1.2, C2.2, JCM 12487T, and DSM 22954T with pairwise ANI values ranging from 98.77% to 100%, confirming their conspecificity.

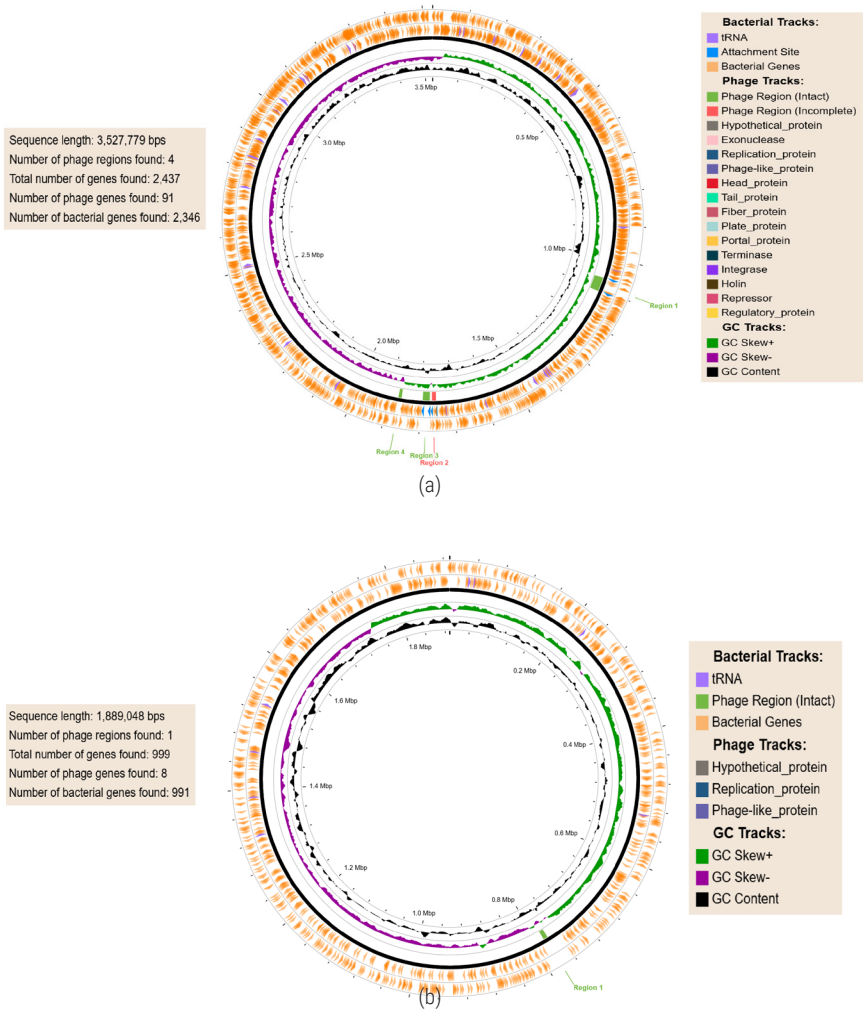

Fig. 4. Circular visualisation of the draft genome assembly of *P. ganghwense* CF1.2 showing the distribution of complete and partial prophage regions in (a) Scaffold 1 and (b) Scaffold 2, as identified using PHASTEST

Fig. 5. Heatmap generated by FastANI illustrating the genomic relationship of *P. ganghwense* CF1.2 with representative *Photobacterium* species based on average nucleotide identity (ANI)

The two genome assemblies representing the type strain FR1311^T (DSM 22954^T/JCM 12487^T) showed 100% ANI, while strain CF1.2 exhibited 98.77% ANI with C2.2 and JCM 12487T and 98.78% ANI with DSM 22954T, suggesting minor strain level genomic differences within the species. ANI values between *P. ganghwense* and other *Photobacterium* species were much lower, ranging from 79.04 to 80.58% for *P. gaetbulicola*, *P. phosphoreum*, *P. profundum*, *P. damselae* and *P. swingsii*, with relatively higher ANI of 86.53% for *P. sanguinicancri* with *P. swingsii*. These values are still far below the species boundary, indicating a clear genomic separation among the *Photobacterium* species studied. The high ANI values (>98.7%) obtained between strains of *P. ganghwense* confirm that strain CF1.2 belongs to *P. ganghwense* and demonstrate its close genomic relationship with the reference and type strains. In contrast, much lower ANI values with other *Photobacterium* species suggest clear species level separation, consistent with the established ANI threshold of 95–96% for prokaryotic species delineation. These results confirmed the taxonomic position of strain CF1.2 and supported its clustering as a new strain of *P. ganghwense* (Richter and Rossello-Mora, 2009; Jain *et al*., 2018).

Genome synteny

Synteny analysis between *P. ganghwense* CF1.2 and the complete reference genome of strain C2.2 showed a high degree of structural conservation between the two genomes (Fig. 6). Scaffold_1 and Scaffold_2 of CF1.2 primarily aligned to Chromosome 1 and Chromosome 2 of strain C2.2, respectively, supporting their assignment to the corresponding chromosomes based on reference-guided scaffolding and synteny. No large chromosomal translocations or inversions were observed, suggesting that the genome architecture is in general highly conserved. Any minor discontinuities in the alignments are more likely to be localised strain specific regions or small insertion/deletion events, rather than extensive structural variation. The high ANI values (>98.7%) and the high collinearity observed between the two genomes are indicative of the conservation of the core genome. Similar conservation of genome organisation has been reported for members of the genus *Photobacterium*, where genomic differences between strains are largely associated with accessory genetic elements rather than large scale chromosomal rearrangements (Machado and Gram, 2017). Such localised genomic variation,

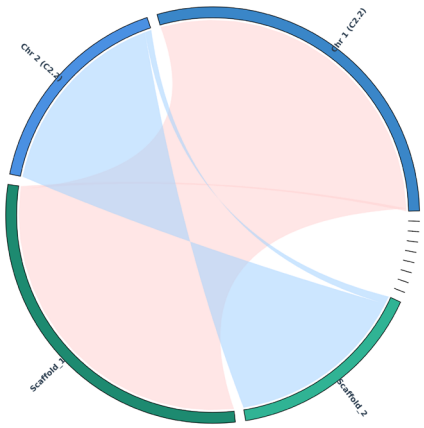

Fig. 6. Circular synteny map representing structural genomic conservation between *P. ganghwense* CF1.2 and reference genome of strain C2.2. Outer blocks represent chromosomes 1 and 2 of C2.2 and corresponding Scaffold_1 and Scaffold_2 of strain CF1.2. Inner ribbons represent homologous genomic regions, showing extensive collinearity and conservation of genome organisation

including prophages, genomic islands, and other mobile genetic elements, leads to strain level diversification, while maintaining the overall chromosomal framework (Treangen and Rocha, 2011). These results further corroborate the taxonomic assignment of strain CF1.2 as *P. ganghwense* and suggest that strain specific diversification has taken place without affecting the conserved genome architecture.

Functional genomic analyses

COG functional classification

The clusters of orthologous groups (COGs) database classifies proteins from complete genomes into evolutionarily conserved functional categories, facilitating the functional annotation of newly sequenced genomes (Shimoyama, 2022). COG-based functional annotation assigned 3,963 predicted protein coding genes (82.44% of the total proteome) of *P. ganghwense* CF1.2 to 26 functional categories (Fig. 7). Among the 26 categories, the largest functional categories were amino acid transport and metabolism (387 proteins), followed by transcription (316), translation, ribosomal structure and biogenesis (305), cell wall/membrane/envelope biogenesis (280), energy production and conversion (277) and general function prediction only (258). Other well represented categories included signal transduction mechanisms (241), carbohydrate transport and metabolism (234), inorganic ion transport and metabolism (212), and coenzyme transport and metabolism (201). In contrast, comparatively fewer proteins were assigned to mobilome related functions (31), extracellular structures (32), intracellular trafficking and secretion (48), cell cycle control (68), secondary metabolite biosynthesis (76), and defence mechanisms (97). In addition, 138 proteins were classified as having unknown functions, indicating the presence of uncharacterised genes in the genome.

The predominance of genes related to amino acid and carbohydrate metabolism, energy production, transcription and cell envelope biogenesis reflects the metabolic versatility and physiological adaptability of strain CF1.2 and its ability to metabolise different nutrient sources in marine environments. Similarly, the high abundance of genes related to signal transduction suggests an improved capacity to sense and respond to various environmental signals, including changes in salinity, nutrient availability and other physicochemical factors common in coastal habitats. The low relative abundance of mobilome associated genes is in line with a limited contribution of mobile genetic elements to the annotated functional repertoire, despite the presence of prophage regions identified elsewhere in the genome. COG profile suggests that *P. ganghwense* CF1.2 possesses a conserved functional repertoire shared by the genus *Photobacterium*.

COG-based methods have been used to functionally annotate *Photobacterium* genomes, from single strain analysis of *P. halotolerans* MELD1 (Mathew *et al*., 2016) and *P. kishitanii* FJ21 (Yin *et al*., 2022), to genus-wide comparative studies, comprising 35 (Machado and Gram, 2017) and 37 (Lau *et al*., 2022) strains. These studies have consistently revealed that the largest functional categories in *Photobacterium* genomes were, carbohydrate transport and metabolism (G), amino acid transport and metabolism (E), transcription (K) and signal transduction mechanisms (T). In contrast, the abundance of mobilome related genes (COG class X, prophages and transposons) varies considerably among strains, reflecting differences in prophage content and the distribution of mobile genetic elements (Mathew *et al*., 2016; Machado and Gram, 2017; Lau *et al*., 2022).

Unique glycerol utilisation pathway and phenotypic validation

Genomic analysis of *P. ganghwense* CF1.2 revealed a reconstructed metabolic pathway linking glycerol utilisation to polyhydroxyalkanoate (PHA) biosynthesis (Fig. 8a), highlighting the metabolic versatility of the strain and its potential significance for biotechnological applications. The strain had a phaCAB biosynthetic operon encoding

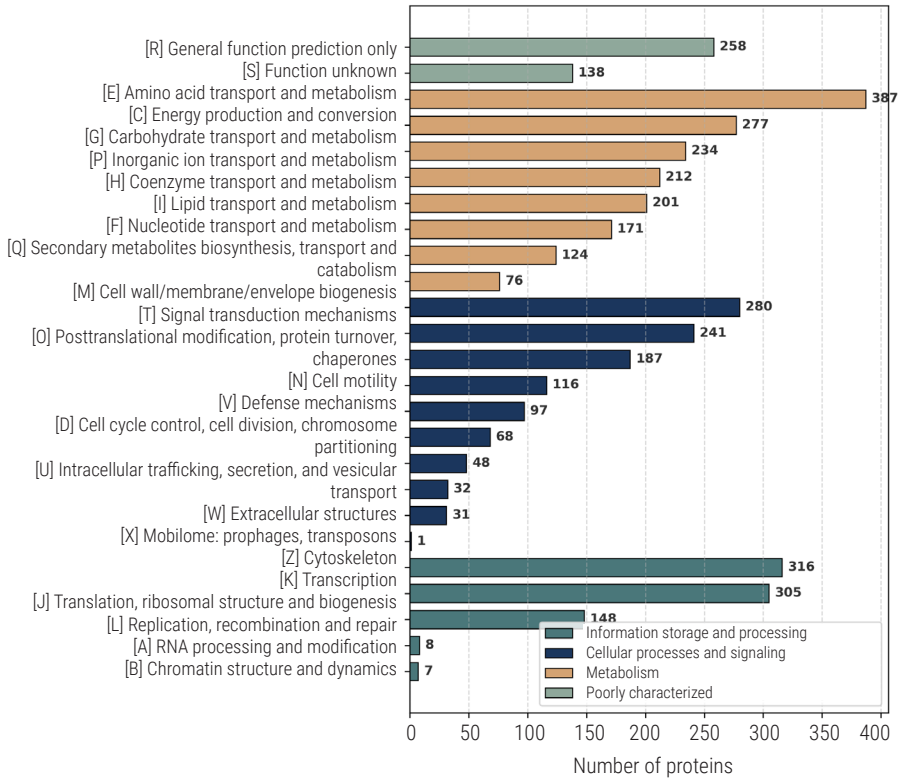

Fig. 7. COG-based functional classification of predicted proteins in *P. ganghwense* CF1.2

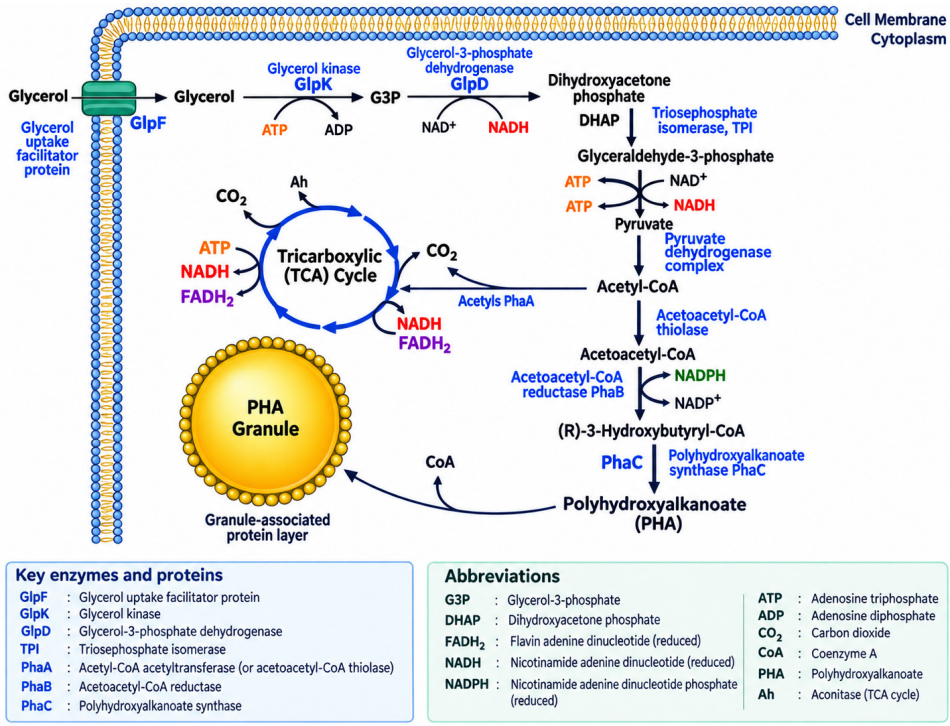

Fig. 8a. Schematic illustration of the reconstructed metabolic pathway linking glycerol utilisation and catabolism with PHA biosynthesis in *P. ganghwense* CF1.2. Glycerol is transported by the glycerol facilitator protein GlpF and phosphorylated to glycerol-3-phosphate by GlpK, which is then oxidised to DHAP by GlpD. The pathway proceeds through central carbon metabolism and acetyl-CoA, which is the precursor for PHA biosynthesis *via* PhaA, PhaB and PhaC

acetyl-CoA C-acetyltransferase (phaA, PG_ASS_03806), acetoacetyl-CoA reductase (phaB, PG_ASS_03805) and Class I PHA synthase (phaC, PG_ASS_03808). The gene organisation was in synteny with that reported for the previously characterised strain C2.2, in which the phaCAB operon is located on Chromosome 2 and is associated with a phasin-family (phaP) gene involved in PHA granule stabilisation (Lascu *et al*., 2021). CF1.2 also encoded genes related to glycerol uptake and metabolism, such as the glycerol facilitator (glpF, PG_ASS_03055), glycerol kinase (glpK, PG_ASS_03056) and glycerol-3-phosphate dehydrogenase components (glpD/glpABC; PG_ASS_03137 and PG_ASS_01965–01967). Glycerol is taken up by the cell through GlpF and phosphorylated by GlpK to glycerol-3-phosphate, which is further converted to dihydroxyacetone phosphate by GlpD (reconstructed as Fig. 8a). DHAP can enter central carbon metabolism and eventually give rise to pyruvate and acetyl-CoA. The acetyl-CoA then acts as a precursor for the biosynthesis of PHA through the sequential activities of PhaA, PhaB and PhaC. Furthermore, the presence of glpC and glpB suggests the potential for anaerobic glycerol-3-phosphate dehydrogenase activity, consistent with the glycerol utilisation pathway described in strain C2.2 (Lascu *et al*., 2021).

Phenotypic observations supported genomic predictions. CF1.2 was able to grow in Bushnell–Haas broth supplemented with glycerol (Fig. 8b), suggesting that the bacterium can use glycerol as the sole carbon and energy source under the tested conditions, as the medium contains mineral nutrients but no other source of carbon (Hossain *et al.*, 2024). This phenotype contrasts with the glycerol negative reactions obtained for the type strain FR1311^T and the coral reef isolates *P. ganghwense* T14 and T25 (Jayasree *et al*., 2020). The genomic and phenotypic evidence suggests that CF1.2 has the genetic potential to utilise glycerol and channel carbon through acetyl-CoA to PHA biosynthesis. However, presence of the corresponding genes and growth in glycerol are not sufficient to demonstrate PHA production from glycerol. Direct quantification and characterisation of PHA accumulation in glycerol supplemented conditions would be necessary to experimentally validate this potential.

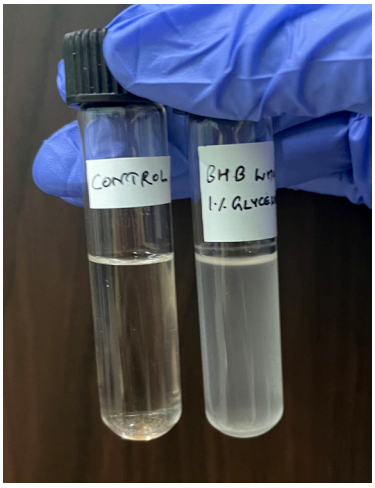

Fig. 8b. Growth of *P. ganghwense* CF1.2 in Bushnell–Haas broth supplemented with 1% glycerol.

KEGG pathway reconstruction

Functional annotation based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database assigned genes of *P. ganghwense* CF1.2 to multiple Level 2 functional categories, reflecting a broad metabolic repertoire (Fig. 9). The largest functional category was amino acid metabolism (206 genes), followed by carbohydrate metabolism (188 genes), metabolism of cofactors and vitamins (182 genes), membrane transport (152 genes), energy metabolism (143 genes), and signal transduction (137 genes). Additional genes were associated with nucleotide metabolism (90 genes), translation (79 genes), lipid metabolism (63 genes), folding, sorting, and degradation (52 genes), replication and repair (51 genes), xenobiotics biodegradation (41 genes), cell growth and death (21 genes), and transcription (4 genes). Detailed pathway reconstruction further revealed complete metabolic modules for central carbon metabolism, including glycolysis (EMP pathway, M00002), gluconeogenesis (M00003), and the Entner-Doudoroff (ED) pathway (M00008). Complete pathways for galactose metabolism (Leloir pathway, M00632) and galacturonate utilisation (M00631), together with biosynthesis of essential amino acids, including serine (M00020), threonine (M00018), cysteine (M00021), and lysine (M00016, 9/9 complete), indicate considerable metabolic flexibility and biosynthetic capability. The genome also encoded complete pathways for KDO_2-lipid A biosynthesis (M00060), menaquinone (M00116), ubiquinone (M00117) and non-mevalonate C5 isoprenoid biosynthesis pathway (M00096).

The predominance of genes involved in amino acid metabolism, carbohydrate metabolism, membrane transport, energy metabolism, and signal transduction indicates the metabolic versatility of *P. ganghwense* CF1.2 and its ability to adapt to nutrient variable marine environments. A similar functional profile was reported for *Photobacterium arenosum* WH24 where amino acid and carbohydrate metabolism were the most important metabolic categories, further highlighting the conserved metabolic framework of marine *Photobacterium* species (Pira *et al*., 2022). The complete glycolytic, Entner–Doudoroff and pentose phosphate pathways encoded by strain CF1.2 further support its ability to utilise a broad array of carbon sources including carbohydrates common in marine habitats, consistent with previous descriptions of *Photobacterium* species (Ogata *et al*., 1999; Park *et al*., 2006; Kanehisa *et al.*, 2025). Similar conservation of amino acid, carbohydrate, nucleotide, cofactor, and lipid metabolic pathways has also been observed among closely related Gammaproteobacteria, indicating that these core metabolic functions are evolutionarily conserved and contribute to ecological fitness (Nourdin *et al*., 2017). Similar metabolic profiles have been reported in other *Photobacterium* genomes, in which central metabolic pathways are highly conserved, whereas strain level variation is largely associated with accessory metabolic functions involved in ecological adaptation (Machado and Gram, 2017; Lau *et al.*, 2022).

Carbohydrate active enzymes (CAZyme) profile

The carbohydrate active enzyme repertoire of *P. ganghwense* CF1.2 was analysed using dbCAN3 to evaluate its potential for carbohydrate metabolism. In total, 101 CAZyme encoding genes were identified, corresponding to 115 CAZyme family assignments based on high confidence annotations supported by at least two tools (Table 3).

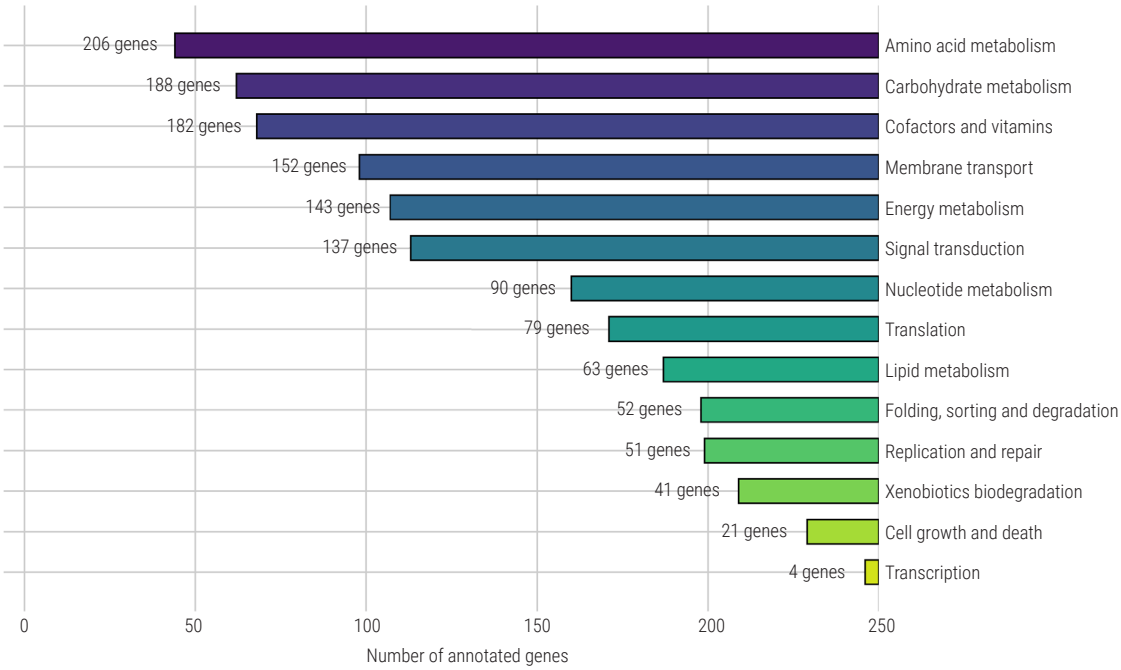

Fig. 9. KEGG Level 2 functional classification of annotated genes in *P. ganghwense* CF1.2. Bars represent the number of annotated genes in each major metabolic and cellular category

The largest class was glycoside hydrolases (GHs) with 43 family assignments (37.4%), followed by glycosyltransferases (GTs; 36; 31.3%), carbohydrate binding modules (CBMs; 19; 16.5%), auxiliary activities (AAs; 9; 7.8%), carbohydrate esterases (CEs; 6; 5.2%) and polysaccharide lyases (PLs; 2; 1.7%) (Fig. 10). The predominance of glycoside hydrolases and glycosyltransferases suggests predicted genomic potential for the degradation, modification, and biosynthesis of structurally diverse carbohydrates, which may contribute towards adaptation to marine environments rich in algal and microbial derived polysaccharides. The presence of carbohydrate binding modules may facilitate substrate recognition and catalytic efficiency by promoting enzyme binding to complex carbohydrate substrates. Similar CAZyme profiles have been reported in bacterial genomes, where carbohydrate active enzymes constitute an important component of the flexible genome and may contribute to nutrient acquisition and ecological adaptation, while the core housekeeping functions remain highly conserved (Machado and Gram, 2017; Lau *et al.,* 2022).

Table 3. Distribution of CAZyme family assignments and associated genes in *P. ganghwense* CF1.2 based on dbCAN3 annotation

CAZyme class	No. of CAZyme family assignments	Percentage
Glycoside hydrolase (GH)	43	37.4
Glycosyltransferases (GT)	36	31.3
Carbohydrate binding modules (CBM)	19	16.5
Auxiliary activities (AA)	9	7.8
Carbohydrate esterases (CE)	6	5.2
Polysaccharide lyases (PL)	2	1.7
Total CAZyme family assignments	115	99.9
Total genes associated with CAZyme families	101	-

Genome plasticity (Mobilome)

Insertion sequences

Insertion sequences (IS elements) are important mobile genetic elements that contribute to genome plasticity through transposition, gene disruption and genome rearrangements. Analysis using pyISFinder identified 25 IS elements belonging to eight IS families in *P. ganghwense* CF1.2 (Table 4). The most abundant category comprised unclassified (or novel) IS elements (nine elements; 36%), followed by the IS21 (five elements, 20%) and IS110 (four elements, 16%) families, while IS630, IS3, IS481, IS1595, and ISAS1 were represented by one or two copies each (Fig. 11). The IS elements were unevenly distributed in the assembled genome (Fig. 12) . Scaffold_1 and Scaffold_2 corresponding to Chromosomes 1 and 2 harboured 17 and 7 IS elements, respectively, while one more IS element was detected on Scaffold_3. No IS elements were found on the remaining small scaffolds. The higher abundance of IS elements on Scaffold 1 suggested the higher concentration of these mobile elements on

Table 4. Families of insertion sequences (IS) identified in *P. ganghwense* CF1.2

IS family	No. of IS elements	Percentage
Unclassified	9	36.0
IS21	5	20.0
IS110	4	16.0
IS630	2	8.0
IS3	2	8.0
IS481	1	4.0
IS1595	1	4.0
ISAS1	1	4.0

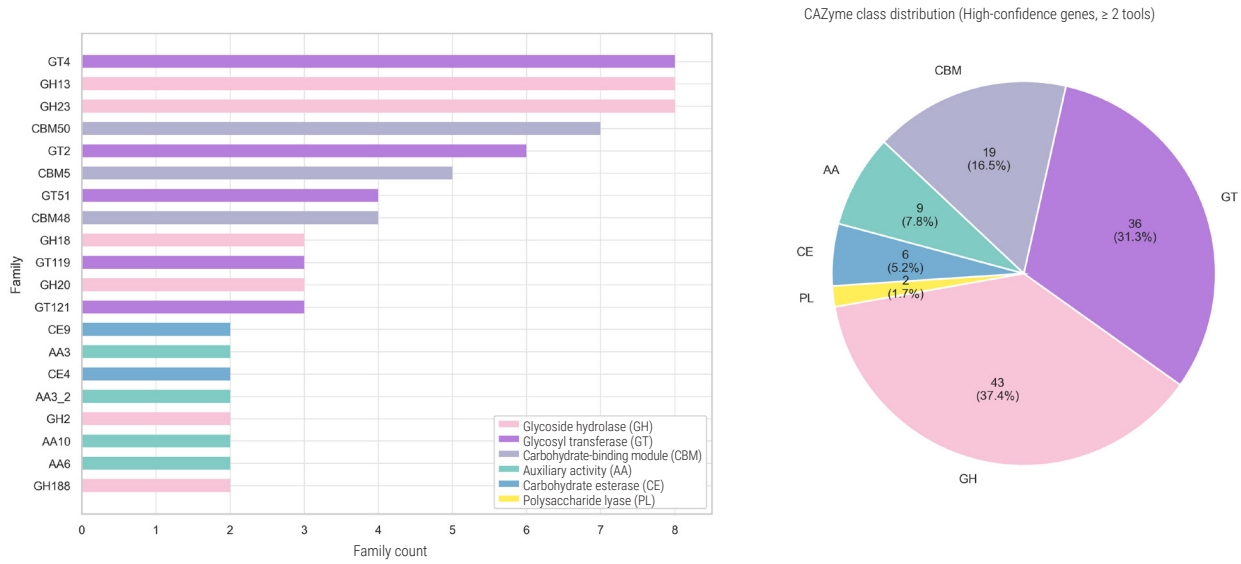

Fig. 10. Distribution of CAZyme families identified in *P. ganghwense* CF1.2. Bar chart shows the abundance of glycoside hydrolases (GH), glycosyltransferases (GT), carbohydrate binding modules (CBM), auxiliary activities (AA), carbohydrate esterases (CE) and polysaccharide lyases (PL)

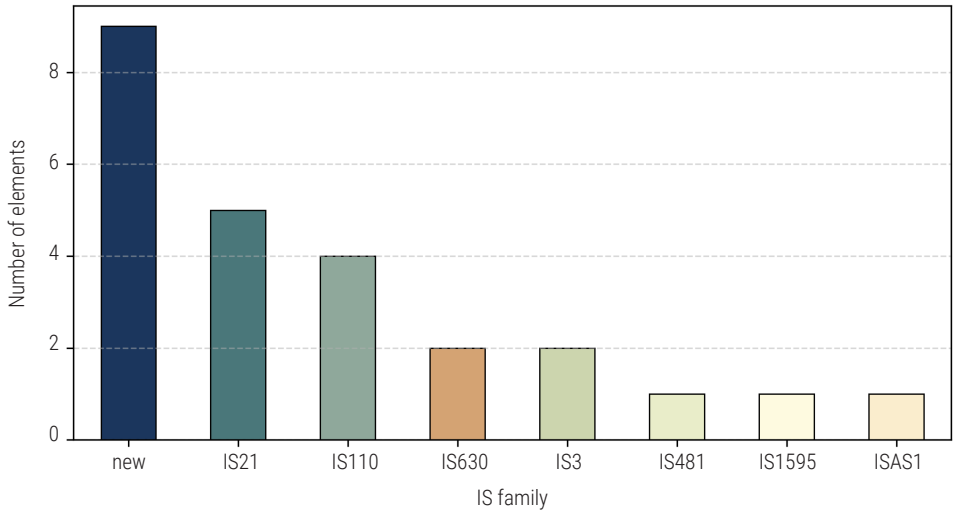

Fig. 11. Distribution of predicted insertion sequence (IS) families in *P. ganghwense* CF1.2, identified using pyISFinder

this chromosome which might contribute to local genome plasticity. Importantly, the core genes of the polyhydroxyalkanoate (PHA) biosynthetic operon were not disrupted by any of the identified IS elements, indicating that this metabolic pathway is conserved.

Comparative genomic analyses revealed the presence of a wide array of insertion sequences in *Photobacterium* species, which are involved in the diversification and adaptation of their genomes. De Palmenaer *et al*. (2008) noted an expansion of IS elements in the non-pathogenic *Photobacterium* spp. compared to pathogenic *Vibrio* spp. (*e.g.*, *V. cholerae*, *V. parahaemolyticus* and *V. vulnificus*), consistent with a lineage specific evolution of the mobilome. In parallel, Terceti *et al*. (2018) identified an IS10-family transposase, integrase genes and tRNA loci in an extremely variable genomic region of *P. damselae* subsp. *damselae*, indicating ongoing horizontal gene transfer and genome rearrangements in this marine pathogen.

Mobile genetic elements were recently found to frequently co-occur with phage defence systems in marine Vibrionaceae, and genomes harbouring IS-rich mobilomes often encode a larger set of antiviral defence determinants (Piel *et al*., 2022). In natural *Vibrio* populations, phage defences are frequently linked to mobile genetic elements and strains with IS-rich mobilomes tend to carry a broader range of defence systems, leading to increased resistance to lytic phage infections (Piel *et al*., 2022). Conversely, the moderate complement of 25 IS elements in *P. ganghwense*

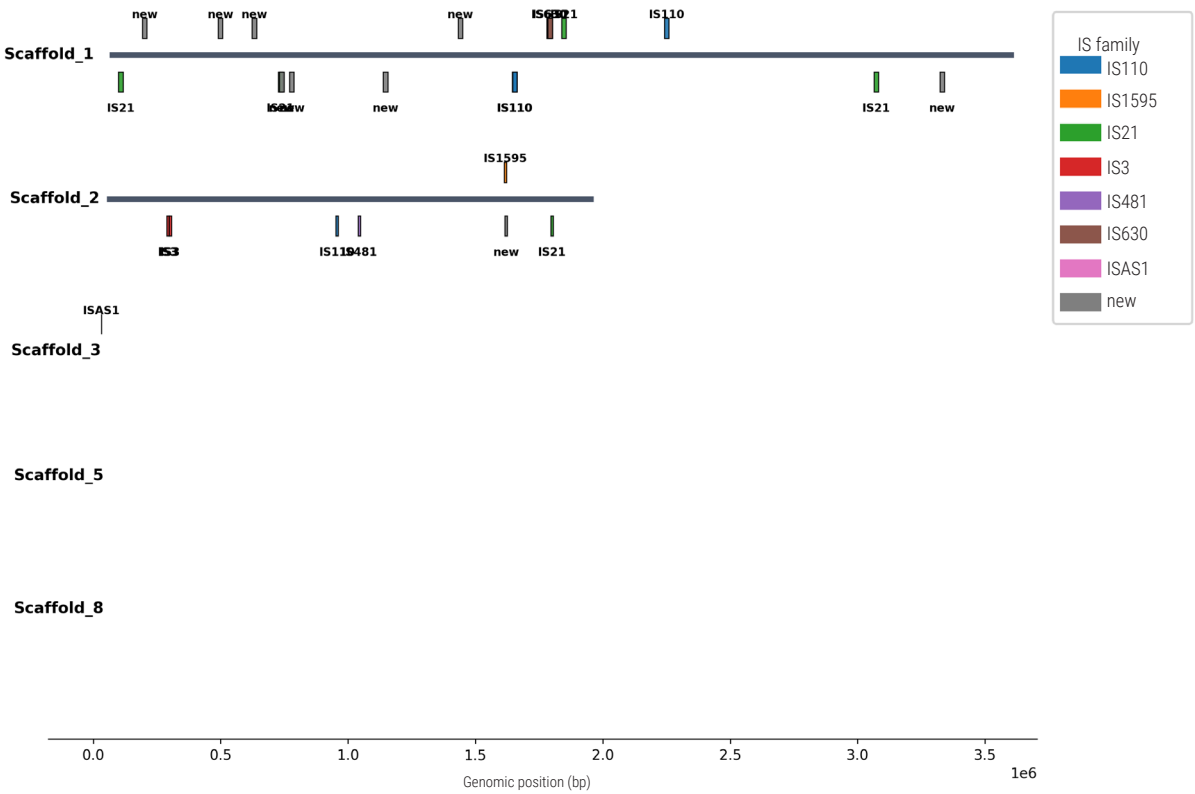

Fig. 12. Genomic distribution of predicted insertion sequence (IS) elements in *P. ganghwense* CF1.2 generated by pyISFinder

CF1.2 suggests a mobilome that can support genome plasticity without large IS expansion. CF1.2 is a permissive host of the lytic marine bacteriophage PhCF1.2. The presence of these IS elements may be relevant to genome plasticity, but their potential role in phage resistance or susceptibility is unclear. Further experimental validation is required to test the functional contribution of individual IS elements to phage-host interactions.

Specialised genomic features

Secondary metabolite biosynthetic gene clusters (BGCs)

Genome mining using antiSMASH 8.0 identified 10 biosynthetic gene clusters (BGCs) in *P. ganghwense* CF1.2 distributed across the two chromosomes. Chromosome 1 (Scaffold_1) contained gene clusters encoding a ribosomally synthesised and post-translationally modified peptide (RiPP-like, Region 1), a betalactone cluster (Region 2), a non-ribosomal peptide synthetase (NRPS, Region 3), a terpene cluster (Region 4), and an azole-containing RiPP cluster (Region 5). Chromosome 2 (Scaffold_2) encoded an NRPS-like cluster (Region 1), an ectoine cluster (Region 2), a Type I Polyketide Synthase (T1PKS) cluster (Region 3), and an ectoine-arylpolyene hybrid cluster (Region 5) (Fig. 13). The presence of two ectoine-associated BGCs suggests a substantial genetic capacity for osmotic stress tolerance, consistent with adaptation to fluctuating salinity in marine environments. The predicted arylpolyene cluster in Region 5 is likely to be involved in protection from photooxidative stress *via* the production of membrane associated pigments that scavenge reactive oxygen species, a role that has been widely reported in marine Gram negative bacteria (Cimermancic *et al*., 2014; Medema *et al*., 2021). The detection of NRPS, T1PKS and RiPP gene clusters also indicates the potential to produce structurally diverse secondary metabolites possibly involved in microbial competition, signalling and ecological interactions. Interestingly, a full biosynthetic cluster of hydrogen cyanide (HCN) (hcnABC) was found on Chromosome 2 (Scaffold_2, Region 4). The hcnABC operon is involved in the synthesis of hydrogen cyanide (HCN), a volatile secondary metabolite that has been shown to have antibacterial, antifungal and anti-predator properties and may increase competitiveness in complex microbial communities (Blumer and Haas, 2000). However, the identification of a complete hcnABC cluster suggests a possible genetic predisposition for HCN biosynthesis, but does not prove HCN production under the tested conditions. Thus, before any possible biotechnological or industrial application, the predicted biosynthetic potential of strain CF1.2 should be experimentally validated and its possible biosafety issues should be assessed.

Comparative genomic analyses show that BGCs are highly dynamic components of marine bacterial genomes and are frequently disseminated by horizontal gene transfer. Many of the secondary metabolite pathways identified in *Vibrio* and *Photobacterium* spp. are

Select genomic region:

Overview | 1.1 | 1.2 | 1.3 | 1.4 | 1.5 | 2.1 | 2.2 | 2.3 | 2.4 | 2.5

Identified secondary metabolite regions using strictness 'relaxed'

Region	Type	From	To	Similarity Confidence	Most similar known cluster	
Region 1.1	RiPP-like	117,329	128,201			
Region 1.2	betalactone	379,839	407,081			
Region 1.3	NRPS	699,295	749,993	High	amonabactin P 750	NRPS:Type I
Region 1.4	terpene-precursor	827,521	848,432			
Region 1.5	azole-containing-RiPP	2,751,043	2,780,219			
Region 2.1	NRPS-like	42,924	85,833			
Region 2.2	ectoine	267,743	278,129			
Region 2.3	T1PKS	972,084	1,014,981	Low	chromobactin	NRPS:Type I
Region 2.4	hydrogen-cyanide	1,202,377	1,215,238			
Region 2.5	ectoine , arylpolyene	1,635,565	1,679,235	Low	APE Vf	other:other

Fig. 13. Distribution of BGCs in *P. ganghwense* CF1.2 predicted by antiSMASH 8.0. The figure shows the position and organisation of BGCs in the genome, including RiPP-like, betalactone, NRPS, terpene, azole containing RiPP, NRPS-like, ectoine, Type I PKS, hydrogen cyanide (HCN) and ectoine-arylpolyene hybrid clusters

shared with distantly related marine bacteria, reflecting extensive exchange of biosynthetic potential within the marine microbiome (Cimermancic *et al*., 2014). For example, the pddGI-1 genomic island and plasmid pPHDP70 in *Photobacterium damselae* contain siderophore biosynthetic gene clusters that were likely acquired by horizontal gene transfer, highlighting the significant contribution of mobile genetic elements to secondary metabolism and ecological adaptation (Osorio *et al.*, 2015; Terceti *et al.*, 2018). The diverse BGC complement of strain CF1.2 also suggests an adaptive genome with the ability to produce multiple specialised metabolites that may potentially contribute to environmental fitness in marine ecosystems.

Bioremediation associated genes

Genome annotation revealed several genes associated with predicted metabolism of organic pollutants and resistance to environmental contaminants in *P. ganghwense* CF1.2, indicating potential relevance to bioremediation. The genome encodes carboxylesterase 2, acetyl esterases, and an arylesterase, enzymes potentially involved in the hydrolysis of ester bonds present in pyrethroid and organophosphate pesticides (Fig. 14). Furthermore, the presence of atrazine chlorohydrolase (AtzA) and N-isopropylammelide isopropylamidohydrolase (AtzC) suggests the genetic potential for

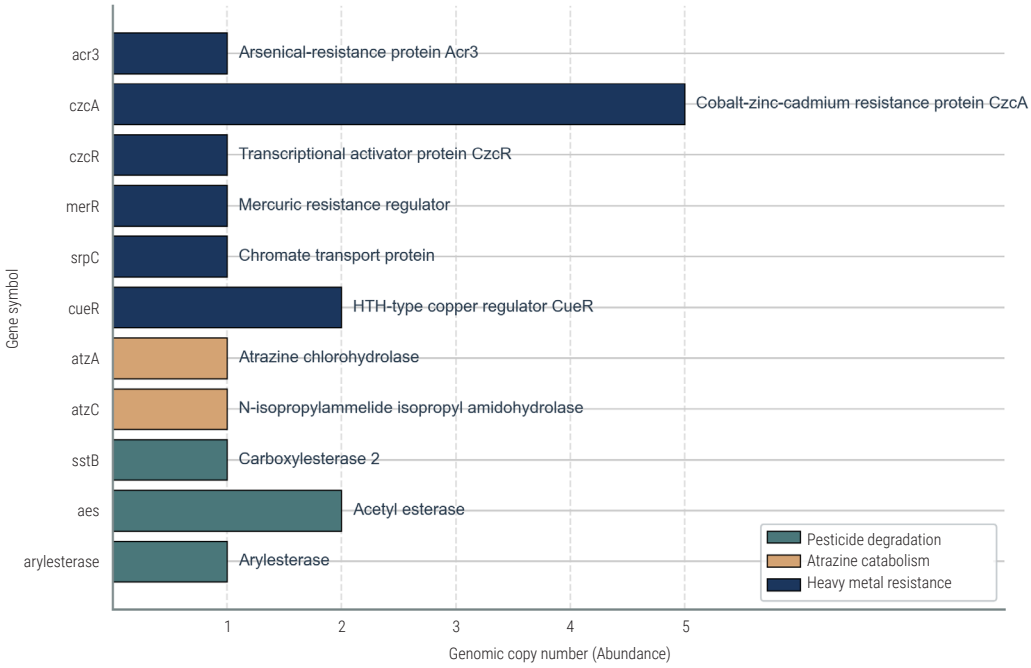

Fig. 14. Functional classification of bioremediation associated genes identified in *P. ganghwense* CF1.2. Genes are grouped according to their predicted roles in pesticide degradation, atrazine catabolism and heavy metal resistance

the initial steps of degradation of the herbicide atrazine through its catabolism (Fan and Song, 2014; Bhatt *et al*., 2021). The genome also harbours genes associated with heavy metal resistance, which may contribute to tolerance or resistance to heavy metal stress in metal polluted marine environments.

The results are consistent with previous studies on the metabolic versatility and bioremediation potential of *P. ganghwense*. *P. ganghwense* strain T14, isolated from a typical reef coral ecosystem, was experimentally confirmed to be able to degrade the pyrethroid pesticide cyfluthrin, emphasising the potential of this species to transform synthetic pesticides in marine environments (Jayasree *et al*., 2020). Comparative metabolomic analyses also showed that the degradation of cyfluthrin by *P. ganghwense* was coupled with the coordinated changes in central carbon metabolism, amino acid metabolism and stress response pathways, providing mechanistic insights into pesticide biodegradation (Wang *et al*., 2019). In addition to these observations, the identification of several pesticide degrading enzymes (for example, esterases and atrazine catabolic enzymes) in strain CF1.2 indicates a broader genetic potential for the transformation of different xenobiotic compounds. These results suggest that *P. ganghwense* may be involved in the natural attenuation of pesticide contaminants in marine ecosystems. However, experimental studies are required to confirm the biodegradation efficiency and substrate specificity of the predicted enzymes in *P. ganghwense* CF1.2.

Biosafety

Genome screening using AMRFinderPlus did not detect any acquired antimicrobial resistance (AMR) genes or human virulence factors in *P. ganghwense* CF1.2. In particular, no clinically relevant beta-lactamase, aminoglycoside, or colistin resistance genes, nor genes encoding major *Vibrio* associated human toxins, were identified. These results indicate that strain CF1.2 lacks the screened genomic determinants associated with antimicrobial resistance and the virulence factors evaluated. However, these genomic data alone do not demonstrate non-pathogenicity, as the pathogenic potential may be determined by factors that are not present in the current reference databases. Screening of genomes for AMR and virulence determinants is a relevant part of preliminary biosafety assessment for microorganisms intended for industrial biotechnology and environmental applications, as it can help identify potential risks associated with antimicrobial resistance and virulence associated traits (Feldgarden *et al*., 2021; EFSA Scientific Committee, 2025). Together with its predicted metabolic and bioremediation capabilities, the favourable genomic safety profile of strain CF1.2 supports its suitability for further evaluation as a candidate for industrial biotechnology applications, including biopolymer production and other marine biotechnological applications. Nevertheless, comprehensive biosafety evaluation should also include phenotypic assessment of antimicrobial susceptibility and pathogenicity before industrial deployment.

Here, we report the whole genome characterisation of the marine bacterium, *P. ganghwense* CF1.2, the specific host of the lytic marine bacteriophage, PhCF1.2. Comparative genomic analyses based on ANI and whole genome synteny confirmed its taxonomic placement in *P. ganghwense* and showed a highly conserved bipartite genome architecture with strain specific genomic variation. Functional annotation by COG, KEGG and CAZyme analyses showed that the bacterium is metabolically versatile, with a wide range of carbohydrate utilisation, nutrient metabolism, adaptation to the environment and response to stress. Genome mining also identified several BGCs, prophages, insertion sequence elements, and bioremediation related genes, thus demonstrating the strain's adaptive potential and genomic plasticity. The absence of acquired AMR genes and known virulence determinants supports its favourable genomic safety profile for future biotechnological applications. Overall, these results provide important insights into the genomic architecture, metabolic potential and evolutionary adaptation of *P. ganghwense* CF1.2 and constitute an important genomic resource for future studies of phage-host interactions, marine microbial ecology and sustainable marine biotechnological applications.

Acknowledgements

The research was supported by the Indian Council of Agricultural Research (ICAR), New Delhi, under the Institute Project No. MBEMD/MMD/39 of ICAR-CMFRI, Kochi. The authors gratefully acknowledge the present Director, ICAR-CMFRI, Kochi and Dr. A. Gopalakrishnan, former Director, ICAR-CMFRI, Kochi for providing all facilities for the research work. The first author gratefully acknowledges CSIR-UGC for the award of a Research Fellowship and CUSAT, Kochi, for granting Ph. D. admission.

References

Andrews, S. 2020. *FastQC: A quality control tool for high throughput sequence data*. https://www. bioinformatics.babraham.ac.uk/projects/fastqc/.

Astashyn, A., Tvedte, E. S., Sweeney, D., Sapojnikov, V., Bouk, N., Joukov, V., Mozes, E., Strope, P. K., Sylla, P. M., Wagner, L., Bidwell, S. L., Brown, L. C., Clar, K., Davis, E. W., Smith-White, B., Hlavina, W., Pruitt, K. M., Schneider, V. A. and Murphy, T. D. 2024. Rapid and sensitive detection of genome contamination at scale with FCS-GX. *Gen. Biol.,* 25: 60. https://doi.org/10.1186/s13059-024-03198-7.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A. and Pevzner, P. A. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.,* 19(5): 455–477. https://doi.org/10.1089/cmb.2012.0021.

Bhatt, P., Zhou, X., Huang, Y., Zhang, W. and Chen, S. 2021. Characterization of the role of esterases in the biodegradation of organophosphate, carbamate and pyrethroid pesticides. *J. Hazardous Mat.,* 411: 125026. https://doi.org/10.1016/j.jhazmat.2020.125026.

Blin K., Shaw S., Vader L., Szenei J., Reitz Z. L., Augustijn H. E., Cediel-Becerra J. D. D., de Crecy-Lagard V., Koetsier R. A., Williams S. E., Cruz-Morales P., Wongwas S., Segurado Luchsinger A. E., Biermann F., Korenskaia A., Zdouc M., Meijer D., Terlouw B. R., van der Hooft J. J. J., Ziemert N., Helfrich E. J. N., Masschelein J., Corre C., Chevrette M. G., van Wezel G. P., Medema M. H. and Weber T. 2025. antiSMASH 8.0: Extended gene cluster detection capabilities and analyses of chemistry, enzymology and regulation. *Nucleic Acids Res.,* 53(W1): W32–W38. https://doi.org/10.1093/nar/gkaf334.

Blumer, C. and Haas, D. 2000. Mechanism, regulation and ecological role of bacterial cyanide biosynthesis. *Arch. Microbiol.,* 173(3): 170–177. https://doi.org/10.1007/s002039900127.

Bobay, L. M., Touchon, M. and Rocha, E. P. C. 2014. Pervasive domestication of defective prophages by bacteria. *Proc. Natl. Acad. Sci. USA*, 111: 12127–12132. https://doi.org/10.1073 /pnas.1405336111.

Bosi, E., Donati, B., Galardini, M., Brunetti, S., Sagot, M. F., Lio, P., Crescenzi, P., Fani, R. and Fondi, M. 2015. MeDuSa: A multi-draft based scaffolder. *Bioinformatics*, 31(15): 2443–2451. https://doi.org/10.1093/bioinformatics/btv171.

Canchaya, C., Fournous, G. and Brussow, H. 2003. The impact of prophages on bacterial chromosomes. *Mol. Microbiol.,* 53: 9-18. https://doi.org/10.1111/j.1365-2958.2004.04113.

Chen, S., Dong, Y. H., Chang, C., Deng, Y., Zhang, X. F. and Zhong, G. 2021. Characterisation of a novel cyfluthrin-degrading bacterial strain *Brevibacterium aureum* and its biochemical degradation pathway. *Bioresour. Technol.,* 132: 16–23. https://doi.org/10.1016/j.biortech. 2013.01.002.

Chen S., Zhou Y., Chen Y. and Gu J. 2018. fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 34(17): i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

Cimermancic, P., Medema, M. H., Claesen, J., Kurita, K., Brown, L. C. W., Mavrommatis, K., Pati, A., Godfrey, P. A., Koehrsen, M., Clardy, J., Birren, B. W., Takano, E., Sali, A., Linington, R. G. and Fischbach, M. A. 2014. Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell,* 158(2): 412-421. https://doi.org/10.1016/ j.cell.2014.06.034.

De Palmenaer, D., Siguier, P. and Mahillon, J. 2008. IS4 family goes genomic. *BMC Evolution. Biol.,* 8: 18. https://doi.org/10.1186/1471-2148-8-18.

EFSA Scientific Committee 2025. Guidance on the characterisation and assessment of microorganisms used in the food chain. *EFSA J.,* 23(11). https://doi.org/10.2903/j.efsa.2025.9705.

Fan, X. and Song, F. 2014. Bioremediation of atrazine: Recent advances and promises, *J. Soils Sediments*, 14: 1727-1737. https://doi.org/10.1007/s11368-014-0921-5.

Feldgarden, M., Brover, V., Gonzalez-Escalona, N., Frye, J. G., Haendiges, J., Haft, D. H., Hoffmann, M., Pettengill, J. B., Prasad, A. B., Tillman, G. E., Tyson, G. H. and Klimke, W. 2021. AMRFinderPlus and the reference gene catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Sci. Rep.,* 11: 12728. https://doi.org/10.1038/s41598-021-91456-0.

Fortier, L. C. and Sekulovic, O. 2013. Importance of prophages to evolution and virulence of bacterial pathogens. *Virulence*, 4: 354–365. https://doi.org/10.4161/viru.24498.

Fuertes-Perez, S., Vogel, R. F. and Hilgarth, M. 2021. Comparative genomics of *Photobacterium* species from terrestrial and marine habitats. *Curr. Res. Microb. Sci.,* 2: 100087.

Galperin, M. Y., Makarova, K. S., Wolf, Y. I. and Koonin, E. V. 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res.,* 43(D1): D261–D269. https://doi.org/10.1093/nar/gku1223.

Gomez-Gil, B., Roque, A., Rotllant, G., Peinado, L., Romalde, J. L., Doce, A., Cabanillas-Beltrán, H., Chimetto, L. A. and Thompson, F. L. 2011. *Photobacterium swingsii* sp. nov., isolated from marine organisms. *IJSEM*, 61(Pt 2): 315–319. https://doi.org/10.1099/ijs.0.019687-0.

Gomez-Gil, B., Roque, A., Rotllant, G., Romalde, J. L., Doce, A., Eggermont, M. and Defoirdt, T. 2016. *Photobacterium sanguinicancri* sp. nov. isolated from marine animals. *Antonie van Leeuwenhoek*, 109(6): 817–825. https://doi.org/10.1007/s10482-016-0681-x.

Grant, J. R., Enns, E., Marinier, E., Mandal, A., Herman, E. K., Chen, C.-Y., Graham, M., Van Domselaar, G. and Stothard, P. 2023. Proksee: In-depth characterization and visualization of bacterial genomes. *Nucleic Acids Res.,* 51(W1): W484–W492. https://doi.org/10.1093/nar /gkad326.

Gurevich, A., Saveliev, V., Vyahhi, N. and Tesler, G. 2013. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics*, 29: 1072–1075. https://doi.org/10.1093/bioinformatics/ btt086.

Hossain, S., Shukri, Z. N. A., Waiho, K., Ibrahim, Y. S., Kamaruzzan, A. S., Rahim, A. I. A., Draman, A. S., Wahab, W., Khatoon, H. and Kasan, N. A. 2024. Biodegradation of polyethylene (PE), polypropylene (PP), and polystyrene (PS) microplastics by floc-forming bacteria, *Bacillus cereus* strain SHBF2, isolated from a commercial aquafarm. *Environ. Sci. Pollut. Res.,* 31(22): 32225–32245. https://doi.org/10.1007/s11356-024-33337-3.

ISO 2017. Microbiology of the food chain: Horizontal method for the determination of *Vibrio* spp., Part 1. Detection of potentially enteropathogenic *Vibrio parahaemolyticus, Vibrio cholerae* and *Vibrio vulnificus*. *ISO 21872-1:2017*, 1st edn. International Organization for Standardization, Geneva, Switzerland, 33 p.

Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T. and Aluru, S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Commun.,* 9: 5114. https://doi.org/10.1038/s41467-018-07641-9.

Jayasree, V., Sobhana, K. S., Poulose, P., Babu, K. R., Jasmine, S., Ranjith, L., Saravanan, R., Kingsly, H. J., Sreenath, K. R., Joshi, K. K., George, R. M. and Gopalakrishnan, A. 2020. Biodegradation of the pyrethroid pesticide cyfluthrin by the halophilic bacterium *Photobacterium ganghwense* isolated from coral reef ecosystem. *Indian J. Fish.,* 67. https://doi.org/10.21077/ ijf.2020.67.4.100553-14.

Kanehisa, M., Furumichi, M., Sato, Y., Matsuura, Y. and Ishiguro-Watanabe, M. 2025. KEGG: Biological systems database as a model of the real world. *Nucleic Acids Res.,* 53(D1): D672–D677. https://doi.org/10.1093/nar/gkae909.

Kim, Y.-O., Kim, K.-K., Park, S., Kang, S.-J., Lee, J.-H., Lee, S.-J., Oh, T.-K. and Yoon, J.-H. 2010. *Photobacterium gaetbulicola* sp. nov., a lipolytic bacterium isolated from a tidal flat sediment. *IJSEM*, 60: 2587–2591. https://doi.org/10.1099/ijs.0.016923-0.

Labella, A. M., Arahal, D. R., Castro, D., Lemos, M. L. and Borrego, J. J. 2017. Revisiting the genus Photobacterium: taxonomy, ecology and pathogenesis. *International microbiology: The official journal of the Spanish Society for Microbiology*, 20(1): 1–10. https://doi.org/10.2436/20.1501.01.280.

Labella, A. M. and Borrego, J. J. 2023. *Photobacterium*. In: Trujillo, M. E., Dedysh, S., DeVos, P., Hedlund, B., Kampfer, P., Rainey, F. A. and Whitman, W. B. (Eds.). *Bergey's manual of systematics of Archaea and Bacteria*. Wiley. https://doi.org/10.1002/9781118960608. gbm01076.pub2.

Labella, A. M., Castro, M. D., Manchado, M. and Borrego, J. J. 2018. Description of new and amended clades of the genus *Photobacterium*. *Microorganisms*, 6: 24. https://doi.org/10.3390/ microorganisms6010024.

Lascu, I., Mereuță, I., Chiciudean, I., Hansen, H., Avramescu, S. M., Tănase, A. M. and Stoica, I. 2021. Complete genome sequence of *Photobacterium ganghwense* C2.2: A new polyhydroxyalkanoate production candidate. *Microbiology Open*, 10: e1182. https://doi.org/10. 1002/mbo3.1182.

Lau, N., Heng, W. L., Miswan, N., Azami, N. A. and Furusawa, G. 2022. Comparative genomic analyses of the genus *Photobacterium* illuminate biosynthetic gene clusters associated with antagonism. *Int. J. Mol. Sci.,* 23: 9712. https://doi.org/10.3390/ijms23179712.

Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27(21): 2987–2993. https://doi.org/10.1093/bioinformatics/btr509.

Li, H. 2018. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*, 34(18): 3094–3100. https://doi.org/10.1093/bioinformatics/bty191.

Li, H. and Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14): 1754–1760. https://doi.org/10.1093/ bioinformatics/btp324.

Machado, H. and Gram, L. 2017. Comparative genomics reveals high genomic diversity in the genus *Photobacterium*. *Front. Microbiol.,* 8: 1204. https://doi.org/10.3389/fmicb.2017.01204.

Medema, M. H., Kottmann, R., Yilmaz, P., Cummings, M., Biggins, J. B., Blin, K., de Bruijn, I., Chooi, Y. H., Claesen, J., Coates, R. C., Cruz-Morales, P., Duddela, S., Düsterhus, S., Edwards, D. J., Fewer, D. P., Garg, N., Geiger, C., Gomez-Escribano, J. P., Greule, A., Hadjithomas, M. and Glöckner, F. O. 2021. Minimum information about a biosynthetic gene cluster. *Nat. Chem. Biol.,* 11: 625–631. https://doi.org/10.1038/nchembio.1890.

Merlin, S. T. P. and Umar, M. 2022. Genomic insights into symbiosis and host adaptation of an ascidian-associated bacterium *Bacillus aryabhattai* MCCB 387. *Symbiosis*, 86: 313-326. https://doi.org/ 10.1007/s13199-022-00839-6.

Mishra, S., Lin, Z., Pang, S., Zhang, W., Bhatt, P. and Chen, S. 2021. Recent advanced technologies for the characterization of xenobiotic-degrading microorganisms and microbial communities. *Front. Bioeng. Biotechnol.,* 9: 632059. https://doi.org/10.3389/fbioe.2021.632059.

Nourdin-Galindo, G., Sanchez, P., Molina, C. F., Espinoza-Rojas, D. A., Oliver, C., Ruiz, P., Vargas-Chacoff, L., Cárcamo, J. G., Figueroa, J. E., Mancilla, M., Maracaja-Coutinho, V. and Yañez, A. J. 2017. Comparative pan-genome analysis of *Piscirickettsia salmonis* reveals genomic divergences within genogroups. *Front. Cell. Infect. Microbiol.,* 7: Article 459. https://doi.org/10.3389/fcimb.2017.00459.

Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H. and Kanehisa, M. 1999. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.,* 27(1): 29–34. https://doi.org/10. 1093/nar/27.1.29.

Osorio, C. R., Rivas, A. J., Balado, M., Fuentes-Monteverde, J. C., Rodríguez, J., Jimenez, C., Lemos, M. L. and Waldor, M. K. 2015. A transmissible plasmid-borne pathogenicity island confers piscibactin biosynthesis in the fish pathogen *Photobacterium damselae* subsp. *piscicida*. *Appl. Environ. Microbiol.,* 81. https://doi.org/10.1128/AEM.01580-15.

Park, Y., Baik, K. S., Seong, C. N., Bae, K. S., Kim, S.-G. and Chun, J. 2006. *Photobacterium ganghwense* sp. nov., a halophilic bacterium isolated from seawater. *IJSEM*, 56: 745–749. https://doi.org/10.1099/ijs.0.63811-0.

Piel, D., Bruto, M., Labreuche, Y., Blanquart, F., Goudenege, D., Barcia-Cruz, R., Chenivesse, S., Le Panse, S., James, A., Dubert, J., Petton, B., Lieberman, E., Lieberman, E., Wegner, K. M., Hussain, F. A., Kauffman, K. M., Polz, M. F., Bikard, D., Gandon, S., Rocha, E. P. C. and Le Roux, F. 2022. Phage–host coevolution in natural populations. *Nat Microbiol.*, 7**:** 1075–1086. https://doi.org/10.1038/s41564-022-01157-1.

Richter, M. and Rossello-Mora, R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. USA,* 106: 19126–19131. https://doi.org/10.1073 /pnas.0906412106.

Rocha, E. P. C. 2008. The organization of the bacterial genome. *Ann. Rev. Gen.,* 42: 211–233. https://doi.org/10.1146/annurev.genet.42.110807.091653.

Schwengers, O., Jelonek, L., Dieckmann, M. A., Beyvers, S., Blom, J. and Goesmann, A. 2021. Bakta: Rapid and standardized annotation of bacterial genomes *via* alignment-free sequence identification. *Microb. Genom.,* 7: 000685. https://doi.org/10.1099/mgen.0.000685.

Seemann, T. 2014. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, 30: 2068–2069. https://doi.org/10.1093/bioinformatics/btu135.

Shimoyama, Y. 2022. *COG classifier: A tool for classifying prokaryote protein sequences into COG functional category*. GitHub.

Siguier, P., Perochon, J., Lestrade, L., Mahillon, J. and Chandler, M. 2006. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res.,* 34 (Database issue): D32–D36. https://doi.org/10.1093/nar/gkj014.

Tamura, K., Stecher, G. and Kumar, S. 2021. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.,* 38(7): 3022–3027. https://doi.org/10.1093/molbev/msab120.

Terceti, M. S., Vences, A., Matanza, X. M., Dalsgaard, I., Pedersen, K. and Osorio, C. R. 2018. Molecular epidemiology of *Photobacterium damselae* subsp. *damselae* outbreaks in marine rainbow trout farms reveals extensive horizontal gene transfer and high genetic diversity. *Front. Microbiol.,* 9: 2155. https://doi.org/10.3389/fmicb.2018.02155.

Touchon, M., Bernheim, A. and Rocha, E. P. C. 2017. Genetic and life-history traits associated with the distribution of prophages in bacteria. *ISME J.,* 11: 2744–2754. https://doi.org/10. 1038/ismej.2016.47.

Treangen, T. J. and Rocha, E. P. C. 2011. Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. *PLoS GENET.,* 7: e1001284. https://doi.org/10.1371/ journal.pgen.1001284.

Umar, M., Merlin, T. S. and Sajeevan, T. P. 2024. Genomic insights into symbiosis and host adaptation of sponge-associated novel bacterium, *Rossellomorea orangium* sp. nov. *FEMS Microbiol. Lett.,* 371, fnae074. https://doi.org/10.1093/femsle/fnae074.

Urbanczyk, H., Ast, J. C. and Dunlap, P. V. 2010. Phylogeny, genomics, and symbiosis of *Photobacterium*. *FEMS Microbiol. Rev.,* 35: 324–342. https://doi.org/10.1111/j.1574-6976.2010.00250.x.

Vezzi, A., Campanaro, S., D'Angelo, M., Simonato, F., Vitulo, N., Lauro, F. M., Cestaro, A., Malacrida, G., Simionati, B., Cannata, N., Romualdi, C., Bartlett, D. H. and Valle, G. 2005. Life at depth: *Photobacterium profundum* genome sequence and expression analysis. *Science*, 307: 1459–1461. https://doi.org/10.1126/science.1103341.

Wang, T., Hu, C. and Zhang, R. 2019. Mechanism study of cyfluthrin biodegradation by *Photobacterium ganghwense* with comparative metabolomics. *Appl. Microbiol. Biotechnol.*, 103: 473-488. https://doi.org/10.1007/s00253-018-9458-7.

Wishart, D. S., Han, S., Saha, S., Oler, E., Peters, H., Grant, J. R., Stothard, P. and Gautam, V. 2023. PHASTEST: faster than PHASTER, better than PHAST. *Nucleic Acids Res.,* 51(W1): W443–W450. https://doi.org/10.1093/nar/gkad382.

Zheng, J., Ge, Q., Yan, Y., Zhang, X., Huang, L. and Yin, Y. 2023. dbCAN3: Automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Res.*, 51(W1): W115–W121. https://doi.org/10.1093/nar/gkad328.