

# Genomic basis of Baikal sponge tissue degradation caused by a free-living *Flavobacterium* with host-glycan-targeting polysaccharide utilization loci

Panova Yu.A.<sup>1,2</sup>, Belikov S.I.<sup>1\*</sup>, Chernogor L.I.<sup>1</sup>

<sup>1</sup>Limnological Institute, Siberian Branch of the Russian Academy of Sciences, Ulan-Batorskaya Str., 3, Irkutsk, 664033, Russia

<sup>2</sup>Siberian Federal University, Svobodny Prospect, 79, Krasnoyarsk, 660041, Russia

**ABSTRACT.** Here we present the complete genome sequence of the *Flavobacterium* sp. strain PLB03 isolated from the primmorphs cell culture of the *Lubomirskia baikalensis* sponge and compare it with pathogenic free-living members of the genus *Flavobacterium*. Strain *Flavobacterium* sp. PLB03 exhibits pronounced lytic activity towards sponge cells and causes death of symbiotic microalgae (the Trebouxiophyceae). The *Flavobacterium* sp. PLB03 genome has 5,925,828 bp and a GC content of 34.3%. An average nucleotide identity (ANI) of 98.9% revealed that strain *Flavobacterium* sp. PLB03 has the highest similarity to *Flavobacterium* CSZ. Comparative genomic analysis of the strain and phylogenetically related *Flavobacterium* indicated that *Flavobacterium* sp. PLB03 has a large genome size characteristic of the environmental genus *Flavobacterium*, and the peptidase (PEP) to glycoside hydrolase (GH) gene ratio is 1.8, suggesting that strain PLB03 exhibits characteristics similar to nonpathogenic strains.

**Keywords:** *Flavobacterium* sp. PLB03, freshwater sponges, CAZymes, PULs, Type IX secretion system, genomic islands, opportunistic pathogenicity

**For citation:** Panova Yu.A., Belikov S.I., Chernogor L.I. Genomic basis of Baikal sponge tissue degradation caused by a free-living *Flavobacterium* with host-glycan-targeting polysaccharide utilization loci // Limnology and Freshwater Biology. 2025. - № 6. - P. 1313-1328. DOI: [10.31951/2658-3518-2025-A-6-1313](https://doi.org/10.31951/2658-3518-2025-A-6-1313)

## 1. Introduction

Lake Baikal is one of the oldest lakes in geological history, containing nearly 20% of the world's freshwater reserves. The lake is distinguished by its diversity of endemic flora and fauna, and Baikal sponges are most abundant organisms in the coastal zone, with a biomass of over 700 g/m<sup>2</sup> (Kozhov, 1972; 2013; Pile et al., 1997; Semiturkina et al., 2009).

A disease affecting freshwater sponges was first discovered in 2011, when pink specimens were found in the central basin of Lake Baikal (Bormotov, 2012). The death of the sponge primary symbiont, the green alga *Choricystis* sp., and its replacement by the *Synechococcus* sp. cyanobacterium caused discoloration of the sponges (Denikina et al., 2016; Belikov et al., 2019). In subsequent years, the external signs of the disease changed, and from 2013 to the present, sponges have shown various symptoms of damage, such as discoloration, tissue necrosis, and the formation of brown and dirty-purple bacterial biofilms. Sponge numbers have declined significantly, and diseased sponges are

now found throughout Lake Baikal. Multiple changes in the littoral ecosystem of Lake Baikal accompany disease and death of sponges (Timoshkin et al., 2016; Khanaev et al., 2018).

Previously, we demonstrated that sponge disease is associated with changes and dramatic shifts in the composition of the microbial community of Baikal sponges. We found that the observed imbalance in the microbial communities of diseased sponges is caused by various microorganisms that, acting in concert, intensify their negative impact, leading to the death of Baikal sponges (Belikov et al., 2019).

We conducted an experimental infection of the cell culture of primmorphs with the bacterial suspension of the diseased sponge to identify microorganisms pathogenic to sponges. A significant increase in the relative content of bacteria of the *Oxalobacteraceae* family, up to 25.65%, and the *Flavobacteriaceae* family, up to 62.04%, was detected along with a simultaneous decrease in the number of other microorganisms (Chernogor et al., 2020a; Chernogor et al., 2020b). Based on the genome analysis of the isolated

\*Corresponding author. E-mail address: [sergeibelikov47@gmail.com](mailto:sergeibelikov47@gmail.com) (S.I. Belikov)

**Received:** November 20, 2025;

**Accepted after revised:** December 09, 2025;

**Available online:** December 25, 2025

© Author(s) 2025. This work is distributed under the Creative Commons Attribution-NonCommercial 4.0 International License.



*Janthinobacterium* SLB01 and *Flavobacterium* SLB02 strains, we hypothesized the possible cooperative interaction of these two bacteria in the development of sponge disease (Petrushin et al., 2020; Belikov et al., 2021).

The *Flavobacterium* is known to be ubiquitous. For example, type strains of *Flavobacterium* spp. were isolated from various sources, including freshwater, the Antarctic Sea ice, the rhizosphere, and diseased fish (Yi et al., 2005; Ali et al., 2009; Chen et al., 2012; Ekwe et al., 2017). The number of officially described *Flavobacterium* species is rapidly increasing and includes 448 species as of October 2025 (<https://lpsn.dsmz.de/genus/Flavobacterium>).

For the first time, the strain *Flavobacterium* sp. PLB03 was isolated from cell culture of the primorphs of diseased sponge *L. baikalensis*. The role of these bacteria and their link with disease development and mortality of sponges remain poorly understood. In this study, we present the assembled and annotated genome of the strain *Flavobacterium* sp. PLB03 and compare with the genomes of fish-pathogenic and non-pathogenic *Flavobacterium*. The genome sequence of the strain *Flavobacterium* sp. PLB03 and its analysis are expected to provide more comprehensive information on disease-associated factors. Identifying these genomic features will help understand the causes of the disease distribution and mortality of Baikal sponges.

## 2. Materials and Methods

### 2.1. Genome sequencing, quality control, and assembly

Raw Illumina sequencing reads were evaluated using FastQC v0.11.9 (Andrews, 2010) to assess base quality, GC content, and adapter contamination. Low-quality bases and adapter sequences were trimmed using Trimmomatic v0.39 (Bolger et al., 2014). The genome was assembled with SPAdes v3.15.4 (Bankevich et al., 2012; Prjibelski et al., 2020) using default k-mer iterations. Assembly quality and contamination were assessed using FCS-GX (Astashyn et al., 2024) to detect and remove putative foreign contigs. Contigs shorter than 200 bp or flagged as contaminants were discarded, yielding the final assembly used for downstream analyses.

### 2.2. Genome annotation

The final assembly was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v6.9 (Tatusova et al., 2016), which integrates *ab initio* gene prediction, HMM-based protein family classification, and curated functional models from RefSeq. Annotation included coding sequences, non-coding RNAs, tRNAs, and rRNA fragments, and pseudogenes.

### 2.3. Genomes Included in the Analysis

To place *Flavobacterium* sp. PLB03 into a phylogenomic and functional context, we selected 14 publicly available genomes representing both free-living and fish-pathogenic lineages of the genus *Flavobacterium*.

Genomes were retrieved from NCBI GenBank between March and May 2025. The final comparative dataset included the following species: *F. sp.* CSZ, *F. piscisymbiosum* F-30, *F. piscis* CCUG 60099, *F. bizetiae* HJ-32-4, *F. bizetiae* FlaQc-53, *F. bizetiae* CIP 105534, *F. aquidurens* DSM 18293, *F. johnsoniae* FJOH, *F. pectinovorum* ZE23VCel01, *F. soyae* SCIV07, *F. collinsii* FlaQc-26, *F. lipolyticum* F-126, *F. chilense* DSM 24724, and *F. psychroterrae* CCM 8827, as well as the fish-pathogenic species *F. columnare* 04017018, *F. psychrophilum* JIP02-86, *F. branchiophilum* FL-15, and *F. tractae* ATCC BAA-2541. Detailed assembly statistics and genome annotation features for all strains are provided in Supplementary Table S1.

### 2.4. Phylogeny and ANI analysis

An exploratory genome-based phylogenetic analysis was conducted using the TYGS platform (Meier-Kolthoff and Göker, 2019; Meier-Kolthoff et al., 2022) to identify the closest taxa and to select appropriate reference genomes for downstream comparative analyses. Because the genome assembly of *Flavobacterium* sp. PLB03 is highly fragmented, and all recovered 16S rRNA gene sequences were incomplete, we used the housekeeping gene *gyrB* for phylogenetic reconstruction. This approach follows previous recommendations for *Flavobacterium* taxonomy, where *gyrB* provides higher phylogenetic resolution and is preferred when 16S rRNA is unavailable or incomplete (Peeters and Willems, 2011).

Whole-genome Average Nucleotide Identity (ANI) was calculated using FastANI (Jain et al., 2018). For this analysis, we included the following genomes representing closely related members of the genus *Flavobacterium*: *F. sp.* CSZ, *F. piscisymbiosum* strain F-30, *F. piscis* CCUG 60099, *F. bizetiae* strain HJ-32-4, *F. sp.* KBS0721, *F. bizetiae* strain FlaQc-53, and *F. bizetiae* strain CIP 105534. These genomes were selected based on TYGS clustering results and on their proximity to *Flavobacterium* sp. PLB03 in preliminary analyses.

### 2.5. Detection of antimicrobial resistance and virulence genes

Potential antimicrobial resistance (AMR) genes and virulence factor candidates were screened using ABRicate v1.0.1 (Seemann, 2016) against the CARD, NCBI AMR, and VFDB (Chen et al., 2016; Liu et al., 2022) databases using Galaxy platform 2024.

### 2.6. Identification of secretion systems

Type IX secretion system (T9SS) components were identified with manual inspection of annotated genes and BLASTp search.

### 2.7. Genomic island prediction

Genomic islands (GIs) were predicted using IslandViewer 4 (Bertelli et al., 2017), which integrates the SIGI-HMM and IslandPath-DIMOB methods. Genes within GIs were annotated using PGAP. Functional cat-

egories including mobile elements, secretion components, CAZymes, TonB-dependent receptors and proteases were recorded.

2.8. Protease annotation and MEROPS comparative analysis

Proteases were annotated by BLASTp searches against the MEROPS (Rawlings et al., 2014; 2018; Rawlings and Bateman, 2021) database (E-value ≤ 1e-10). Family-level counts were summarized for *Flavobacterium* sp. PLB03 and compared with reference species, particularly *Flavobacterium johnsoniae*. Proteases were classified by catalytic type (serine, metalloprotease, cysteine, and aspartic) and by predicted subcellular localization.

2.9. Identification of CAZymes and polysaccharide utilization loci (PULs) and analysis across *Flavobacterium* species

Carbohydrate-active enzymes (CAZymes) were annotated using dbCAN3 (Zheng et al., 2023). Only enzymes supported by at least two methods were retained. Polysaccharide utilization loci (PULs) were predicted using the dbCAN CGC pipeline, which detects co-localized CAZymes, SusC/D-like TonB-dependent transporters and transcriptional regulators. PUL substrates were inferred based on CAZy family composition, dbCAN substrate prediction, and literature on known PUL architectures in *Flavobacteria* and other *Bacteroidetes*.

To contextualize *Flavobacterium* sp. PLB03 within the genus, we first compared CAZyme repertoires across 15 *Flavobacterium* genomes, including environmental representatives and fish-pathogenic species. This broad-scale analysis revealed major patterns of CAZyme diversity and showed that *Flavobacterium* sp. PLB03 possesses one of the most expanded CAZyme repertoires, particularly in host-glycan-associated families.

Because PUL architectures differ greatly between distant species and lack locus-level homology across the genus, detailed PUL comparison was restricted to six representative genomes: *Flavobacterium* sp. PLB03, *F. johnsoniae*, *F. piscisymbiosum*, *F. bizetiae*, *F. tractae*, and *F. columnare*.

Comparative analyses of CAZyme profiles, PUL repertoires and MEROPS family distributions were performed across 15 *Flavobacterium* genomes. Data matrices were constructed from family-level counts and analyzed by hierarchical clustering using Ward’s method and Euclidean distances in R.

3. Results

3.1. Genome assembly and annotation results

The draft genome assembly comprised 334 contigs with a total length of 6,003,505 base pairs (bp). After removing short contigs (< 200 bp) and contaminants, the final assembly contained 93 contigs span-

ning a total of 5,925,828 bp, the N50 value of 542,557 bp, and a GC content of 34.3%. This genome assembly was submitted to GenBank under accession number JBLIYN000000000 and prepared for downstream annotation and analysis with PGAP. Key assembly and annotation metrics are summarized in Table 1.

The genome size is larger than that of fish-pathogenic *Flavobacterium* species and comparable to environmental strains with expanded metabolic repertoires (Supplementary Table S1).

3.2. Phylogeny and ANI analysis

An exploratory analysis using the TYGS service and an analysis of BLAST results in NCBI indicated that the sequences of *Flavobacterium* sp. PLB03 are most closely related to the sequences of unspecified species *Flavobacterium* sp. CSZ, *Flavobacterium* sp. KBS0721 and *F. bizetiae*, *F. piscisymbiosum*, and *F. piscis*. ANI-based genomic comparisons with these genomes revealed that *Flavobacterium* sp. PLB03 forms a tight species-level cluster only with *Flavobacterium* sp. CSZ, with ANI of 98.96%, clearly above the 95% species threshold. All other analyzed *Flavobacterium* genomes showed ANI values of 89.3–90.8% relative to *Flavobacterium* sp. PLB03, demonstrating interspecies-level divergence. The closest species outside the *Flavobacterium* sp. PLB03-CSZ pair was *F. piscisymbiosum* F-30 (90.8% ANI), a known fish-associated taxon, although still far below the species boundary (Table S2 Supplementary). Together, these results indicate that *Flavobacterium* sp. PLB03 and *F. sp.* CSZ form a distinct species-level lineage.

Phylogenetic analysis of *gyrB* placed *Flavobacterium* sp. PLB03 within an environmental lineage of *Flavobacterium*, which was most closely related to *Flavobacterium* sp. CSZ and at a significant distance from *F. psychrophilum*, *F. columnare*, and *F. branchiophilum*, being fish pathogens (Fig. 1).

Table 1. Raw reads and genomic features statistics of *Flavobacterium* sp. PLB03

| Property                    | <i>Flavobacterium</i> sp. PLB03 |
|-----------------------------|---------------------------------|
| Raw reads                   | 7,220,904                       |
| Assembly length, bp         | 5,925,828                       |
| Number of contigs           | 93                              |
| Longest contig’s length, bp | 940,434                         |
| N50, bp                     | 542,557                         |
| N90, bp                     | 198,207                         |
| GC content                  | 34.3%                           |
| Total number of genes       | 5,104                           |
| Protein-coding sequences    | 4,975                           |
| tRNAs                       | 64                              |
| Noncoding RNAs              | 3                               |
| Pseudogenes                 | 32                              |
| GenBank accession number    | JBLIYN000000000                 |

### 3.3. Detection of antimicrobial resistance and virulence genes

At standard stringency thresholds ( $\geq 75$ –90% identity and coverage), no virulence factors were detected by Abricate in the genome of *Flavobacterium* sp. PLB03. When the thresholds were relaxed to 60% identity and 60% coverage, Abricate reported four hits showing partial similarity to virulence-associated proteins described in other bacterial taxa: *katA*, *htpB* (GroEL/Hsp60), *neuC1* (UDP-N-acetylglucosamine 2-epimerase), and one protein annotated in VFDB as a Dot/Icm type IV secretion system effector (*lpg2370*).

All matches displayed low amino acid identity (66–68%) and incomplete alignment coverage (70–90%), indicating distant homology rather than functional equivalence. These genes encode core house-keeping functions widely conserved in bacteria and are not considered bona fide virulence determinants in the genus *Flavobacterium*. Thus, *Flavobacterium* sp. PLB03 does not encode conventional virulence systems described in fish pathogens such as *F. psychrophilum* and *F. columnare*.

### 3.4. Genomic islands

IslandViewer predicted 23 genomic islands (GIs), ranging from 4.1 to 74.9 kbp and totaling approximately 540 kbp (10–12% of the genome). GI-associated genes included hypothetical proteins, mobile elements, TonB-dependent receptors, and selected Type IX Secretion System components. Genomic islands were dispersed across the genome rather than forming large clusters (Fig. 2). Most GIs were predicted by at least two independent algorithms, indicating high confidence in their boundaries.

The predicted genomic islands were enriched in hypothetical proteins (32%,  $n = 129$ ), along with transposases and integrases (4%), membrane-associated proteins (4%), and regulatory elements such as helix–turn–helix (HTH) transcriptional regulators (4%). Several islands also contained components typical of *Flavobacterium*, including Type IX secretion system-associated genes, TonB-dependent receptors, and *SusC/RagA* TonB-linked outer membrane proteins ( $n = 4$ ), and IS6-family transposases ( $n = 4$ ).

### 3.5. Carbohydrate-active enzymes (CAZymes)

The *Flavobacterium* sp. PLB03 genome encodes 392 CAZymes, including 185 glycoside hydrolases, 53 glycosyltransferases, 36 carbohydrate esterases, 14 polysaccharide lyases, 28 carbohydrate-binding modules, and 7 auxiliary activity enzymes. This places *Flavobacterium* sp. PLB03 among the most CAZyme-rich freshwater *Flavobacterium* species studied (Table 2). The CAZyme repertoire not only covers a wide range of plant and algal polysaccharides (xylan, pectin, xyloglucan, beta-glucan, mannan, alginate, and alpha-glucan) but also includes families typically associated with host glycans, such as GH29, GH33, GH92, GH95, GH105,

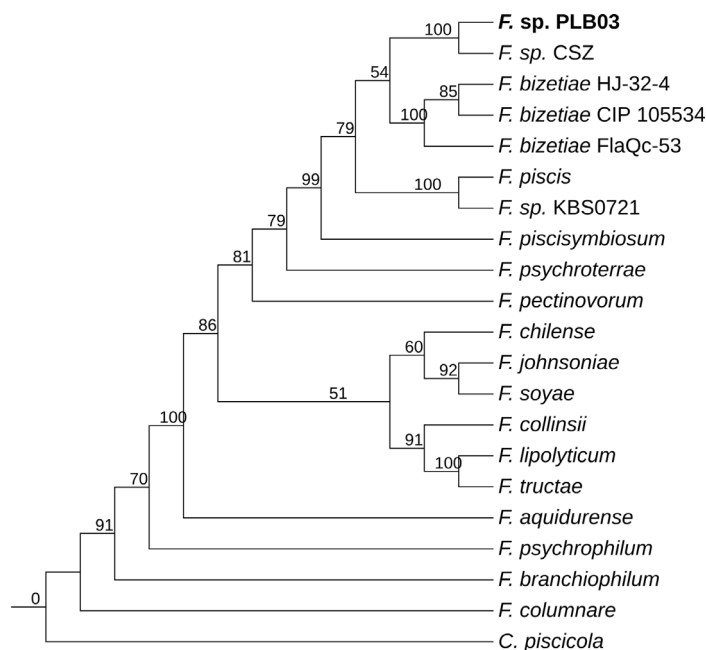


Fig.1. Neighbor-Joining tree based on complete *gyrB* sequences.

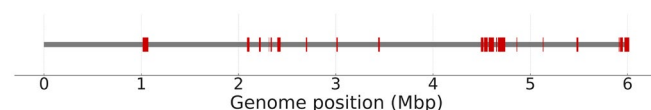


Fig.2. Genomic island distribution across the genome of the *Flavobacterium* sp. PLB03.

GH127, PL12, PL29, and CBM32. The latter are key to degradation of mucins, glycosaminoglycans, and eukaryotic glycoproteins.

### 3.6. Polysaccharide utilization loci

Using dbCAN3 and the CGC pipeline, we identified 32 polysaccharide utilization loci (PULs) in the *Flavobacterium* sp. PLB03 genome. Comparative analysis with five species (*F. johnsoniae*, *F. bizetiae*, *F. piscisymbiosum*, *F. tractae*, and *F. columnare*) revealed a total of 134 PULs (Supplementary Table S3). *Flavobacterium* sp. PLB03 is among the species with the highest number of PULs, comparable only to *F. piscisymbiosum* and *F. bizetiae*, whereas fish pathogens encode far fewer PULs (1–11).

Structure-based comparison of PUL architectures revealed at least 11 PULs unique to *Flavobacterium* sp. PLB03 across all evaluated genomes. The most notable and unique clusters target host-associated glycans, as shown in Table 3.

None of these architectures were identified in the other five species.

### 3.7. Protease analysis

MEROPS-based annotation of proteases and protease inhibitors revealed a broad enzymatic diversity across the analyzed *Flavobacterium* genomes. After normalization of MEROPS annotation to unified fam-

**Table 2.** Comparative distribution of CAZy classes across 15 *Flavobacterium* strains.

| Species                                 | CAZy-genes | GH  | GT | CE | PL | CBM | AA |
|-----------------------------------------|------------|-----|----|----|----|-----|----|
| <i>Flavobacterium</i> sp. PLB03         | 392        | 185 | 53 | 36 | 14 | 28  | 7  |
| <i>F. soyae</i> SCIV07                  | 408        | 183 | 66 | 32 | 10 | 36  | 6  |
| <i>F. piscisymbiosum</i> F-30           | 386        | 175 | 57 | 38 | 13 | 24  | 7  |
| <i>F. pectinovorum</i> ZE23VCel01       | 392        | 172 | 60 | 33 | 11 | 34  | 3  |
| <i>F. johnsoniae</i> FJOH               | 382        | 165 | 59 | 31 | 12 | 29  | 4  |
| <i>F. bizetiae</i> HJ-32-4              | 363        | 153 | 58 | 33 | 9  | 25  | 3  |
| <i>F. aquidurens</i> DSM 18293          | 343        | 149 | 39 | 42 | 14 | 18  | 5  |
| <i>F. chilense</i> DSM 24724            | 348        | 134 | 54 | 34 | 15 | 28  | 4  |
| <i>F. psychroterrae</i> CCM 8827        | 331        | 127 | 46 | 28 | 15 | 32  | 3  |
| <i>F. tractae</i> ATCC BAA-2541         | 276        | 55  | 57 | 25 | 7  | 7   | 6  |
| <i>F. lipolyticum</i> F-126             | 278        | 55  | 60 | 24 | 7  | 8   | 6  |
| <i>F. collinsii</i> FlaQc-26            | 270        | 59  | 52 | 21 | 9  | 7   | 5  |
| <i>F. branchiophilum</i> FL-15 (пат.)   | 180        | 29  | 51 | 14 | 2  | 7   | 2  |
| <i>F. columnare</i> 04017018 (пат.)     | 154        | 18  | 44 | 12 | 3  | 1   | 1  |
| <i>F. psychrophilum</i> JIP02-86 (пат.) | 137        | 10  | 45 | 9  | 0  | 0   | 1  |

**Note:** GH – glycoside hydrolases; GT – glycosyltransferases; CE – carbohydrate esterases; PL – polysaccharide lyases; CBM – carbohydrate-binding modules; AA – auxiliary activities.

ily codes, all species showed a conserved protease core dominated by the I39, S33, M14, S12, and S9 families (Supplementary Table S3). Among these, the inhibitor family I39 was the most abundant across nearly every genome. Despite this shared core, the species-by-family protease matrix showed marked differences in the degree of family expansion. Some species possessed compact repertoires, while others exhibited substantial enrichment of multiple S- and I-families. The *Flavobacterium* sp. PLB03 encoded high counts in several major protease families (Fig.3).

Further comparison of protease families within Cluster 2 revealed four MEROPS families present exclusively in the *Flavobacterium* sp. PLB03 but absent in all other members of its cluster: M10, S1, T2, and T3. All four families are present in species from Clusters 1 and 3 but absent from the other cluster-2 genomes.

4. Discussion

The *Flavobacterium* sp. PLB03 genome exhibits a combination of traits typical of free-living members of the genus together with modules commonly associated with host interaction. Phylogenetic placement and high

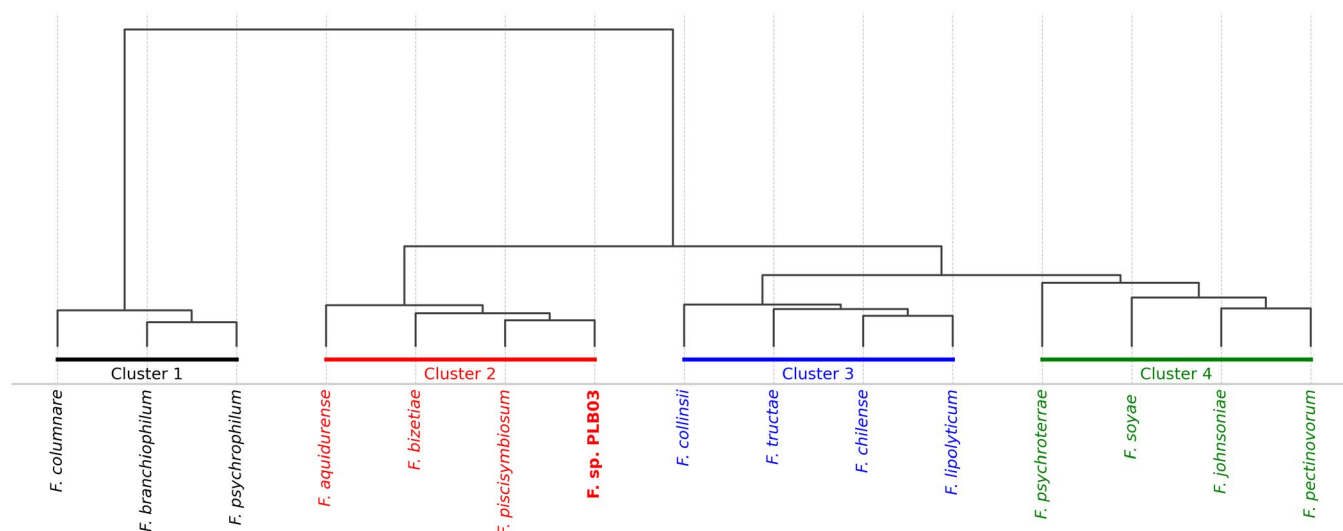
ANI values clearly demonstrate that *Flavobacterium* sp. PLB03 forms a distinct species-level lineage together with *Flavobacterium* sp. CSZ, while remaining distant from well-characterized fish pathogens such as *F. psychrophilum*, *F. columnare*, and *F. branchiophilum*. This positioning supports the view that *Flavobacterium* sp. PLB03 is primarily an environmental taxon rather than a specialized pathogen.

Despite the absence of classical virulence determinants, even with relaxed detection thresholds, the *Flavobacterium* sp. PLB03 genome has several functional features that may contribute to its ability to degrade sponge tissues. The genome is larger than the genomes of fish-pathogenic species and encodes an expanded metabolic repertoire. For instance, 392 CAZyme genes place *Flavobacterium* sp. PLB03 among the richest carbohydrate-degrading freshwater *Flavobacterium* species analyzed previously. Noteworthy is the presence of CAZy families typically linked to the degradation of eukaryotic glycoconjugates, including GH29, GH33, GH92, GH95, GH105, GH127, PL12, PL29, and CBM32. These enzymes enable the breakdown of mucins, glycosaminoglycans, and complex animal glycoproteins—key structural polysaccharides in sponge tissues.

**Table 3.** Host-glycan and algal polysaccharide-targeting PUL identified in *Flavobacterium* sp. PLB03

| PUL type                              | Key CAZy families             | Notes                                                              |
|---------------------------------------|-------------------------------|--------------------------------------------------------------------|
| Mucin-degrading PUL                   | GH105, GH88, PL12/PL8, CBM32  | Canonical mucin-targeting architecture                             |
| Glycosaminoglycan (GAG) PUL (×2)      | GH154, GH159, GH128, PL12/PL8 | Two separate GAG-specific loci                                     |
| Host glycan / animal glycoprotein PUL | GH29, GH33, GH92, GH130       | Large locus enriched in fucosidases, sialidases and α-mannosidases |
| Arabinogalactan / AGP-type PUL        | GH105, GH128, GH154, PL42     | Targets arabinogalactan and AGP-like substrates                    |
| Alginate-specific PUL                 | PL6, PL17, CE20               | Unique alginate-degrading locus absent from other species          |





**Fig.3.** Hierarchical clustering of *Flavobacterium* species based on MEROPS protease family profiles.

This metabolic potential is reflected in the architecture of the 32 identified PULs. Among them, there are at least five large *Flavobacterium* sp. PLB03 specific clusters targeting host-associated glycans: a canonical mucin-degrading PUL, two independent glycosaminoglycan (GAG) loci, a large host-glycan/animal glycoprotein PUL enriched in fucosidases, sialidases and  $\alpha$ -mannosidases, an AGP-type arabinogalactan PUL, and a unique alginate-degrading locus absent from all comparative genomes. Such a diversified PUL repertoire suggests a broad substrate range and is characteristic of “opportunistic” *Flavobacteria* capable of exploiting both environmental polysaccharides and those of eukaryotic origin.

The presence of 23 genomic islands, comprising 10–12% of the genome, provides additional evidence for adaptive versatility. These islands are enriched in hypothetical proteins, mobile elements, TonB-dependent receptors, components of the Type IX secretion system, and SusC/RagA-like transporters, indicating active genome modularity and potential horizontal gene acquisition. The dispersed distribution of GI-associated TonB–SusC modules aligns with the typical organization of PULs in *Flavobacterium*, further supporting the idea of substrate-driven genome plasticity.

Protease analysis revealed a broadened MEROPS repertoire in *Flavobacterium* sp. PLB03, including four families (M10, S1, T2, and T3) absent from other members of its cluster but present in more metabolically flexible species. These proteases may contribute to the degradation of proteinaceous components of sponge tissues. In *Flavobacterium*, proteases frequently act synergistically with CAZymes during the breakdown of complex organic matrices, suggesting a complementary role in substrate acquisition.

Taken together, the genomic features of *Flavobacterium* sp. PLB03 indicate that it is not a classical pathogen but rather an environmentally adapted bacterium with an expanded capacity for the degradation of complex organic substrates, including those of eukaryotic origin. The combination of a rich CAZyme and PUL repertoire, unique host glycan-targeting loci, diverse proteases, and modular genomic islands pro-

vides a mechanistic basis for the observed ability of *Flavobacterium* sp. PLB03 to degrade sponge tissues and green microalgae. These findings support the speculation that *Flavobacterium* sp. PLB03 represents an ecologically versatile organism capable of opportunistic interactions with host organisms.

## 5. Conclusions

The analyzed strain *Flavobacterium* sp. PLB03 is a new species belonging to a group of free-living (environmental) *Flavobacterium* with a large genome. ANI calculations with values of 89.4–90.8% that are significantly below the thresholds used to distinguish members of closely related species confirm that the *Flavobacterium* sp. PLB03 is a unique species. The strain *Flavobacterium* sp. PLB03 lacks typical virulence factors but possesses extensive metabolic capabilities for invasion and destruction of host tissue. *Flavobacterium* sp. PLB03 is a likely symbiont of sponges, as it has a unique set of specific enzymes that degrade host glucans, which is unusual for the genus *Flavobacterium*. An analysis of its ability to degrade polysaccharides revealed that the strain *Flavobacterium* sp. PLB03 has significant potential in this field. However, further research is needed to identify the specific polysaccharides that this strain degrades, which will contribute to our understanding of the ecological role and biotechnological applications of the *Flavobacterium* strains.

## Acknowledgements

This study is funded by the State Project of Limnological Institute Siberian Branch of Russian Academy of Sciences No. 0279-2021-0011. This material is based upon work supported by the Russian Science Foundation Project No. 19-14-00088 and the Russian Foundation for Basic Research No. 16-04-00065; 16-54-150007; 18-04-00224.

## Conflict of interests

The authors declare no conflicts of interest.

## References

- Ali Z., Cousin S., Fruhling A. et al. 2009. *Flavobacterium rivuli* sp. nov., *Flavobacterium sub-saxonicum* sp. nov., *Flavobacterium swingsii* sp. nov. and *Flavobacterium reichenbachii* sp. nov., isolated from hard water rivulet. International Journal of Systematic and Evolutionary Microbiology 59: 2610–2617.
- Andrews S. 2010. FastQC: A quality control tool for high throughput sequence data. Babraham Bioinformatics.
- Astashyn A., Tvedte E.S., Sweeney D. et al. 2024. Rapid and sensitive detection of genome contamination at scale with FCS-GX. Genome Biology 25: 1–25. DOI: [10.1186/S13059-024-03198-7](https://doi.org/10.1186/S13059-024-03198-7)/FIGURES/4
- Bankovich A., Nurk S., Antipov D. et al. 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. Journal of Computational Biology 19(5):455–477. DOI: [10.1089/cmb.2012.0021](https://doi.org/10.1089/cmb.2012.0021)
- Belikov S., Belkova N., Butina T. et al. 2019. Diversity and shifts of the bacterial community associated with Baikal sponge mass mortalities. PLOS One 14: 3. DOI: [10.1371/journal.pone.0213926](https://doi.org/10.1371/journal.pone.0213926)
- Belikov S.I., Petrushin I.S., Chernogor L.I. 2021. Genome analysis of the *Janthinobacterium* sp. strain SLB01 from the Diseased Sponge of the *Lubomirskia baikalensis*. Current Issues in Molecular Biology 43: 2220–2237. DOI: [10.3390/cimb43030156](https://doi.org/10.3390/cimb43030156)
- Bertelli C., Laird M.R., Williams K.P. et al. 2017. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Research 45: W30–W35. DOI: [10.1093/nar/gkx343](https://doi.org/10.1093/nar/gkx343)
- Bolger A.M., Lohse M., Usadel B. 2014. Trimmomatic: A flexible read trimming tool for Illumina NGS data. Bioinformatics 30: 2114–2120.
- Bormotov A.E. 2012. What has happened to Lake Baikal sponges? Science first hand 2: 32. URL: <https://scfh.ru/en/papers/what-has-happened-to-lake-baikal-sponges/>
- Chen L., Zheng D., Liu B. et al. 2016. VFDB 2016: Hierarchical and refined dataset for big data analysis - 10 years on. Nucleic Acids Research 44: D694–D697. DOI: [10.1093/NAR/GKV1239](https://doi.org/10.1093/NAR/GKV1239)
- Chen W.-M., Huang W.-C., Young C.-C. et al. 2012. *Flavobacterium tilapiae* sp. nov., isolated from a freshwater pond, and emended descriptions of *Flavobacterium defluvii* and *Flavobacterium johnsoniae*. International Journal of Systematic and Evolutionary Microbiology 63: 827–834. DOI: [10.1099/ijs.0.041178-0](https://doi.org/10.1099/ijs.0.041178-0)
- Chernogor L., Klimenko E., Khanaev I. et al. 2020. Microbiome analysis of healthy and diseased sponges *Lubomirskia baikalensis* by using cell cultures of primmorphs. PeerJ 8: e9080. DOI: [10.7717/peerj.9080](https://doi.org/10.7717/peerj.9080)
- Chernogor L.I., Petrushin I.S., Khanaev I.V. et al. 2020. Study of changes in microbiomes of sponges *Lubomirskia baikalensis* using the cell culture of primmorphs. Limnology and Freshwater Biology 4: 647–648. DOI: [10.31951/2658-3518-2020-A-4-647](https://doi.org/10.31951/2658-3518-2020-A-4-647)
- Denikina N.N., Dzyuba E.V., Belkova N.L. et al. 2016. The first case of disease of the sponge *Lubomirskia baikalensis*: investigation of its microbiome. Translated in Biology Bulletin 43: 263–270. DOI: [10.1134/S106235901603002X](https://doi.org/10.1134/S106235901603002X)
- Ekwe A.P., Ahn J.-H., Kim S.B. 2017. *Flavobacterium keumense* sp. nov., isolated from freshwater. International Journal of Systematic and Evolutionary Microbiology 67: 2166–2170. DOI: [10.1099/ijsem.0.001920](https://doi.org/10.1099/ijsem.0.001920)
- Galaxy platform for accessible, reproducible, and collaborative data analyses. 2024. Nucleic Acids Research. Oxford Academic, n.d. URL: <https://academic.oup.com/nar/article/52/W1/W83/7676834?login=false> (accessed: 11.16.25).
- Jain C., Rodriguez-R L.M., Phillippy A.M. et al. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nature communications 9(1): 5114. DOI: [10.1038/S41467-018-07641-9](https://doi.org/10.1038/S41467-018-07641-9)
- Khanaev I.V., Kravtsova L.S., Maikova O.O. et al. 2018. Current state of the sponge fauna (Porifera: Lubomirskiidae) of Lake Baikal: sponge disease and the problem of conservation of diversity. Journal of Great Lakes Research 44: 77–85. DOI: [10.1016/j.jglr.2017.10.004](https://doi.org/10.1016/j.jglr.2017.10.004)
- Kozhov M. 2013. Lake Baikal and its life. Springer Science business Media. DOI: [10.1007/978-94-015-7388-7](https://doi.org/10.1007/978-94-015-7388-7)
- Kozhov M.M. 1972. Essays on Baikal studies. Irkutsk: East Siberian Publisher. (in Russian)
- Liu B., Zheng D., Zhou S. et al. 2022. VFDB 2022: A general classification scheme for bacterial virulence factors. Nucleic Acids Research 50: D912–D917. DOI: [10.1093/nar/gkab1107](https://doi.org/10.1093/nar/gkab1107)
- Meier-Kolthoff J.P., Carbasse J.S., Peinado-Olarte R.L. et al. 2022. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. Nucleic Acids Research 50: D801–D807. DOI: [10.1093/NAR/GKAB902](https://doi.org/10.1093/NAR/GKAB902)
- Meier-Kolthoff J.P., Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nature Communications 10: 1–10. DOI: [10.1038/s41467-019-10210-3](https://doi.org/10.1038/s41467-019-10210-3)
- Peeters K., Willems A. 2011. The *gyrB* gene is a useful phylogenetic marker for exploring the diversity of *Flavobacterium* strains isolated from terrestrial and aquatic habitats in Antarctica. FEMS Microbiology Letters 321: 130–140. DOI: [10.1111/J.1574-6968.2011.02326.X](https://doi.org/10.1111/J.1574-6968.2011.02326.X)
- Petrushin I., Belikov S., Chernogor L. 2020. Cooperative interaction of *Janthinobacterium* sp. SLB01 and *Flavobacterium* sp. SLB02 in the diseased sponge *Lubomirskia baikalensis*. International journal of molecular sciences 21: 1–13. DOI: [10.3390/ijms21218128](https://doi.org/10.3390/ijms21218128)
- Pile A.J., Patterson M.R., Savarese M. et al. 1997. Trophic effects of sponge feeding within Lake Baikal's littoral zone. 2. Sponge abundance, diet, feeding efficiency, and carbon flux. Limnology and Oceanography 42: 178–184. DOI: [10.4319/lo.1997.42.1.0178](https://doi.org/10.4319/lo.1997.42.1.0178)
- Prjibelski A., Antipov D., Meleshko D. et al. 2020. Using SPAdes De Novo Assembler. Current Protocols in Bioinformatics 70: e102. DOI: [10.1002/CPBI.102;PAGE=STRING:ARTICLE/CHAPTER](https://doi.org/10.1002/CPBI.102;PAGE=STRING:ARTICLE/CHAPTER)
- Rawlings N.D., Barrett A.J., Thomas P.D. et al. 2018. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Research 46: D624–D632. DOI: [10.1093/nar/gkx1134](https://doi.org/10.1093/nar/gkx1134)
- Rawlings N.D., Bateman A. 2021. How to use the MEROPS database and website to help understand peptidase specificity. Protein science: a publication of the Protein Society 30: 83–92. DOI: [10.1002/pro.3948](https://doi.org/10.1002/pro.3948)
- Rawlings N.D., Waller M., Barrett A.J. et al. 2014. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Research 42: D503–509. DOI: [10.1093/nar/gkt953](https://doi.org/10.1093/nar/gkt953)
- Seemann T. 2016. ABRicate: mass screening of contigs for antibiotic resistance genes. URL: <https://github.com/tseemann/abricate>
- Semiturkina N.A., Efremova S.M., Timoshkin O.A. 2009. State-of-the art of biodiversity and ecology of spongiofauna of Lake Baikal with special attention to the diversity, peculiarities of ecology and vertical distribution of Porifera on Berezovy ecology test site. In: Timoshkin O.A. (Ed.), Index of animal species inhabiting Lake Baikal and its catchment area. Novosibirsk, pp. 891–901. (in Russian)
- Tatusova T., Dicuccio M., Badretin A. et al. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Research 44: 6614–6624. DOI: [10.1093/NAR/GKW569](https://doi.org/10.1093/NAR/GKW569)

Timoshkin O.A., Samsonov D.P., Yamamuro M. et al. 2016. Rapid ecological change in the coastal zone of Lake Baikal (East Siberia): Is the site of the world's greatest freshwater biodiversity in danger? *Journal of Great Lakes Research* 42: 487–497. DOI: [10.1016/j.jglr.2016.02.011](https://doi.org/10.1016/j.jglr.2016.02.011)

Yi H., Oh H.-M., Lee J.-H. et al. 2005. *Flavobacterium antarcticum* sp. nov., a novel psychrotolerant bacterium isolated from the Antarctic. *International Journal of Systematic and Evolutionary Microbiology* 55: 637–641. DOI: [10.1099/ijs.0.63423-0](https://doi.org/10.1099/ijs.0.63423-0)

Zheng J., Ge Q., Yan Y. et al. 2023. dbCAN3: automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Research* 51: W115–W121. DOI: [10.1093/nar/gkad328](https://doi.org/10.1093/nar/gkad328)



# Геномные основы деградации тканей байкальских губок за счет свободно живущих флавобактерий с локусами утилизации полисахаридов, ориентированными на гликаны хозяина

Панова Ю.А.<sup>1,2</sup>, Беликов С.И.<sup>1\*</sup>, Черногор Л.И.<sup>1</sup>

<sup>1</sup> Лимнологический институт СО РАН, ул. Улан-Баторская, 3, Иркутск, 664033, Россия

<sup>2</sup> Сибирский федеральный университет, пр. Свободный, 79, Красноярск, 660041, Россия

**АННОТАЦИЯ.** В данной работе мы представляем полную последовательность генома штамма *Flavobacterium* sp. PLB03, выделенного из культуры клеток примморф губки *Lubomirskia baikalensis*, и сравниваем его с патогенными и свободно живущими представителями рода *Flavobacterium*. Штамм *Flavobacterium* sp. PLB03 проявляет выраженную литическую активность в отношении клеток губки, а также вызывает гибель симбиотических микроводорослей (*Trebouxiophyceae*). Геном *Flavobacterium* sp. PLB03 имеет длину 5 925 828 п.н. и содержание ГЦ 34,3%. Средняя идентичность нуклеотидов (ANI) 98,9% показала, что штамм *Flavobacterium* sp. PLB03 имеет наибольшее сходство с *Flavobacterium* CSZ. Сравнительный геномный анализ штамма и филогенетически родственного рода *Flavobacterium* показал, что изолированный штамм *Flavobacterium* sp. PLB03 имеет большой размер генома, характерный для экологического рода *Flavobacterium*, а соотношение генов пептидаз и гликозидгидролаз составляет 1,8, что позволяет предположить, что штамм *Flavobacterium* sp. PLB03 обладает характеристиками, сходными с непатогенными штаммами.

**Ключевые слова:** *Flavobacterium* sp. PLB03, пресноводные губки, CAZymes, локусы утилизации полисахаридов, система секреции IX типа, геномные острова, условная патогенность

**Для цитирования:** Панова Ю.А., Беликов С.И., Черногор Л.И. Геномные основы деградации тканей байкальских губок за счет свободно живущих флавобактерий с локусами утилизации полисахаридов, ориентированными на гликаны хозяина // Limnology and Freshwater Biology. 2025. - № 6. - С. 1313-1328. DOI: 10.31951/2658-3518-2025-A-6-1313

## 1. Введение

Озеро Байкал — одно из древнейших озёр в геологической истории, содержащее почти 20% мировых запасов пресной воды. Озеро отличается разнообразием эндемичной флоры и фауны, а байкальские губки являются наиболее многочисленными организмами в прибрежной зоне с биомассой более 700 г/м<sup>2</sup> (Кожов, 1972; 2013; Pile et al., 1997; Semiturkina et al., 2009).

Заболевание, поражающее пресноводные губки, было впервые обнаружено в 2011 году, когда в центральной котловине озера Байкал были обнаружены розовые особи (Bormotov, 2012). Изменение цвета губок было вызвано гибелью их основного симбионта, зелёной водоросли *Choricystis* sp. и дальнейшей заменой на цианобактерии *Synechococcus* sp. (Denikina et al., 2016; Belikov et al., 2019). В после-

дующие годы внешние признаки заболевания губок быстро менялись, затем с 2013 года и по настоящее время наблюдались различные симптомы поражения тканей губок, такие как изменение цвета, некроз тканей и образование коричневых и грязно-фиолетовых бактериальных биопленок. Численность губок значительно сократилась, и в настоящее время больные губки встречаются по всему озеру Байкал. Кроме того, отмечено, что заболевание и гибель байкальских губок сопровождается многочисленными изменениями в экосистеме литорали озера Байкал (Timoshkin et al., 2016; Khanaev et al., 2018).

Ранее нами было показано, что заболевание губок связано с изменениями и резкими сдвигами в составе микробных сообществ байкальских губок. Нами установлено, что наблюдаемый дисбаланс в микробных сообществах больных губок обусловлен

\*Автор для переписки. Адрес e-mail: [sergeibelikov47@gmail.com](mailto:sergeibelikov47@gmail.com) (С.И. Беликов)

**Поступила:** 20 ноября 2025;

**Принята после доработки:** 09 декабря 2025;

**Опубликована online:** 25 декабря 2025

© Автор(ы) 2025. Эта работа распространяется под международной лицензией Creative Commons Attribution-NonCommercial 4.0.



различными микроорганизмами, которые, действуя согласованно, усиливают свое негативное воздействие, приводящее к гибели байкальских губок (Belikov et al., 2019).

Для выявления патогенных для губок микроорганизмов нами было проведено экспериментальное заражение клеточной культуры примморф бактериальной суспензией от больных губок. Выявлено значительное увеличение относительного содержания бактерий семейства *Oxalobacteraceae* – до 25,65% и семейства *Flavobacteriaceae* – до 62,04% при одновременном снижении численности других микроорганизмов (Chernogor et al., 2020a; Chernogor et al., 2020b). На основании анализа геномов выделенных штаммов *Janthinobacterium* SLB01 и *Flavobacterium* SLB02 была выдвинута гипотеза о возможном совместном взаимодействии этих двух бактерий в развитии болезни губок (Petrushin et al., 2020; Belikov et al., 2021).

Известно, что бактерии представители рода *Flavobacterium* широко распространены. Например, типовые штаммы *Flavobacterium* spp. были выделены из различных источников, включая пресную воду, морской лед Антарктики, ризосферу и больных рыб (Yi et al., 2005; Ali et al., 2009; Chen et al., 2012; Ekwe et al., 2017). Число официально описанных видов *Flavobacterium* быстро растет и по состоянию на октябрь 2025 года включает 448 видов (<https://lpsn.dsmz.de/genus/Flavobacterium>).

Штамм *Flavobacterium* sp. PLB03 был впервые выделен из клеточных культур примморф больной губки *L. baikalensis*. Роль этих бактерий и их связь с развитием заболеваний и гибелью губок остаются недостаточно изученными. В данном исследовании мы представляем собранный и аннотированный геном *Flavobacterium* sp. PLB03, и сравниваем его с геномами патогенных для рыб и непатогенных *Flavobacterium*. Ожидается, что последовательность генома *Flavobacterium* sp. PLB03 и проведенный анализ предоставят более полную информацию о факторах, связанных с заболеваниями. Выявление этих геномных особенностей поможет понять причины столь широкого распространения заболеваний и гибели байкальских губок.

## 2. Материалы и методы

### 2.1. Секвенирование генома, контроль качества и сборка

Необработанные прочтения секвенирования Illumina оценивали с помощью FastQC v0.11.9 (Andrews, 2010) для оценки качества оснований, содержания GC и загрязнения адаптерами. Низкокачественные основания и последовательности адаптеров обрезали с помощью Trimmomatic v0.39 (Bolger et al., 2014). Геном собирали с помощью SPAdes v3.15.4 (Bankevich et al., 2012; Prjibelski et al., 2020) с использованием стандартных итераций k-меров. Качество сборки и загрязнение оценивали с помощью FCS-GX (Astashyn et al., 2024) для обнаружения и удаления предполагаемых чужеродных контигов. Контиги короче 200 п.н. или поме-

ченные, как загрязняющие отбрасывали, что давало окончательную сборку, используемую для последующих анализов.

### 2.2. Аннотация генома

Финальная сборка была аннотирована с помощью инструмента NCBI Prokaryotic Genome Annotation Pipeline (PGAP) версии 6.9 (Tatusova et al., 2016), который объединяет *ab initio* предсказание генов, классификацию семейств белков на основе HMM и подобранные функциональные модели RefSeq. Аннотация включала кодирующие последовательности, некодирующие РНК, тРНК, фрагменты рРНК и псевдогены.

### 2.3. Геномы, включенные в анализ

Для изучения филогеномного и функционального контекста *Flavobacterium* sp. PLB03 мы отобрали 14 общедоступных геномов, представляющих, как свободноживущие, так и патогенные для рыб линии рода *Flavobacterium*. Геномы были получены из NCBI GenBank в период с марта по май 2025 года. Итоговый сравнительный набор данных включал следующие виды: *F. sp. CSZ*, *F. piscisymbiosum* F-30, *F. piscis* CCUG 60099, *F. bizetiae* HJ-32-4, *F. bizetiae* FlaQc-53, *F. bizetiae* CIP 105534, *F. aquidurens* DSM 18293, *F. johnsoniae* FJOH, *F. pectinovorum* ZE23VCel01, *F. soyae* SCIV07, *F. collinsii* FlaQc-26, *F. lipolyticum* F-126, *F. chilense* DSM 24724, *F. psychroterrae* CCM 8827, as well as the fish-pathogenic species *F. columnare* 04017018, *F. psychrophilum* JIP02-86, *F. branchiophilum* FL-15, and *F. tractae* ATCC BAA-2541. Подробная статистика сборки и характеристики аннотации генома для всех штаммов приведены в Таблице S1 Приложения.

### 2.4. Филогения и анализ средней идентичности нуклеотидов

Был проведен филогенетический анализ генома с использованием платформы TYGS (Meier-Kolthoff and Göker, 2019; Meier-Kolthoff et al., 2022) для выявления наиболее близких таксонов и выбора подходящих референтных геномов, для последующего сравнительного анализа. Поскольку геном *Flavobacterium* sp. PLB03 сильно фрагментирован, а все восстановленные последовательности генов 16S рРНК были неполными, для филогенетической реконструкции мы использовали ген домашнего хозяйства *gyrB*. Этот подход соответствует предыдущим рекомендациям по таксономии флавобактерий, где *gyrB* обеспечивает более высокое филогенетическое разрешение и предпочтителен в случаях, когда 16S рРНК недоступна или неполна (Peeters and Willems, 2011). Средняя идентичность нуклеотидов (ANI) по всему геному была рассчитана с помощью FastANI (Jain et al., 2018). Для этого анализа мы включили следующие геномы, представляющие близкородственных представителей рода *Flavobacterium*: *F. sp. CSZ*, *F. piscisymbiosum*

штамм F-30, *F. piscis* CCUG 60099, *F. bizetiae* штамм HJ-32-4, *F. sp.* KBS0721, *F. bizetiae* штамм FlaQc-53 и *F. bizetiae* штамм CIP 105534. Эти геномы были отобраны на основе результатов кластеризации TYGS и их близости к *Flavobacterium sp.* PLB03 в предварительных анализах.

## 2.5. Выявление генов устойчивости к противомикробным препаратам и генов вирулентности

Потенциальные гены устойчивости к противомикробным препаратам и факторы вирулентности были проверены с помощью ABRicate v1.0.1 (Seemann, 2016) по базам данных CARD, NCBI AMR и VFDB (Chen et al., 2016; Liu et al., 2022) с использованием платформы Galaxy 2024.

## 2.6. Идентификация систем секреции

Компоненты системы секреции IX типа (T9SS) были идентифицированы с помощью ручного просмотра аннотированных генов и поиска BLASTp.

## 2.7. Прогнозирование геномных островов

Геномные острова (ГО) были предсказаны с помощью IslandViewer 4 (Bertelli et al., 2017), который объединяет методы SIGI-HMM и IslandPath-DIMOB. Гены в пределах ГО были аннотированы с помощью PGAP. Были зарегистрированы функциональные категории, включая мобильные элементы, компоненты секреции, углевод-активные ферменты (CAZymes), ТонВ-зависимые рецепторы и протеазы.

## 2.8. Аннотация протеаз и сравнительный анализ по базе MEROPS

Протеазы были аннотированы с помощью поиска BLASTp в базе данных MEROPS (Rawlings et al., 2014; 2018; Rawlings and Bateman, 2021) ( $E\text{-value} \leq 1e-10$ ). Количество протеаз на уровне семейств было суммировано для штамма *Flavobacterium sp.* PLB03 и сравнено с референтными видами, в частности, *Flavobacterium johnsoniae*. Протеазы были классифицированы по каталитическому типу (сериновые, металлопротеазы, цистеиновые, аспарагиновые) и по предполагаемой субклеточной локализации.

## 2.9. Идентификация CAZymes и локусов утилизации полисахаридов (ПУЛ) и анализ по видам *Flavobacterium*

Углевод-активные ферменты (CAZymes) были аннотированы с помощью dbCAN3 (Zheng et al., 2023). Были отобраны только ферменты, подтвержденные как минимум двумя методами. Локусы утилизации полисахаридов (ПУЛ) были предска-

заны с помощью конвейера dbCAN CGC, который обнаруживает со-локализованные CAZy-рецепторы, SusC/D-подобные ТонВ-зависимые транспортеры и регуляторы транскрипции. Субстраты ПУЛ были определены на основе состава семейства CAZymes, прогнозирования субстратов dbCAN и данных литературы об известной архитектуре ПУЛ у *Flavobacterium* и других *Bacteroidetes*.

Чтобы определить контекст *Flavobacterium sp.* PLB03 в пределах рода, мы сравнили репертуары CAZymes в 15 геномах флавобактерий, включая представителей окружающей среды и виды, патогенные для рыб. Этот широкомасштабный анализ выявил основные закономерности разнообразия CAZymes и показал, что *Flavobacterium sp.* PLB03 обладает одним из самых расширенных репертуаров CAZymes, особенно в семействах, ассоциированных с гликанами хозяина. Поскольку архитектура ПУЛ значительно различается у разных видов и отсутствует гомология на уровне локусов в пределах рода, детальное сравнение ПУЛ было ограничено шестью репрезентативными геномами: *Flavobacterium sp.* PLB03, *F. johnsoniae*, *F. piscisymbiosum*, *F. bizetiae*, *F. tractae* и *F. columnare*.

Сравнительный анализ профилей CAZymes, репертуаров ПУЛ и распределений семейств базы MEROPS был проведен по 15 геномам флавобактерий. Матрицы данных были построены на основе подсчетов на уровне семейств и проанализированы методом иерархической кластеризации с использованием метода Уорда и евклидовых расстояний в R.

## 3. Результаты

### 3.1. Результаты сборки и аннотации генома

Черновик сборки генома включал 334 контига общей длиной 6 003 505 пар оснований (п.н.). После удаления коротких контигов (<200 п.н.) и контаминантов, окончательная сборка содержала 93 контига общей длиной 5 925 828 п.н., значение N50 составляло 542 557 п.н., а содержание GC – 34,3%. Эта сборка генома была представлена в GenBank под номером доступа JBLIYN000000000 и подготовлена для последующей аннотации и анализа с помощью PGAP. Основные показатели сборки и аннотации представлены в Таблице 1.

Размер генома больше, чем у патогенных для рыб видов *Flavobacterium*, и сопоставим с природными штаммами с расширенным метаболическим репертуаром (Таблице S1 Приложения).

### 3.2. Филогения и анализ средней идентичности нуклеотидов

Анализ исследования, с применением сервиса TYGS и анализ результатов BLAST в NCBI показали, что последовательности *Flavobacterium sp.* PLB03 наиболее тесно связаны с последовательностями неустоановленного вида *Flavobacterium sp.* CSZ, и *Flavobacterium sp.* KBS0721 и *F. bizetiae*, *F.*

*piscisymbiosum* и *F. piscis*. Геномное сравнение этих штаммов на основе средней идентичности нуклеотидов показало, что *Flavobacterium* sp. PLB03 образует тесный кластер на видовом уровне только с *Flavobacterium* sp. CSZ с ANI 98,96%, что значительно превышает 95% порог видовой принадлежности. Все остальные проанализированные геномы *Flavobacterium* показали значения средней идентичности нуклеотидов 89,3–90,8% относительно *Flavobacterium* sp. PLB03, что указывает на межвидовую дивергенцию. Ближайшим видом вне пары *Flavobacterium* sp. PLB03 - CSZ был *F. piscisymbiosum* F-30 (90,8% ANI), известный таксон, связанный с рыбами, хотя и значительно ниже видовой границы (таблица S2 Приложения). В совокупности эти результаты показывают, что *Flavobacterium* sp. PLB03 и *F. sp. CSZ* образуют отдельную видовую линию.

Филогенетический анализ *gyrB* поместил *Flavobacterium* sp. PLB03 в экологическую линию флавобактерий, наиболее близкую к *Flavobacterium* sp. CSZ и находящуюся на значительном расстоянии от *F. psychrophilum*, *F. columnare* и *F. branchiophilum*, которые являются патогенами рыб (Рис. 1).

3.3. Выявление генов устойчивости к противомикробным препаратам и генов вирулентности

При стандартных пороговых значениях строгости (идентичность и покрытие  $\geq 75\text{--}90\%$ ), Abricate не обнаружил факторов вирулентности в геноме *Flavobacterium* sp. PLB03. При снижении пороговых значений до 60% идентичности и 60% покрытия, Abricate сообщил о четырёх совпадениях, демонстрирующих частичное сходство с белками, ассоциированными с вирулентностью, описанными в других бактериальных таксонах: *katA*, *htpB* (GroEL/Hsp60), *neuC1* (UDP-N-ацетилглюкозамин-2-эпимераза) и одним белком, аннотированным в VFDB как эффектор системы секреции Dot/Icm типа IV (*lpg2370*).

Все совпадения показали низкую аминокислотную идентичность (66–68%) и неполное покрытие выравнивания (70–90%), что указывает на отдалённую гомологию, а не на функциональную эквивалентность. Эти гены кодируют основные функции «домашнего хозяйства», широко распространённые у бактерий, и не считаются полноценными детерминантами вирулентности рода *Flavobacterium*. Таким образом, штамм *Flavobacterium* sp. PLB03 не кодирует традиционные системы вирулентности, описанные у патогенных для рыб штаммов, таких как *F. psychrophilum* и *F. columnare*.

3.4. Геномные острова

IslandViewer предсказал 23 геномных острова размером от 4,1 до 74,9 т.п.н. и общим размером около 540 т.п.н. (10–12% генома). Гены, ассоциированные с геномными островами, включали гипотетические белки, мобильные элементы, рецеп-

торы, зависящие от TonB, и отдельные компоненты системы секреции IX типа. Геномные острова были рассредоточены по всему геному и не образовывали крупные кластеры (Рис. 2). Большинство геномных островов были предсказаны как минимум двумя независимыми алгоритмами, что свидетельствует о высокой достоверности их границ.

Таблица 1. Статистика необработанных прочтений и геномных признаков *Flavobacterium* sp. PLB03

| Свойство                            | <i>Flavobacterium</i> sp. PLB03 |
|-------------------------------------|---------------------------------|
| Необработанные прочтения            | 7 220 904                       |
| Длина сборки, п.н.                  | 5 925 828                       |
| Количество контигов                 | 93                              |
| Длина самого длинного контига, п.н. | 940 434                         |
| N50, п.н.                           | 542 557                         |
| N90, п.н.                           | 198 207                         |
| Содержание GC                       | 34,3%                           |
| Общее количество генов              | 5 104                           |
| Белок-кодирующие последовательности | 4 975                           |
| тРНК                                | 64                              |
| Некодирующие РНК                    | 3                               |
| Псевдогены                          | 32                              |
| Номер доступа в GenBank             | JBLIYN000000000                 |

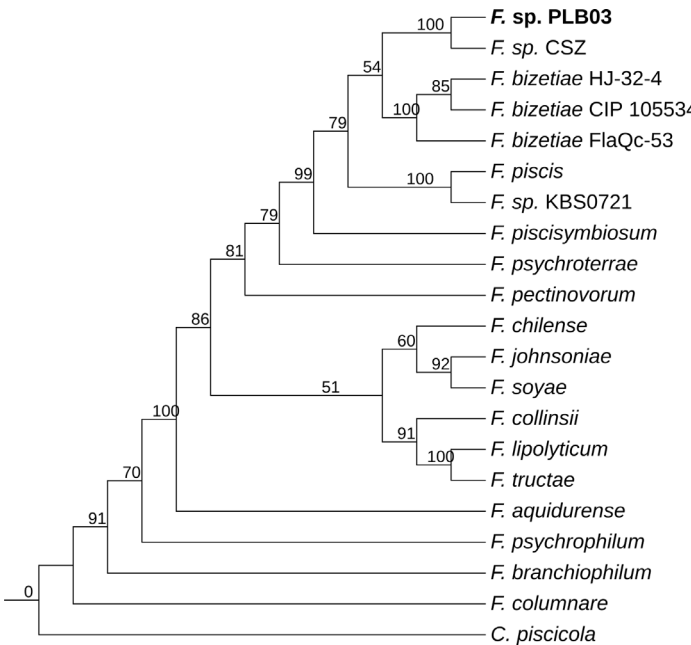


Рис.1. Дерево объединения соседей, основанное на полных последовательностях *gyrB*.

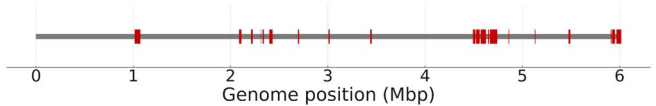


Рис.2. Распределение геномных островов в геноме *Flavobacterium* sp. PLB03.



Предсказанные геномные острова были обогащены гипотетическими белками (32%, n = 129), а также транспозазами и интегразами (4%), мембранно-ассоциированными белками (4%) и регуляторными элементами, такими, как транскрипционные регуляторы типа «спираль-поворот-спираль» (4%). Некоторые острова содержали компоненты, типичные для флавобактерий, включая гены, ассоциированные с системой секреции типа IX, рецепторы, зависимые от TonB, и белки наружной мембраны, связанные с SusC/RagA TonB (n = 4), а также транспозазы семейства IS6 (n = 4).

3.5. Углевод-активные ферменты (CAZymes)

Геном *Flavobacterium* sp. PLB03 кодирует 392 CAZymes, включая 185 гликозид гидролаз, 53 гликозил трансферазы, 36 углеводов эстераз, 14 полисахарид лиаз, 28 углеводов связывающих модулей и 7 ферментов вспомогательной активности. Это делает *Flavobacterium* sp. PLB03 одним из наиболее богатых углеводов-активными ферментами микроорганизмов среди изученных пресноводных видов флавобактерий (Таблица 2). Репертуар CAZymes охватывает широкий спектр полисахаридов растений и водорослей (ксилан, пектин, ксилоглюкан, бета-глюкан, маннан, альгинат, альфа-глюкан), а также семейства, обычно ассоциированные с гликанами хозяина, такие как GH29, GH33, GH92, GH95, GH105, GH127, PL12, PL29 и CBM32. Последние играют ключевую роль в деградации муцинов, гликозаминогликанов и гликопротеинов.

3.6. Локусы утилизации полисахаридов

С помощью dbCAN3 и конвейера CGC в геноме *Flavobacterium* sp. PLB03 было идентифицировано 32 локуса утилизации полисахаридов

(ПУЛ). Сравнительный анализ с пятью видами (*F. johnsoniae*, *F. bizetiae*, *F. piscisymbiosum*, *F. tractae*, *F. columnare*) выявил в общей сложности 134 ПУЛ (см. Дополнительную таблицу S3). *Flavobacterium* sp. PLB03 входит в число видов с наибольшим количеством ПУЛ, сопоставимых только с *F. piscisymbiosum* и *F. bizetiae*, тогда, как патогенные для рыб штаммы кодируют гораздо меньшее количество ПУЛ (от 1 до 11).

Сравнение архитектур ПУЛ на основе структур выявило по меньшей мере 11 ПУЛ, уникальных для штамма *Flavobacterium* sp. PLB03, среди всех исследованных геномов. Наиболее заметные уникальные кластеры нацелены на гликаны, ассоциированные с хозяином (Таблица 3).

Ни одна из этих структур не была выявлена у остальных пяти видов флавобактерий.

3.7. Анализ протеаз

Аннотация протеаз и ингибиторов протеаз с помощью базы MEROPS выявила широкое ферментативное разнообразие в проанализированных геномах флавобактерий. После нормализации аннотации MEROPS по унифицированным кодам семейств, все виды имели консервативное ядро протеаз, в котором доминировали семейства I39, S33, M14, S12 и S9 (Дополнительная таблица S4). Среди них семейство ингибиторов I39 было наиболее распространенным практически во всех геномах. Несмотря на общее ядро, матрица протеаз по видам и семействам показала существенные различия в степени расширения семейств. Некоторые виды обладали компактными репертуарами, в то время как другие демонстрировали значительное обогащение несколькими S- и I-семействами. Геном штамма *Flavobacterium* sp. PLB03 кодировал большое количество генов нескольких основных семейств протеаз (Рис. 3).

Таблица 2. Сравнительное распределение классов CAZy среди 15 штаммов

| Виды                                    | CAZy-гены | GH  | GT | CE | PL | CBM | AA |
|-----------------------------------------|-----------|-----|----|----|----|-----|----|
| <i>Flavobacterium</i> sp. PLB03         | 392       | 185 | 53 | 36 | 14 | 28  | 7  |
| <i>F. soyae</i> SCIV07                  | 408       | 183 | 66 | 32 | 10 | 36  | 6  |
| <i>F. piscisymbiosum</i> F-30           | 386       | 175 | 57 | 38 | 13 | 24  | 7  |
| <i>F. pectinovorum</i> ZE23VCel01       | 392       | 172 | 60 | 33 | 11 | 34  | 3  |
| <i>F. johnsoniae</i> FJOH               | 382       | 165 | 59 | 31 | 12 | 29  | 4  |
| <i>F. bizetiae</i> HJ-32-4              | 363       | 153 | 58 | 33 | 9  | 25  | 3  |
| <i>F. aquidurensense</i> DSM 18293      | 343       | 149 | 39 | 42 | 14 | 18  | 5  |
| <i>F. chilense</i> DSM 24724            | 348       | 134 | 54 | 34 | 15 | 28  | 4  |
| <i>F. psychroterrae</i> CCM 8827        | 331       | 127 | 46 | 28 | 15 | 32  | 3  |
| <i>F. tractae</i> ATCC BAA-2541         | 276       | 55  | 57 | 25 | 7  | 7   | 6  |
| <i>F. lipolyticum</i> F-126             | 278       | 55  | 60 | 24 | 7  | 8   | 6  |
| <i>F. collinsii</i> FlaQc-26            | 270       | 59  | 52 | 21 | 9  | 7   | 5  |
| <i>F. branchiophilum</i> FL-15 (пат.)   | 180       | 29  | 51 | 14 | 2  | 7   | 2  |
| <i>F. columnare</i> 04017018 (пат.)     | 154       | 18  | 44 | 12 | 3  | 1   | 1  |
| <i>F. psychrophilum</i> JIP02-86 (пат.) | 137       | 10  | 45 | 9  | 0  | 0   | 1  |

Примечание: GH – гликозид гидролазы; GT – гликозил трансферазы; CE – углеводов эстеразы; PL – полисахарид лиазы; CBM – углеводов связывающие модули; AA – вспомогательные активности.

**Таблица 3.** Гликаны хозяина и водорослевые полисахариды, нацеленные на локусы утилизации полисахаридов, выявленные в *Flavobacterium* sp. PLB03

| Тип ЛУП                                                 | Ключевые семейства CAZymes    | Примечания                                                             |
|---------------------------------------------------------|-------------------------------|------------------------------------------------------------------------|
| Муцин-деградирующие ЛУП                                 | GH105, GH88, PL12/PL8, CBM32  | Каноническая архитектура, нацеленная на муцин                          |
| Гликозаминогликан (GAG) ЛУП                             | GH154, GH159, GH128, PL12/PL8 | Два отдельных GAG-специфичных локуса                                   |
| Гликан хозяина/гликопротеин животного происхождения ЛУП | GH29, GH33, GH92, GH130       | Большой локус, обогащенный фукозидазами, сиалидазами и α-маннозидазами |
| Арабиногалактан/AGP-тип                                 | GH105, GH128, GH154, PL42     | Нацелен на арабиногалактан и AGP-подобные субстраты                    |
| Альгинат-специфичные ЛУП                                | PL6, PL17, CE20               | Уникальный Локус, разрушающий альгинат, отсутствует у других видов     |

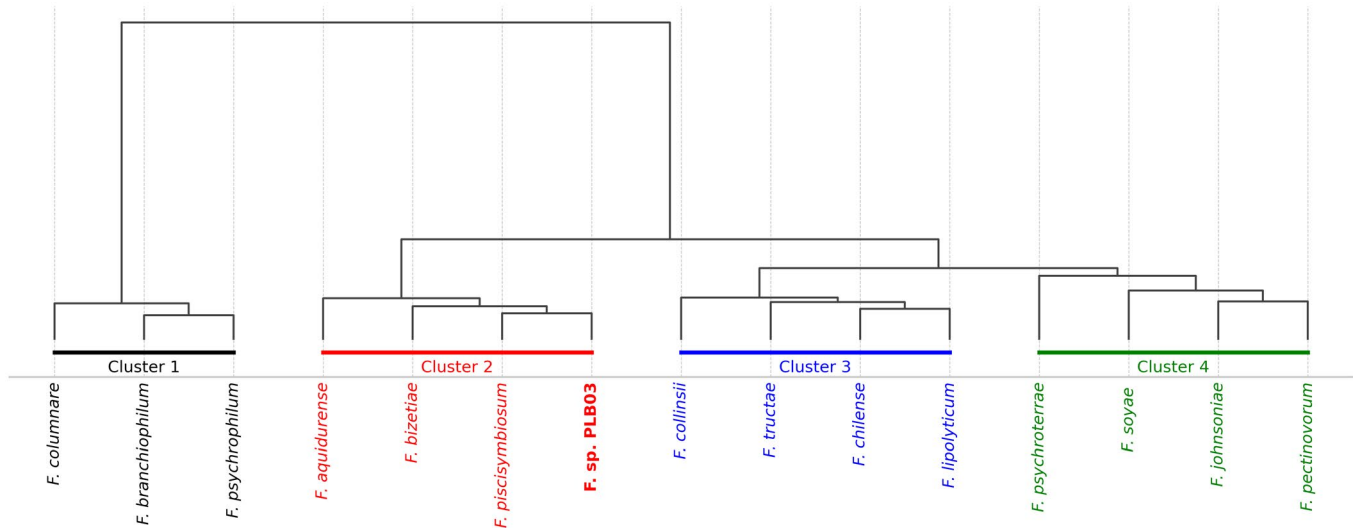
Дальнейшее сравнение семейств протеаз в кластере 2 выявило четыре семейства MEROPS: M10, S1, T2, T3, которые присутствуют исключительно в геноме штамма *Flavobacterium* sp. PLB03, но отсутствуют у остальных представителей этого кластера. Все четыре семейства присутствуют у отдельных видов из кластеров 1 и 3, но отсутствуют в геномах других представителей кластера 4.

4. Обсуждение

Геном изолированного штамма *Flavobacterium* sp. PLB03 демонстрирует сочетание признаков, типичных для свободноживущих представителей рода *Flavobacterium*, с модулями, обычно связанными с взаимодействием, с хозяином. Филогенетическое расположение и высокие значения средней идентичности нуклеотидов ясно демонстрируют, что *Flavobacterium* sp. PLB03 образует отдельную видовую линию вместе с *Flavobacterium* sp. CSZ, оставаясь при этом далёким от хорошо охарактеризованных патогенных для рыб штаммов, таких как *F. psychrophilum*, *F. columnare* и *F. branchiophilum*. Такое расположение подтверждает мнение о том, что штамм *Flavobacterium* sp. PLB03 является, прежде всего, экологическим таксоном, а не специализированным патогеном.

Несмотря на отсутствие классических детерминант вирулентности — даже при смягченных порогах обнаружения — геном *Flavobacterium* sp. PLB03 содержит ряд функциональных особенностей, которые могут усиливать его способности разрушать ткани губок и зеленых симбиотических микроводорослей. Отмечено, что геном *Flavobacterium* sp. PLB03 больше, чем у патогенных для рыб штаммов и кодирует расширенный метаболический репертуар. Так, 392 гена CAZymes делают геном штамма *Flavobacterium* sp. PLB03 одним из наиболее богатых углевод-утилизирующих видов среди пресноводных *Flavobacterium*, изученных ранее. Особое внимание привлекает наличие семейств CAZymes, обычно связанных с деградацией эукариотических гликоконъюгатов, включая GH29, GH33, GH92, GH95, GH105, GH127, PL12, PL29 и CBM32. Эти ферменты обеспечивают расщепление муцинов, гликозамино-гликанов и сложных животных гликопротеинов — ключевых структурных полисахаридов в тканях губок.

Этот метаболический потенциал отражен в архитектуре 32 идентифицированных локусов утилизации полисахаридов (ПУЛ). Среди них есть как минимум пять крупных кластеров, специфичных для *Flavobacterium* sp. PLB03, нацеленных на ассоциированные с хозяином гликаны: канонический ПУЛ, деградирующий муцин, два независимых локуса



**Рис.3.** Иерархическая кластеризация видов *Flavobacterium* на основе профилей семейства протеаз по MEROPS.

гликозаминогликанов (GAG), большой ПУЛ, связанный с гликанами хозяина/животным гликопротеином, обогащенный фукозидазами, сиалидазами и  $\alpha$ -маннозидазами, ПУЛ, связанный с арабиногалактаном типа AGP, и уникальный локус, деградирующий альгинат, отсутствующий во всех сравниваемых геномах. Такой разнообразный репертуар ПУЛ предполагает широкий спектр субстратов и характерен для «оппортунистических» флавобактерий, способных использовать как полисахариды окружающей среды, так и полисахариды эукариотического (животного) происхождения. Дополнительным доказательством адаптивной универсальности служит наличие 23 геномных островов, составляющих 10–12% генома. Эти острова обогащены гипотетическими белками, мобильными элементами, TonB-зависимыми рецепторами, компонентами системы секреции IX типа и SusC/RagA-подобными транспортерами, что указывает на активную модульность генома и потенциальное горизонтальное приобретение генов. Распределенное распределение модулей TonB–SusC, ассоциированных с геномными островами, согласуется с типичной организацией ПУЛ у флавобактерий, что дополнительно подтверждает идею о субстрат-зависимой пластичности генома.

Анализ протеаз выявил расширенный репертуар MEROPS у *Flavobacterium* sp. PLB03, включающий четыре семейства (M10, S1, T2, T3), которые отсутствуют у других представителей этого кластера, но присутствуют у более метаболически гибких видов. Эти протеазы могут способствовать деградации белковых компонентов тканей губок. У флавобактерий протеазы часто действуют синергетически с CAZymes при разрушении сложных органических матриц, что указывает на их взаимодополняющую роль в усвоении субстрата.

В совокупности, геномные особенности штамма *Flavobacterium* sp. PLB03 указывают на то, что это не классический патоген, а скорее экологически адаптированные бактерии с расширенной способностью к деградации сложных органических субстратов, в том числе животного и растительного происхождения. Сочетание богатого репертуара CAZymes и ПУЛ, уникальных локусов, ориентированных на гликаны хозяина, разнообразных протеаз и модульных геномных островов обеспечивает механистическую основу наблюдаемой способности *Flavobacterium* sp. PLB03 деградировать ткани губок и зеленых микроводорослей. Эти данные подтверждают интерпретацию того, что изолированный штамм *Flavobacterium* sp. PLB03 представляет собой экологически универсальный организм, способный к оппортунистическим взаимодействиям с организмами-хозяевами.

## 5. Выводы

Исследованный нами геном штамма *Flavobacterium* sp. PLB03 является новым видом и относится к группе свободноживущих (экологически распространенных) флавобактерий, с большим геномом. Расчеты средней идентичности нуклеоти-

дов со значениями 89,4–90,8%, значительно ниже пороговых значений, используемых для различения представителей близкородственных видов, подтверждают, что штамм *Flavobacterium* sp. PLB03 является уникальным видом. Штамм *Flavobacterium* sp. PLB03 не обладает типичными факторами вирулентности, но обладает широкими метаболическими возможностями для инвазии и разрушения тканей хозяина. Кроме того, *Flavobacterium* sp. PLB03, вероятно, является симбионтом губок, поскольку обладает уникальным набором специфических ферментов, расщепляющих гликаны хозяина, что необычно для рода *Flavobacterium*. Анализируя его способность расщеплять полисахариды, мы установили, что штамм *Flavobacterium* sp. PLB03 обладает значительным потенциалом в этой области. Однако необходимы дальнейшие исследования для определения конкретных полисахаридов, которые расщепляет этот штамм, что будет способствовать нашему пониманию экологической роли и биотехнологических применений штаммов флавобактерий.

## Благодарности

Исследование выполнено в рамках государственного задания Лимнологического института Сибирского отделения Российской академии наук № 0279-2021-0011. Материал подготовлен на основе работ, поддержанных проектом Российского научного фонда № 19-14-00088 и Российским фондом фундаментальных исследований № 16-04-00065; 16-54-150007; 18-04-00224.

## Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

## Список литературы

- Кожов М.М. 1972. Очерки по байкаловедению. Иркутск: Восточно-Сибирское книжное издательство.
- Ali Z., Cousin S., Fruhling A. et al. 2009. *Flavobacterium rivuli* sp. nov., *Flavobacterium sub-saxonicum* sp. nov., *Flavobacterium swingsii* sp. nov. and *Flavobacterium reichenbachii* sp. nov., isolated from hard water rivulet. *International Journal of Systematic and Evolutionary Microbiology* 59: 2610–2617.
- Andrews S. 2010. FastQC: A quality control tool for high throughput sequence data. *Babraham Bioinformatics*.
- Astashyn A., Tvedte E.S., Sweeney D. et al. 2024. Rapid and sensitive detection of genome contamination at scale with FCS-GX. *Genome Biology* 25: 1–25. DOI: [10.1186/S13059-024-03198-7/FIGURES/4](https://doi.org/10.1186/S13059-024-03198-7/FIGURES/4)
- Bankevich A., Nurk S., Antipov D. et al. 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology* 19(5):455–477. DOI: [10.1089/cmb.2012.0021](https://doi.org/10.1089/cmb.2012.0021)
- Belikov S., Belkova N., Butina T. et al. 2019. Diversity and shifts of the bacterial community associated with Baikal sponge mass mortalities. *PLOS One* 14: 3. DOI: [10.1371/journal.pone.0213926](https://doi.org/10.1371/journal.pone.0213926)
- Belikov S.I., Petrushin I.S., Chernogor L.I. 2021. Genome analysis of the *Janthinobacterium* sp. strain SLB01 from

the Diseased Sponge of the *Lubomirskia baikalensis*. *Current Issues in Molecular Biology* 43: 2220–2237. DOI: [10.3390/cimb43030156](https://doi.org/10.3390/cimb43030156)

Bertelli C., Laird M.R., Williams K.P. et al. 2017. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Research* 45: W30–W35. DOI: [10.1093/nar/gkx343](https://doi.org/10.1093/nar/gkx343)

Bolger A.M., Lohse M., Usadel B. 2014. Trimmomatic: A flexible read trimming tool for Illumina NGS data. *Bioinformatics* 30: 2114–2120.

Bormotov A.E. 2012. What has happened to Lake Baikal sponges? *Science first hand* 2: 32. URL: <https://scfh.ru/en/papers/what-has-happened-to-lake-baikal-sponges/>

Chen L., Zheng D., Liu B. et al. 2016. VFDB 2016: Hierarchical and refined dataset for big data analysis - 10 years on. *Nucleic Acids Research* 44: D694–D697. DOI: [10.1093/NAR/GKV1239](https://doi.org/10.1093/NAR/GKV1239)

Chen W.-M., Huang W.-C., Young C.-C. et al. 2012. *Flavobacterium tilapia* sp. nov., isolated from a freshwater pond, and emended descriptions of *Flavobacterium defluvii* and *Flavobacterium johnsoniae*. *International Journal of Systematic and Evolutionary Microbiology* 63: 827–834. DOI: [10.1099/ijs.0.041178-0](https://doi.org/10.1099/ijs.0.041178-0)

Chernogor L., Klimenko E., Khanaev I. et al. 2020. Microbiome analysis of healthy and diseased sponges *Lubomirskia baikalensis* by using cell cultures of primmorphs. *PeerJ* 8: e9080. DOI: [10.7717/peerj.9080](https://doi.org/10.7717/peerj.9080)

Chernogor L.I., Petrushin I.S., Khanaev I.V. et al. 2020. Study of changes in microbiomes of sponges *Lubomirskia baikalensis* using the cell culture of primmorphs. *Limnology and Freshwater Biology* 4: 647–648. DOI: [10.31951/2658-3518-2020-A-4-647](https://doi.org/10.31951/2658-3518-2020-A-4-647)

Denikina N.N., Dzyuba E.V., Belkova N.L. et al. 2016. The first case of disease of the sponge *Lubomirskia baikalensis*: investigation of its microbiome. *Translated in Biology Bulletin* 43: 263–270. DOI: [10.1134/S106235901603002X](https://doi.org/10.1134/S106235901603002X)

Ekwe A.P., Ahn J.-H., Kim S.B. 2017. *Flavobacterium keumense* sp. nov., isolated from freshwater. *International Journal of Systematic and Evolutionary Microbiology* 67: 2166–2170. DOI: [10.1099/ijsem.0.001920](https://doi.org/10.1099/ijsem.0.001920)

Galaxy platform for accessible, reproducible, and collaborative data analyses. 2024. *Nucleic Acids Research*. Oxford Academic, n.d. URL: <https://academic.oup.com/nar/article/52/W1/W83/7676834?login=false> (accessed: 11.16.25).

Jain C., Rodriguez-R L.M., Phillippy A.M. et al. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature communications* 9(1): 5114. DOI: [10.1038/S41467-018-07641-9](https://doi.org/10.1038/S41467-018-07641-9)

Khanaev I.V., Kravtsova L.S., Maikova O.O. et al. 2018. Current state of the sponge fauna (Porifera: Lubomirskiidae) of Lake Baikal: sponge disease and the problem of conservation of diversity. *Journal of Great Lakes Research* 44: 77–85. DOI: [10.1016/j.jglr.2017.10.004](https://doi.org/10.1016/j.jglr.2017.10.004)

Kozhov M. 2013. Lake Baikal and its life. Springer Science business Media. DOI: [10.1007/978-94-015-7388-7](https://doi.org/10.1007/978-94-015-7388-7)

Liu B., Zheng D., Zhou S. et al. 2022. VFDB 2022: A general classification scheme for bacterial virulence factors. *Nucleic Acids Research* 50: D912–D917. DOI: [10.1093/nar/gkab1107](https://doi.org/10.1093/nar/gkab1107)

Meier-Kolthoff J.P., Carbasse J.S., Peinado-Olarte R.L. et al. 2022. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Research* 50: D801–D807. DOI: [10.1093/NAR/GKAB902](https://doi.org/10.1093/NAR/GKAB902)

Meier-Kolthoff J.P., Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nature Communications* 10: 1–10. DOI: [10.1038/s41467-019-10210-3](https://doi.org/10.1038/s41467-019-10210-3)

Peeters K., Willems A. 2011. The gyrB gene is a useful phylogenetic marker for exploring the diversity of *Flavobacterium* strains isolated from terrestrial and aquatic habitats in Antarctica. *FEMS Microbiology Letters* 321: 130–140. DOI: [10.1111/J.1574-6968.2011.02326.X](https://doi.org/10.1111/J.1574-6968.2011.02326.X)

Petrushin I., Belikov S., Chernogor L. 2020. Cooperative interaction of *Janthinobacterium* sp. SLB01 and *Flavobacterium* sp. SLB02 in the diseased sponge *Lubomirskia baikalensis*. *International journal of molecular sciences* 21: 1–13. DOI: [10.3390/ijms21218128](https://doi.org/10.3390/ijms21218128)

Pile A.J., Patterson M.R., Savarese M. et al. 1997. Trophic effects of sponge feeding within Lake Baikal's littoral zone. 2. Sponge abundance, diet, feeding efficiency, and carbon flux. *Limnology and Oceanography* 42: 178–184. DOI: [10.4319/lo.1997.42.1.0178](https://doi.org/10.4319/lo.1997.42.1.0178)

Prijbelski A., Antipov D., Meleshko D. et al. 2020. Using SPAdesDeNovoAssembler. *Current Protocols in Bioinformatics* 70: e102. DOI: [10.1002/CPBI.102;PAGE=STRING:ARTICLE/CHAPTER](https://doi.org/10.1002/CPBI.102;PAGE=STRING:ARTICLE/CHAPTER)

Rawlings N.D., Barrett A.J., Thomas P.D. et al. 2018. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Research* 46: D624–D632. DOI: [10.1093/nar/gkx1134](https://doi.org/10.1093/nar/gkx1134)

Rawlings N.D., Bateman A. 2021. How to use the MEROPS database and website to help understand peptidase specificity. *Protein science: a publication of the Protein Society* 30: 83–92. DOI: [10.1002/pro.3948](https://doi.org/10.1002/pro.3948)

Rawlings N.D., Waller M., Barrett A.J. et al. 2014. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research* 42: D503–509. DOI: [10.1093/nar/gkt953](https://doi.org/10.1093/nar/gkt953)

Seemann T. 2016. ABRicate: mass screening of contigs for antibiotic resistance genes. URL: <https://github.com/tseemann/abricate>

Semiturkina N.A., Efremova S.M., Timoshkin O.A. 2009. State-of-the art of biodiversity and ecology of spongiofauna of Lake Baikal with special attention to the diversity, peculiarities of ecology and vertical distribution of Porifera on Berezovy ecology test site. In: Timoshkin O.A. (Ed.), *Index of animal species inhabiting Lake Baikal and its catchment area*. Novosibirsk, pp. 891–901. (in Russian)

Tatusova T., Dicuccio M., Badretdin A. et al. 2016. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Research* 44: 6614–6624. DOI: [10.1093/NAR/GKW569](https://doi.org/10.1093/NAR/GKW569)

Timoshkin O.A., Samsonov D.P., Yamamuro M. et al. 2016. Rapid ecological change in the coastal zone of Lake Baikal (East Siberia): Is the site of the world's greatest freshwater biodiversity in danger? *Journal of Great Lakes Research* 42: 487–497. DOI: [10.1016/j.jglr.2016.02.011](https://doi.org/10.1016/j.jglr.2016.02.011)

Yi H., Oh H.-M., Lee J.-H. et al. 2005. *Flavobacterium antarcticum* sp. nov., a novel psychrotolerant bacterium isolated from the Antarctic. *International Journal of Systematic and Evolutionary Microbiology* 55: 637–641. DOI: [10.1099/ijs.0.63423-0](https://doi.org/10.1099/ijs.0.63423-0)

Zheng J., Ge Q., Yan Y. et al. 2023. dbCAN3: automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Research* 51: W115–W121. DOI: [10.1093/nar/gkad328](https://doi.org/10.1093/nar/gkad328)