Iranian Journal of Microbiology

Volume 17 Number 6 (December 2025) 912-928
DOI: http://doi.org/10.18502/ijm.v17i6.20358

ORIGINAL ARTICLE

Core genome expansion in *Brevibacterium* across marine provinces reveals genomic footprint for long-term marine adaptation

Jagannath Sarkar*

Department of Biological Sciences, Bose Institute, Kolkata, India

Received: July 2025**, Accepted:** October 2025

ABSTRACT

Background and Objectives: Actinobacteria are ubiquitous across diverse environmental niches. *Brevibacterium* strains within this phylum are widely distributed in both marine and terrestrial ecosystems worldwide. Marine environments are defined by distinct physicochemical properties—high salinity, alkaline pH, fluctuating O_2 levels, and dynamic nutrient availability—which set them apart from terrestrial habitats. The broad ecological range of *Brevibacterium* strains raises questions about genome-encoded metabolic features that have evolved to adapt in marine environments.

Materials and Methods: Genomics of *Brevibacterium* strains from various marine provinces was analyzed, focusing on core genome and pan-genome structure.

Results: Core genome and pan-genome derived phylograms reveal a distinct polyphyletic origin of marine strains, as evidenced by their phylogenetic proximity despite diverse species affiliations. Only 1.16% of gene clusters from the total nonredundant gene repertoire were part of the core genome. Core genome size is shaped by geographical distribution. Notably, when strains from localized regions are analyzed, the core genome expands, indicating specialized functional requirements of additional genes within that environment. In marine isolates, the core genome includes genes involved in nutrient uptake, osmoregulation, and resistance to sediment genotoxicity. Additionally, a marine province-specific core genome analysis reveals genomic adaptations essential for acclimatization across different environments, regardless of species-level taxonomy.

Conclusion: Microbial genome evolution is shaped by ecological niche differentiation. The emergence and spread of habitats driven by tectonic plate movements may contribute to province-specific genomic divergence in *Brevibacterium*. This hypothesis merits further investigation, particularly as genomic data from deeper, geologically stable environments such as marine sediments become more accessible.

Keywords: Actinobacteria; Biological adaptation; Genome; Marine ecology; Horizontal gene transfer

INTRODUCTION

Actinobacteria are ubiquitous across a wide range of natural environments, regardless of the environmental harshness. Their habitat ranges from freshwater bodies to saline waterbodies and underlying sediment, hot springs, and soil of arid deserts (1-5). *Brevibacterium* can be regarded as an ideal genus within the phylum Actinobacteria due to its similarly wide habitat range (6-10). The genus includes strains that are obligately aerobic, Gram-positive, and non-spore-forming (11).

*Corresponding author: Jagannath Sarkar, Ph.D, Department of Biological Sciences, Bose Institute, Kolkata, India.

Current address: Materials Science and Environmental Engineering, Tampere University, Tampere, Finland. Tel: +91-8515999850

Fax: +91-8515999850 Email: jsarkarind@gmail.com

The global ocean and its sediment systems represent the largest interconnected microbial habitat (12). This vast oceanic territory is generally divided into four major provinces: Pacific Ocean, Atlantic Ocean, Indian Ocean, and Antarctic/Southern Ocean. Different physical and geochemical characteristics of an environment are the driving forces behind the selection of only those microorganisms possessing adaptive physiological traits necessary to survive in situ (13). In a particular environmental niche, ecological selective pressures influence genome evolution through different mechanisms of genome alteration, such as spontaneous mutation and the gain or loss of gene functions. Among these mechanisms, horizontal gene transfer, especially of genes encoding key survival features between taxonomically diverse individuals, is a highly effective means of genomic adaptation (14). *Brevibacterium* strains are ubiquitous in marine environments. They have been isolated from water and sediment samples collected in the Pacific (15, 16), Atlantic (16), and the Indian Ocean (6, 17).

This study aims to examine the genome-encoded adaptive features that play a pivotal role in the marine habitat-specific adaptation of strains within the *Brevibacterium* genus. The genomics of *Brevibacterium* strains has been studied elsewhere in relation to their adaptation to the cheese habitat (18, 19) and the exploration of their biosynthetic gene clusters (20); however, none of these studies have addressed the adaptive genomic features required for acclimatization to the marine environment. Here, a core- and pan-genome analysis was performed for two purposes: first, to trace the origin and relatedness of marine strains, and second, to identify how the varying physicochemical properties of distinct marine provinces influence genome evolution by analyzing subsets of genomes from each province. The expansion of the core genome for adaptation to specific marine provinces was examined by identifying genes that were added to the core genome only when genomes of strains from a particular province were considered. Thus, progressively narrow habitat-specific genome datasets were analyzed to address marine province-specific genome evolution. Furthermore, the genomic plasticity of marine isolates was assessed by identifying recently acquired genes through horizontal gene transfer in each genome. This study contributes to our understanding of how genome divergence is shaped by niche differentiation resulting from the diverse geophysical conditions of the Earth.

MATERIALS AND METHODS

Genomic data extraction. A total of 138 genome assemblies are available in the GenBank database of the National Center for Biotechnology Information (NCBI), under the genus *Brevibacterium* (as of October 2024). These include only genome assemblies derived from pure culture isolates that have proper annotation and valid taxonomic status, following NCBI's categorization, and that were retrieved in October 2024. The 138 genome assemblies include 25 complete genomes and 33 genomes from 'type material' (within the genus, 37 species have been validly published). These strains were found to be isolated from diverse habitats, including milk products, human infections, soil, freshwater, and marine sediment systems. The genomes of *Brevibacterium* sp. BDJS002 and *Brevibacterium* sp. JSBI002, which are represented by a single complete circular chromosome and were isolated from the Arabian Sea oxygen minimum zone sediment system (21), were also included among the 138 assemblies. TYGS (type strain genome server, DSMZ) was used to determine genome-based taxonomic position when needed, particularly for strains lacking species-level taxonomic status (22). According to TYGS, when the digital DNA-DNA hybridization (dDDH) value is lower than 70% against the closest type strain, the genome in the query is considered to belong to a new species (22).

Clustering of *Brevibacterium* strains based on marine provinces. Six hierarchical genome groups were formed based on the niche characteristics and geographical regions of the habitats from which the 138 *Brevibacterium* strains were isolated. The All_Genomes dataset includes all 138 genomes, which were isolated from diverse environmental habitats diverse environmental habitats, ranging from human infection to marine sediment. The Marine_Genomes dataset includes the genomes of strains isolated from seawater or marine sediment (genomes obtained from strains present in seafood were excluded, as the geographic origin of the seafood may not represent the actual geographic location of the natural habitat). The Atlantic_Ocean_Genomes, Pacific_Ocean_Genomes, and Indian_Ocean_Genomes datasets include

three, four, and six genomes, respectively, isolated from marine water or sediment systems of the corresponding oceans. The Arabian_Sea_Genomes dataset includes genomes obtained from two pure culture strains isolated from the OMZ sediment of the Arabian Sea.

Analysis of core genome and pan-genome with phylogeny. The Bacterial Pan-Genome Analysis pipeline (BPGA, version 1.3; 23) was used to determine the pan-genome and core genome of six hierarchical habitat-based groups of strains classified under the genus *Brevibacterium*. For orthologous gene clustering, the USEARCH algorithm (24) was employed, using a 50% identity cutoff to cluster all genes present in the genomes of each dataset. Each of the six genome groups was analyzed separately to determine the respective pan-genome and core genome. The pan-genome describes the distribution of all genes in a dataset across different orthologous gene families and includes core genes (present in every genome of the dataset), accessory genes (present in two or more genomes), and unique genes (present in only one genome). Non-redundant reference sequences (RefSeq) of core, accessory, and unique genes from each genomic dataset were extracted using the BPGA script. After orthologous clustering of all genes across the genomes in a dataset, pan-genome and core genome curves were plotted. The pan-genome curve depicts the number of new gene families added with each additional genome, while the core genome curve illustrates how the number of different gene families decreases as more genomes are added. Using BPGA scripts, 20 random permutations were carried out for the addition of each genome. The median values of the number of all distinct gene families and the number of commonly shared gene families in a given dataset were considered to determine the pan-genome and core genome, thereby minimizing bias. A power regression model was used for pan-genome data, and an exponential model was used for core genome data to plot pan-genome and core genome curves and to determine whether the pan-genome is closed or open (23). All genes in a genome that contain atypical G+C content (G+C content of a gene deviates from the average G+C content of the genome by more than twice the standard deviation) were identified using the atypical GC content analysis script of BPGA (23). Such atypical GC content in a gene indicates possible inclusion in the respective genome through horizontal gene transfer.

Phylogenetic analysis was performed by considering both the core genome and pan-genome of all 138 strains. For core genome-based phylogenetic analysis, core genes were used for multiple sequences alignments using MUSCLE software (25). For pan-genome-based phylogenetic analysis, a binary pan-matrix based on gene presence and absence for all 138 genomes was considered. A distance matrix was then calculated from this pan-matrix based on the similarity and dissimilarity of gene content. Phylogenetic trees for both the core genome and pan-genome were constructed using the neighbor-joining approach in Bio-Phylo-v2.0.1 (26). Average nucleotide identity (ANI) for groups of genomes was determined using fastANI (27).

Annotation of core, accessory, and unique genes. Web-based utilities of eggNOG-mapper were used for functional identification of core, accessory, and unique genes in the genomic dataset when required (28). For this, the eggNOG 5 database was searched for a given query using the following search filters: minimum hit e-value of 0.001, minimum hit bit-score of 60, percentage identity of 40, minimum query coverage of 20%, and minimum subject coverage of 20%.

RESULTS

General genomic characteristics. Genome assembly and metadata of 138 *Brevibacterium* strains (all obtained from pure cultures) were retrieved from the GenBank database of NCBI. These strains were isolated from diverse environmental habitats, including human infections, dairy products, soil, freshwater, and marine sediments (Table 1). Among the 138 strains, 43 were isolated from samples originating from human infections and 40 were isolated from milk-based products.

Thirteen *Brevibacterium* strains were isolated from marine environments and included accurate geographic sampling locations in their metadata. These marine strains were distributed across three marine provinces: 3 were isolated from the Atlantic Ocean, 6 from the Indian Ocean (including 2 from Arabian Sea OMZ sediment), and 4 from the Pacific Ocean (Fig. 1A). Of these 13 strains, 9 were previously reported and taxonomically classified to the species

Table 1. Genomic features of all studied marine isolates under the genus *Brevibacterium*

Isolates name	Assembly accession number	Isolated from	Geographic location	Size (Mb)	Type strain status	TYGS result	GC %	Complete status (circular or not)	CheckM Completeness (%)
Brevibacterium atlanticum WO024	GCA_011617245.1	Seawater	Atlantic Ocean	4.2	Type	Not required	66	Complete	94.5
Brevibacterium pigmentatum YB235	GCA_011617465.1	Sediment	Atlantic Ocean	3.9	Type	Not required	65	Complete	95.0
Brevibacterium oceani WW007	GCA_017948325.1	Sediment	Atlantic Ocean	4.3	-	Not required	66	Complete	96.0
Brevibacterium sp. BDJS002	GCA_028201555.1	Sediment	Arabian Sea; Indian Ocean	4.2	-	Potential new species	63	Complete	95.1
Brevibacterium sp. JSBI002	GCA_026013965.1	Sediment	Arabian Sea; Indian Ocean	3.7	-	Potential new species	65	Complete	95.2
Brevibacterium oceani BBH7	GCA_013623835.1	Sediment	Indian Ocean	4.5	Type	Not required	66	Draft	96.0
Brevibacterium sediminis FXJ8.269	GCA_013623905.1	Sediment	Indian Ocean	4.2	Type	Not required	65	Draft	97.2
Brevibacterium sediminis CGMCC 1.15472	GCA_014643055.1	Sediment	Indian Ocean	4.2	-	Not required	65	Draft	96.8
Brevibacterium sediminis COD27	GCA_024895295.1	Sediment	Indian Ocean	4.3	-	Not required	65	Draft	97.2
Brevibacterium limosum o2	GCA_011617705.1	Sediment	Pacific Ocean	4.3	Type	Not required	65	Complete	96.2
Brevibacterium marinum DSM 18964	GCA_011927955.1	Seawater	Hwasun Beach; Pacific Ocean	4.2	Type	Not required	65	Draft	88.3
Brevibacterium sp. CCUG 69071	GCA_021023135.1	Sediment	Comau fjord; Pacific Ocean	4.1	-	Potential new species	66	Draft	95.1
Brevibacterium sp. Marine	GCA_012844365.1	Sediment	Pacific Ocean	4.2	-	Potential new species	65	Complete	97.7

level. To determine the species-level taxonomic positions of the remaining 4 strains, digital DNA-DNA hybridization (dDDH) was performed by comparing each genome against all available type strain genomes using the TYGS web server. Each of these strains was found to represent a new species, as the dDDH value with the nearest type strain was below 70% (Table 1).

The genome size of all 138 strains in the All_Genomes dataset ranges from 2.3 Mb to 4.7 Mb, with CheckM-calculated completeness values ranging from 75.5% to 98.7%. The G+C content ranges from 58% to 73%. For the 13 strains in the Marine_Genomes dataset, genome sizes range from 3.7 Mb to 4.5 Mb, CheckM-calculated completeness ranges from 88.3% to 97.7%, and G+C content ranges from 63% to 66% (Figs. 1B, C and Table 1). In this study, no filtration based on CheckM completeness was applied because the genome of *Brevibacterium* sp. CS2 (NCBI assembly accession number: GCA_005280295.1), despite being assembled into a single contig, showed 89.9% completeness according to the CheckM algorithm.

Structural and phylogenetic analysis of the core genome and pan-genome reveals distinct marine clade. The total gene repertoire from the genomes of 138 *Brevibacterium* strains was used for orthologous clustering through the Bacterial Pan Genome Analysis (BPGA) pipeline. It was observed that 325 non-redundant RefSeq genes are present in each strain and can be regarded as the core genome. Additionally, 27914 non-redundant RefSeq genes were identified, encompassing the entire gene repertoire of the 138 genomes, thus constituting the pan-genome.

To determine whether the pan-genome is open or

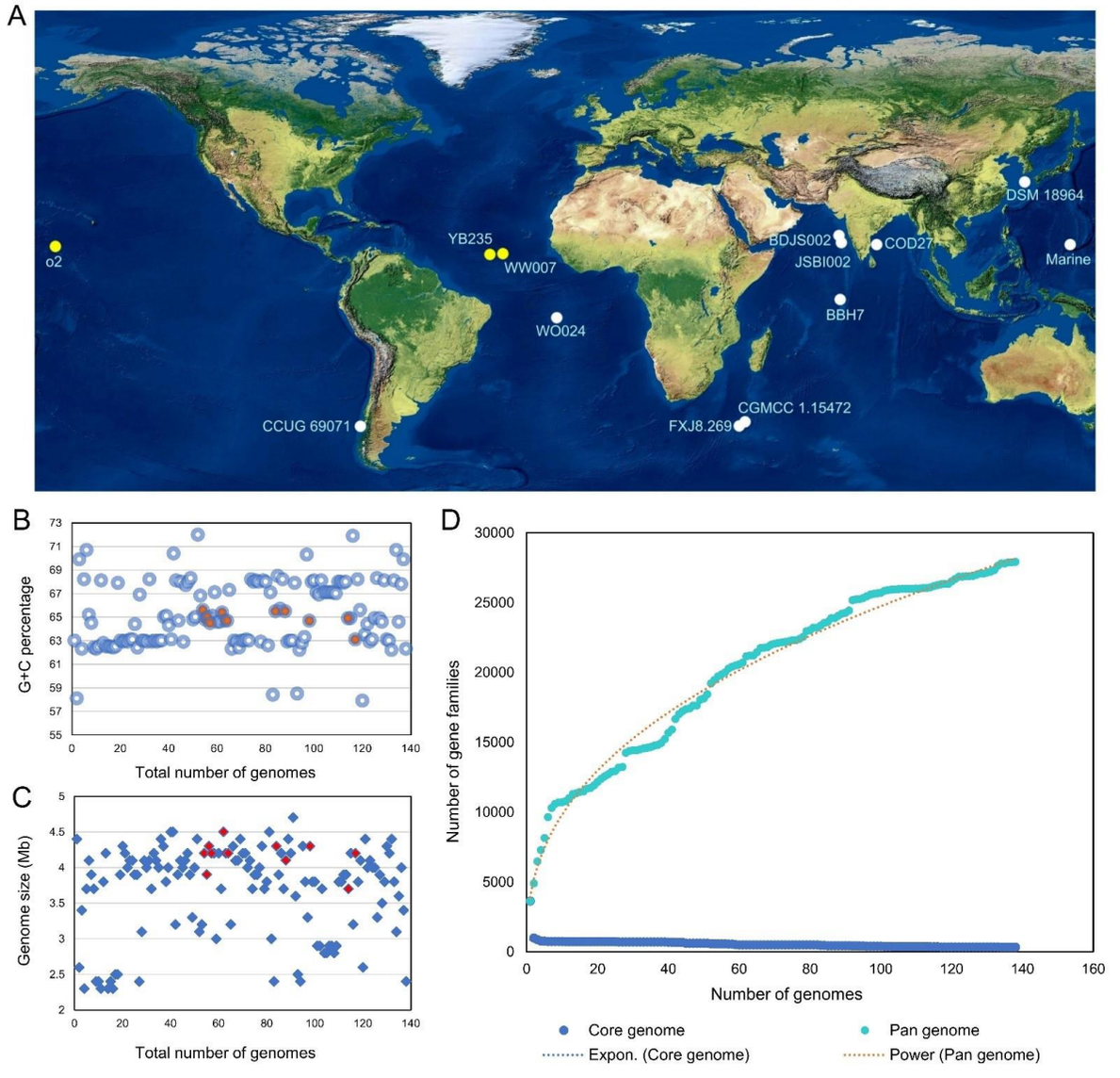


Fig. 1. (A) Approximate geographical locations from which sampling was reported for the isolation of 13 marine pure culture strains of *Brevibacterium*. White dots indicate stains with precise geographical locations for the isolation inoculum, while yellow dots represent strains with only broadly reported geographical provinces (i.e., Atlantic Ocean, Indian Ocean, and Pacific Ocean). (B) Distribution of G+C percentages across all 138 genomes of pure culture strains of *Brevibacterium*. (C) Genomic size distribution across all 138 genomes of pure culture strains of *Brevibacterium*. In both plots, red-filled shapes denote genomic data derived from marine strains. (D) Core and pan-genome plot of the 138 genomes from pure culture strains of *Brevibacterium*, showing genome-by-genome addition and corresponding changes in core and pan-genome structure.

closed, pan-genome data were plotted on a graph where the x-axis represents the sequential addition of genomes, and the y-axis represents the number of new gene families added to the pan-genome. When the graph was fitted with a power regression model [equation: $f(x) = a \cdot x \wedge b$], the values were found to be a = 3674.6 and b = 0.4 (Fig. 1D). This indicates that the overall pan-genome is still open (as the value of $b < 1$), suggesting that genes from new orthologous families continue to be added to the gene pool of *Brevibacterium* strains.

A phylogenetic tree was constructed based on the multiple sequence alignment of the 325 core genes (Fig. 2). In the phylogram, different species-level clusters were evident at various nodes. Twelve out of the thirteen marine strains were positioned adjacently

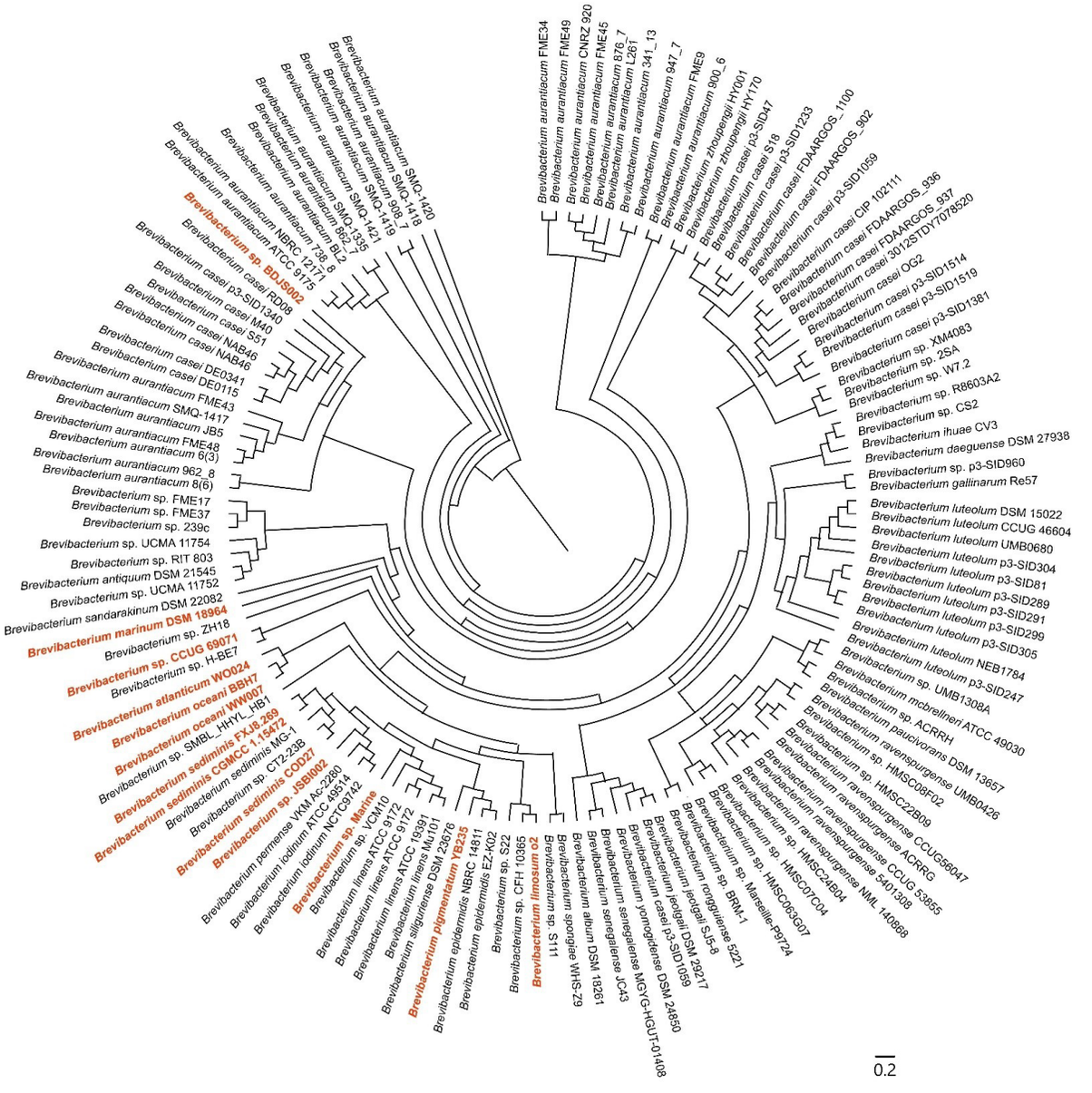


Fig. 2. Phylogenetic tree based on 325 core genes from 138 strains of *Brevibacterium*, constructed using neighbor-joining approach. Strain names in red indicate pure culture isolates obtained from marine seawater or sediment, with accurately reported geographical isolation sites.

on the phylogenetic tree, while only *Brevibacterium* sp. BDJS002 was positioned distantly. A pan-genome-based phylogenetic tree was also constructed, based on each genome's contribution to the pan-genome structure through a gene presence-absence data matrix (data not shown). Notably, *Brevibacterium* sp. BDJS002 also occupied a distant position in this tree.

To determine the distribution of different components of the pan-genome, the non-redundant gene sets of core, accessory (genes present in two or more strains), and unique genes (genes present in a single strain) were analyzed for their distribution across various COG categories. RefSeq gene sequences for all these categories were determined using BPGA, as described earlier. The highest percentage of non-redundant core genes was distributed in the J (Translation, ribosomal structure, and biogenesis) COG category. The majority of non-redundant accessory genes were distributed in the K (Transcription) category. Similar-

ly, the highest number of non-redundant unique genes across 138 genomes was also found in the K (Transcription) category. For both accessory and unique genes, the second-highest distribution was observed in the E (Amino acid transport and metabolism) category. The R category was excluded from this ranking, as it includes genes predicted to be involved in general functions.

Additionally, core genes were found to be distributed with values exceeding 5% in the M (Cell wall/membrane/envelope biogenesis), O (Post-translational modification, protein turnover, and chaperones), L (Replication, recombination, and repair), C (Energy production and conversion), E (Amino acid transport and metabolism), H (Coenzyme transport and metabolism), and I (Lipid transport and metabolism) COG categories. More than 5% of accessory genes were distributed in the L (Replication, recombination, and repair), G (Carbohydrate transport and metabolism), and P (Inorganic ion transport and metabolism) COG categories. More than 5% of unique genes were individually distributed in the G (Carbohydrate transport and metabolism), I (Lipid transport and metabolism), Q (Secondary metabolites biosynthesis, transport, and catabolism), and P (Inorganic ion transport and metabolism) COG categories (Fig. 3A).

ANI analysis corroborates the adjacent phylogenetic position of marine strains. A heatmap with a dendrogram was generated from the ANI data obtained through an all-vs-all genome ANI analysis, including 13 genomes of marine *Brevibacterium* strains and the genomes of type strains of all 33 species. All the marine strains clustered closely together in the dendrogram, except for *Brevibacterium* sp. BDJS002 (Fig. 4). The central yellow-red zone in Fig. 4 distinctly indicates a high level of genomic relatedness (> 90% ANI value), predominantly represented by marine strains.

Expansion of core genome for adaptation in marine habitat. To examine the core genome and pan-genome structure of 13 marine strains, orthologous clustering of the total genomic repertoire was performed using the BPGA pipeline. It was observed that 1963 non-redundant RefSeq genes are present in each strain and can be regarded as the core genome. Additionally, 7788 non-redundant RefSeq genes were identified, encompassing the entire gene repertoire of the 13 marine strains, thus constituting the pan-genome.

To evaluate whether the pan-genome is open or closed, the pan-genome data were plotted on a graph as described previously. When the graph was fitted with a power regression model [equation: $f(x) = a \cdot x \wedge b$], the values were found to be $a = 3456.8$ and $b = 0.3$ (data not shown). This reveals that the overall pan-genome is still open (as the value of $b \leq 1$), indicating that genes from new orthologous families are still being added to the gene pool of marine *Brevibacterium* strains.

Different COG categories were found to be enriched in the core, accessory, and unique gene sets of the pan-genome. The highest percentage (13.5%) of core genes was distributed in the E (Amino acid transport and metabolism) COG category. The majority of accessory genes were predominantly distributed in the K (Transcription, 13.8%) and E (Amino acid transport and metabolism, 13.1%) categories. For unique genes,

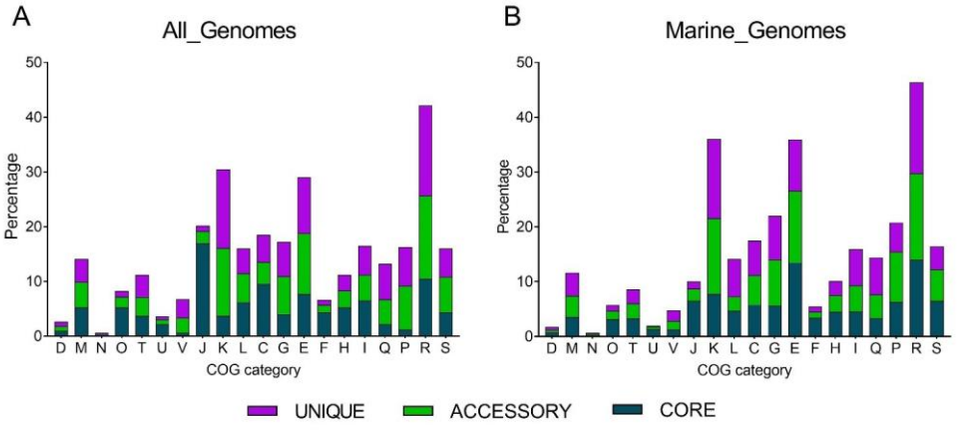


Fig. 3. (A) Distribution of core, accessory and unique genes from 138 genomes across different COG categories (B) Distribution of core, accessory and unique genes from 13 genomes of marine isolates across different COG categories

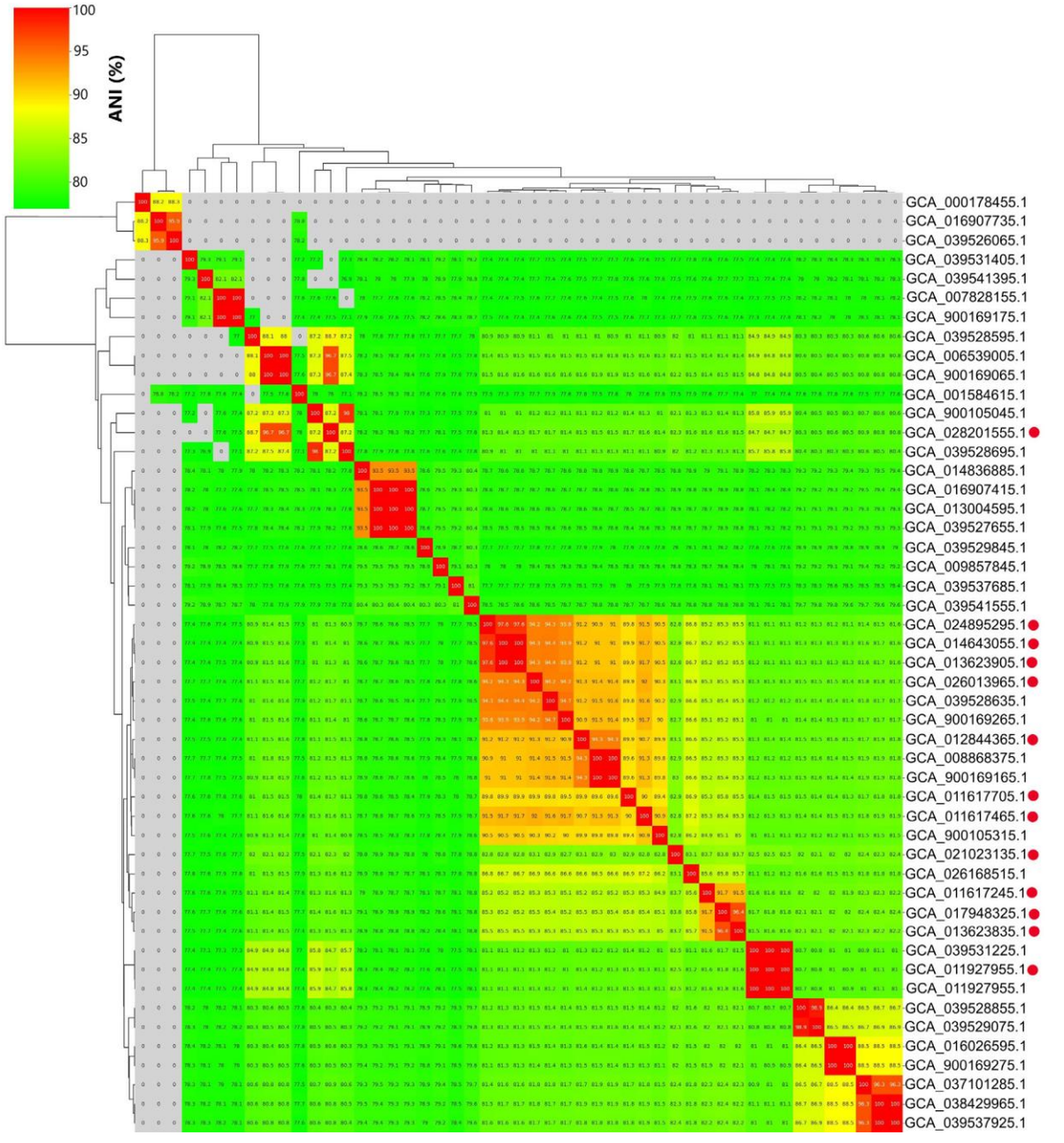


Fig. 4. All-vs-all ANI analysis including genome of type strains and marine isolates. Red dot on the right side of genome accession number signifies that the strain was isolated/found in marine environment.

the K (Transcription) category exhibited the highest distribution. Additionally, more than 5% of the core, accessory, and unique genes was individually distributed in K (Transcription), C (Energy production and conversion), G (Carbohydrate transport and metabolism), E (Amino acid transport and metabolism) and P (Inorganic ion transport and metabolism) COG categories (Fig. 3B).

After annotating all 1963 core genes against the KEGG database, it was found that several KEGG sub-categories were enriched in the core genome of marine strains. These sub-categories include Membrane transport and Signal transduction (under the major KEGG category of Environmental Information Processing), Folding, sorting and degradation, Replication and repair, Translation (under the major KEGG category of Genetic Information Processing), Amino acid metabolism, Carbohydrate

metabolism, Energy metabolism, and Nucleotide metabolism (under the major KEGG category of Metabolisms).

The core genome contributing to marine adaptation can be identified by subtracting the 325 non-redundant core genes common to all strains from the non-redundant core genes of marine strains. Thus, the 1,638 core genes resulting from this subtraction (1,963 - 325) are potential candidates essential for marine adaptation. When annotated against the KEGG database, these genes were found to be distributed across all COG categories. The important genes within each COG category, along with the functional descriptions of their encoded proteins, reveal diverse biological roles. Under COG category D (cell cycle control, cell division, chromosome partitioning), genes such as *ftsK*, *ftsX*, *ftsW*, *parA*, *scpB*, *sepF*, and *smc* encode proteins involved in cell division and chromosome segregation. In COG category M (cell wall/membrane/envelope biogenesis), genes including *mraY*, *murB*, *murD*, *murG*, and *murI* are involved in cell wall formation, while *mscS* and *mscL* regulate mechanosensitive ion channels and osmotic pressure within the cell. COG category O (post-translational modification, protein turnover, and chaperones) includes *ftsH* for quality control of an integral membrane protein, *groL* for proper folding of polypeptides under stress, and *grpE* for preventing aggregation of stress-denatured proteins during hyperosmotic and heat shock conditions. In COG category T (signal transduction mechanisms), genes such as *glnE*, *relA*, *citA*, *pknA*, and *cseB* are involved in ammonia assimilation, stringent response coordination in response to changes in nutritional abundance, and kinase-mediated transcriptional regulation. COG category U (intracellular trafficking, secretion, and vesicular transport) includes components of the Sec protein translocase complex (*secA*, *secD*, *secE*), type II/IV secretion system protein (*cpaF*), and membrane targeting protein (*ffh*). COG category V (defense mechanisms) features antibiotic resistance genes such as an uncharacterized beta-lactamase, and proteins from the AcrB/AcrD/AcrF family. In COG category J (translation, ribosomal structure and biogenesis), genes encode ribosomal proteins (*rpsA* to *rpsD*, *rpsF* to *rpsH*, *rpsJ*, *rpsK*, *rpsM* to *rpsT*, *rplA*, *rplC*, *rplD*, *rplF*, *rplI* to *rplQ*, *rplS*, *rplU* to *rplX*) and translational termination proteins (prfA to prfC). COG category K (transcription) includes genes such as *sigA* (primary sigma factor responsible for sigmoidal growth), *argR* (arginine biosynthesis regulator), *hspR* (regulatory protein), *scoF4* (cold shock protein), and *rpoZ* (a protein needed for assembly of RNA polymerase). In COG category L (replication, recombination and repair), genes encode a DNA alkylation repair enzyme (*alkB*, *alkD*), DNA polymerase and associated proteins (*dnaA*, *dnaB*, *dnaE2*, *dnaG*, *dnaN*, *dnaQ*, *dnaX*, *polA*), recombination-related proteins (*recB*, *recF*, *recG*, *recN*, *recQ*, *recR*), and DNA damage recognition and processing protein machinery (*uvrA*, *uvrD2*, *uvrC*). COG category C (energy production and conversion) includes genes for pyruvate dehydrogenase components (*aceA*, *aceE*), an arginine biosynthesis regulator (*argR*), ATP synthase subunits (*atpA*, *atpB*, *atpC*, *atpE*, *atpF*, *atpG*), and a sodium dicarboxylate symporter (*gltT*). In COG category G (carbohydrate transport and metabolism), genes such as *manB*, *manA*, *otsB*, *otsA*, *pfkA* and *smoE* encode enzymes for phosphoglucomutase/phosphomannomutase, phosphomannose isomerase type I, glycosyl hydrolases, glycosyltransferase, phosphofructokinase, and bacterial extracellular solute-binding protein. COG category E (amino acid transport and metabolism) includes genes for arginine synthase and lyase family proteins (*argB*, *argC*, *argE*, *argF*, *argG*, *argH*, *argJ*), shikimate pathway enzymes (*aroA*, *aroB*, *aroE*, *aroK*, *aroQ*), and tryptophan biosynthesis proteins (*trpA*, *trpB*, *trpC*). In COG category F (nucleotide transport and metabolism), genes encode *add* (adenosine/AMP deaminase), *carA* and *carB* (carbamoyl-phosphate synthase), *pyrF* (orotidine 5'-phosphate decarboxylase), *pyrH* (reversible phosphorylation of UMP to UDP), *xdhA* (CO dehydrogenase), *xdhB* (aldehyde oxidase and xanthine dehydrogenase), and *guaA* (GMP synthase). Under COG category H (coenzyme transport and metabolism) includes genes such as *coaE* (phosphorylation of the 3'-hydroxyl group of dephosphocoenzyme A to form coenzyme A), *coaD* (adenylyl group transfer), *hemA* (NADPH-dependent reduction of glutamyl-tRNA), and various coenzyme biosynthesis genes (*nadA* to *nadE*, *thiD* to *thiG*, *thiL*, *thiM*). In COG category I (lipid transport and metabolism), genes encode transferases (*accD1*, *acpS*, *acpP*), AMP-binding enzyme (*fadD*), thiolase (*fadI*), and acyl transferase (*fabD*). Finally, COG category P (inorganic ion transport and metabolism) includes genes encoding, *katA* (catalase), *katG* (bifunctional catalase/peroxidase), *kefB* (sodium/hydrogen exchanger), *mrpA/mrpB* (NADH-Ubiquinone oxidore-

ductase complex I), *mrpC* (NADH-ubiquinone/plastoquinone oxidoreductase), *mrpD* (proton-conducting membrane transporter), *mrpE* and *mrpG* (Na+/H+ ion antiporter subunits), oligopeptide transport system components (*oppB*, *oppC*, *oppC4*), and *soda*, which detoxifies radicals harmful to cellular systems.

Marine province-specific core genome expansion. Core genome analysis was performed for three individual marine province-specific genomic datasets. It was observed that core genome was represented by 2572, 2376, and 2243 non-redundant core genes for Atlantic_Ocean_Genomes, Pacific_Ocean_Genomes, and Indian_Ocean_Genomes. The pan-genome was represented by 4711, 5780, and 5462 number of non-redundant genes, respectively. The pan-genome of each dataset was found to be open. Thus, the core genome of *Brevibacterium* strains living in Atlantic Ocean, Indian Ocean and Pacific Ocean includes 609, 413 and 280 additional core genes respectively, in addition to the core genome of 13 marine strains (Fig. 5A). Annotation of these additional core genes with eggNOG 5 database revealed that they are distributed across most of the COG categories (Fig. 5B).

The Atlantic, Indian, and Pacific Oceans exhibit overall similar physicochemical properties, such as temperature, salinity, and nutrient profiles, which may influence the distribution and adaptation of microbial communities within these water bodies. Despite this, the strains reported from the Indian Ocean were found to inhabit regions within or near the oxygen minimum zone. COG classifications of the core genes added to the core genome of strains from the Indian Ocean are listed in Table 2.

The two *Brevibacterium* strains isolated from of the Arabian Sea OMZ share an additional 208 core genes, in addition to the shared core genome of the Indian Ocean isolates. These 208 genes are distributed across most of the COG categories.

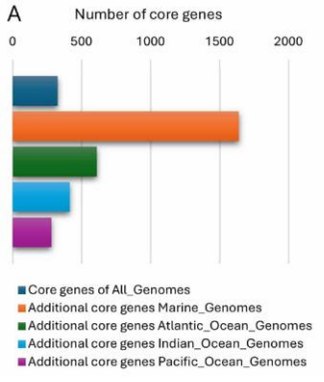
Active genome evolution in marine. The G+C content of genes inherited vertically from parent organisms generally does not deviate appreciably, unlike genes acquired through horizontal gene transfer (HGT) mechanisms (29, 30). When a gene's G+C content deviates by more than two standard deviations from the mean (atypical G+C content), it is likely the result of a previous HGT event. Genes with G+C content deviations exceeding this threshold were identified using BPGA scripts (23).

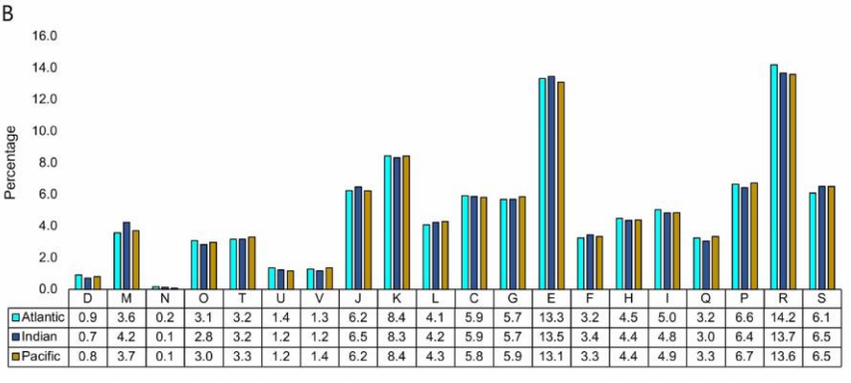
Analysis of the 13 genomes of marine *Brevibacterium* strains revealed that each genome contains several unique genes with atypical G+C content. Annotating these genes using the eggNOG mapper against the KEGG database indicated that they belong to different orthologous groups (eggNOG_OGs).

Table 3 lists the detailed number of unique genes, along with their origins and functions. The source organisms for these potential HGT genes include both closely related and distant lineages.

DISCUSSION

Microbial metabolism in marine environments is influenced by the physicochemical features of the habitat; the diversity of metabolisms differs between





	D	M	N	O	T	U	V	J	K	L	C	G	E	F	H	I	Q	P	R	S
Atlantic	0.9	3.6	0.2	3.1	3.2	1.4	1.3	6.2	8.4	4.1	5.9	5.7	13.3	3.2	4.5	5.0	3.2	6.6	14.2	6.1
Indian	0.7	4.2	0.1	2.8	3.2	1.2	1.2	6.5	8.3	4.2	5.9	5.7	13.5	3.4	4.4	4.8	3.0	6.4	13.7	6.5
Pacific	0.8	3.7	0.1	3.0	3.3	1.2	1.4	6.2	8.4	4.3	5.8	5.9	13.1	3.3	4.4	4.9	3.3	6.7	13.6	6.5

Fig. 5. (A) Increase in the number of core genes when considering only marine strains and those from individual geographical provinces (B) Distribution of core genes from strains inhabiting the Atlantic, Indian, and Pacific Ocean provinces, categorized by COG functional groups. Distribution of core genes across COG categories differs more in detail

Table 2. COG category-wise important genes and their functional descriptions, which were added to the core genome of *Brevibacterium* strains living in the Indian Ocean province

COG Category	Associated Genes / Functions
M – Cell wall/membrane/envelope biogenesis	*murF, ispH, wecB, capD, dapA_1* - involved in cell wall formation, peptide synthesis, and CoA-binding
T – Signal transduction mechanisms	*pdtaR,* forkhead associated domain - transcriptional antitermination
U – Intracellular trafficking, secretion	*secF, secG, FimV* - membrane protein translocation and TonB-independent uptake
J – Translation, ribosomal structure, biogenesis	*pnp, rplB, rplT, rpsL, rpsI* - ribosomal assembly and mRNA degradation *nusG,*
K – Transcription	*nusB, higA* - includes transcription regulators and antidote proteins
L – Replication, recombination and repair	*topA, gyrB2, sdrA, holB, ku, recO, orn, ruvA* - topoisomerases, polymerases, DNA repair enzymes
C – Energy production and conversion	*cat1, atpD, sucC, ctaF* - enzymes in ATP production and citric acid cycle
G – Carbohydrate transport and metabolism	*gntT, eno, nagB* - hydrolases, permeases, sugar metabolism enzymes *aroG,*
E – Amino acid transport and metabolism	*aroD, arr, hisE, hisH* - involved in aromatic amino acid biosynthesis *pyrG,*
F – Nucleotide transport and metabolism	*pyrD, purA, apt* - purine/pyrimidine biosynthesis enzymes
H – Coenzyme metabolism	*coaBC, bioB, apbA, hemH, pncB* - coenzyme A, biotin, and NADPH-related pathways
I – Lipid transport and metabolism	*atoE, ispF–ispH, crtB* - fatty acid and isoprenoid compound synthesis
P – Inorganic ion transport and metabolism	*amt, focA, corA, hmuO* - transporters for citrate, ammonium, formate, and magnesium

coastal seawater-sediment systems and open ocean seawater-sediment systems (31-33). Seawater and the underlying sediment of a particular marine environment function as an environmental continuum, where microbial metabolism within the sediment system is influenced by the overlying seawater, which serves as the primary source of nutrient supply (34). Depending on physicochemical features such as sedimentation rate, bottom water oxygen availability, and the abundance of oxidized/reduced metallic ions, different metabolic strata exist in the marine sedimentary environment (32, 35, 36). Although microbial communities vary from coastal regions to open ocean sites, aerobic strains of Proteobacteria and Actinobacteria are found to be ubiquitously distributed throughout the global marine habitat (21, 37, 38). Despite the extreme environmental conditions in different marine provinces, the widespread distribution of aerobic bacteria across the marine system raises questions about their adaptation. Genomic adaptation to marine habitats was studied in strains of the aerobic genus *Brevibacterium*, revealing its marine origin and expanded core genome structure that supports adaptation to different marine provinces.

Polyphyletic genome evolution for marine strains of *Brevibacterium*. G+C content percentage and genomic size are distributed within a narrow range when only marine strains are considered (Figs. 1B and C), in comparison to all 138 strains. This indicates that, structurally, genomes of the marine strains remain consistent. After orthologous clustering of all genes from the 138 genomes, 325 non-redundant genes were identified, representing the core genome. A major proportion of these 325 core genes was assigned to translation, ribosomal structure, and biogenesis metabolism, indicating their essential and central role in microbial survival and growth. However, since all strains under the genus are Gram-positive and capable of adapting to diverse environmental habitats, a significant number of genes related to cell wall formation, replication, energy production and conversion, amino acid transport, and metabolism become conserved. The pan-genome remains open, indicating the adaptive potential of strains in adverse environmental conditions.

When these 325 core genes were used for multiple sequence alignment, and a phylogenetic tree was constructed using a neighbor-joining approach, it was found that marine strains occupy adjacent positions on the phylogram (Fig. 2). A similar observation was made when a gene presence-absence-based phylogenetic tree was constructed using the same approach, based on each genome's contribution to the pan-genome (data not shown). The only strain that appeared distant in position was *Brevibacterium* sp. BDJS002. Furthermore, an alignment-free

IRAN. J. MICROBIOL. Volume 17 Number 6 (December 2025) 912-928

http://ijm.tums.ac.ir

Table 3. Unique genes, their origins, and functions that are inherited in marine strains through horizontal gene transfer mechanism

Strain name	Number of unique genes (Atypical G+C content)	Orthologous groups	Functions
Brevibacterium atlanticum WO024	33	Bacteria (4), Actinobacteria (24, 1 from Brevibacteriaceae), Proteobacteria (3), Firmicutes (2)	Transcription regulation (carD, ybcM), putative sugar transporter, mercury resistance proteins, type I restriction modification DNA specificity domain (hsdM)
Brevibacterium pigmentatum YB235	16	Bacteria (1), Actinobacteria (11, 1 from Brevibacteriaceae), Proteobacteria (3), Cyanobacteria (1)	Response regulator, adenine-specific DNA methylase, type III restriction enzyme (res subunit), ATPase, P-type transporter (cybH)
Brevibacterium oceani WW007	39	Bacteria (4), Actinobacteria (32, 1 from Brevibacteriaceae), Proteobacteria (1), Firmicutes (2)	Transcriptional regulator, ABC transporter, bacterial extracellular solute-binding protein, sugar transporter, ABC-type nitrate/sulfonate/bicarbonate transport systems
Brevibacterium sp. BDJS002	29	Bacteria (4), Actinobacteria (24, 9 from Brevibacteriaceae), Bacteroidetes (1)	Restriction endonuclease, death-on-curing family protein, resolvase, cytochrome c biogenesis protein transmembrane region, region found in RelA/SpoT proteins (ywaC)
Brevibacterium sp. JSBI002	21	Actinobacteria (19, none from Brevibacteriaceae), Proteobacteria (1), Firmicutes (1)	Glycosyltransferase, endonuclease, transposase, putative ATP-dependent Lon protease, peptidase S8, bacterial toxin 35
Brevibacterium oceani BBH7	23	Viruses (4), Bacteria (2), Actinobacteria (15, none from Brevibacteriaceae), Firmicutes (1), Bacteroidetes (1)	Transcriptional regulator, DNA alkylation repair, endonuclease, viral recombinase domain
Brevibacterium sediminis COD27	31	Bacteria (2), Actinobacteria (22, 6 from Brevibacteriaceae), Proteobacteria (4), Bacteroidetes (1), Firmicutes (2)	Amidinotransferase, transposase, nucleotidyl transferase AbiEii toxin, arsenite transmembrane transporter, Na+/H+ antiporter
Brevibacterium limosum o2	34	Bacteria (4), Actinobacteria (30, 3 from Brevibacteriaceae)	Glycosyltransferases (families 1, 2, 4), sigma-70 region 2, sugar transporter
Brevibacterium marinum DSM 18964	30	Bacteria (6), Actinobacteria (23, 1 from Brevibacteriaceae), Firmicutes (1)	Transposases, proteins involved in the utilization of glycolate and propanediol, hydrolase activities
Brevibacterium sp. CCUG 69071	30	Viruses (1), Actinobacteria (28, 1 from Brevibacteriaceae), Firmicutes (1)	Glycosyltransferase, transcriptional regulator, endonucleases, recombinase
Brevibacterium sp. Marine	37	Bacteria (3), Archaea (1), Actinobacteria (26, 5 from Brevibacteriaceae), Proteobacteria (4), Firmicutes (3)	Glycosyltransferases, aromatic amino acid lyase, reverse transcriptase, sodium solute symporter, DNA mismatch endonuclease (vsr)
Brevibacterium sediminis FXJ8.269 & *Brevibacterium sediminis* CGMCC 1.15472	0	None Identified	N/A

ANI analysis-based dendrogram also corroborated these findings (Fig. 4), indicating that the genomic relatedness of marine strains is influenced not only by coding regions but also by non-coding regions. The polyphyletic yet close phylogenetic relationship among marine strains suggests that acclimatization to marine environments is not a straightforward process for microbes introduced stochastically from terrestrial habitats; rather, marine strains have not diverged recently but have evolved through continuous adaptation to marine conditions.

Genomics of nutrient uptake, osmotic pressure regulation, and resistance mechanisms to genotoxicity of sediment, enriched in marine strains. The number of core genes increased when the genomes of marine strains were considered, compared to all 138 genomes. Genome analysis of 13 marine *Brevibacterium* isolates revealed that 1963 core genes represent the core genome. The highest percentage of these core genes was found to be distributed in the amino acid transport and metabolism COG category, signifying the importance of these genes in marine adaptation. Different strains of *Brevibacterium* follow an aerobic chemoorganoheterotrophic mode of nutrition (6, 16, 17). The uptake and utilization of readily available mono/oligomers of structural and functional cell components from the environment is an adaptive strategy in energy-limited environments (39, 40).

Marine waterbodies and their sediment systems are the active zones where organic compound remineralization takes place, resulting in bioavailable amino acids and short peptide chains in both the water column and underlying sediment (41, 42). Amino acids are available in aquatic environments as simple monomers, short peptide chains, and proteins (40). α- and β-amino acids, along with their derivatives, are often used by bacteria as compatible solutes for osmoadaptation. The uptake of readily available amino acids from the surroundings and their incorporation into biomass are adaptive features of marine life, enabling microbes to economically funnel available energy while withstanding the pressure of the water column and high salinity (39, 40). In addition to amino acid transport and metabolism, the core genome of marine strains includes a high percentage of core genes from the carbohydrate transport and metabolism and energy production and conversion COG categories. Genes in these categories support proper energy budgeting within microbial cells by utilizing readily available compounds from the environment. Polymeric carbohydrates and sugars are also often used by microorganisms as osmolytes to cope with high osmotic pressure (43), making their transport an adaptive strategy for marine habitats. Furthermore, for marine adaptation, an enhancement of core genes was observed in the COG category related to inorganic ion transport and metabolism. The exchange of inorganic ions with the surrounding environment is a key adaptation for marine habitats, alongside the transport of compatible solutes (44, 45). Additional core genes associated with marine adaptation were observed across most of the COG categories (Fig. 3B). These include mechanosensitive ion channels, proteins involved in the proper folding of polypeptides in stressed conditions, DNA damage repair machinery, and membrane channels responsible for solute homeostasis of bacterial cytoplasm. Various mixtures of chemical and organic compounds in marine sediments cause damage to cellular DNA (46). To overcome this genotoxic potential, the presence of DNA repair machinery is a crucial adaptive strategy for *Brevibacterium* strains.

Marine province-specific genome evolution sheds light on oceanic habitat evolution. Strains of *Brevibacterium* living in the seawater-sediment systems of the three major marine provinces—Atlantic Ocean, Indian Ocean, and Pacific Ocean—were found to possess three differently sized core genomes consisting of core genes (Fig. 5A). While the distribution of core genes across different COG categories in province-specific datasets does not show a major difference (Fig. 5B), a detailed analysis reveals distinct variations. Genomes from each marine province share an additional set of core genes among themselves, in addition to core genes associated with marine adaptation. Although connected through the water, sediment systems of individual oceanic regions become isolated due to their vastness. This suggests that marine adaptation is not a simple, ubiquitous metabolic strategy for *Brevibacterium* strains; rather, specific genome-guided metabolic traits must be acquired. The isolated nature of each marine province influences genome evolution in *Brevibacterium* strains, indicating that marine strains are not

recently introduced microorganisms. Instead, they have evolved through long term divergence and adaptation to marine habitats shaped by geological and plate tectonic processes.

For adaptation in the Indian Ocean province, core genome expansion includes genes involved in cell division regulation and determination of the peptidoglycan layer diameter. Various proteins responsible for releasing DNA supercoiling, non-homologous end-joining DNA repair machinery, transcriptional regulators, and inorganic ion transporter proteins are integrated into the core genome of strains from the Indian Ocean territory. The Indian Ocean harbors the Arabian Sea OMZ and Bay of Bengal OMZ, and the isolation sites of *Brevibacterium* strains were located near these OMZs. For aerobic bacteria adapting to stressed environments – characterized by O_2 limitation, nutrient scarcity, heavy metal enrichment, temperature variation, and higher salinity - cell size adjustment, DNA repair mechanisms, and transcriptional regulation at various stages are known to be important (47-50). The expanded core genome of the two Arabian Sea isolates includes genes related to metabolism under carbon starvation, arsenic-resistant machinery, sodium and proton transporters, and cold shock proteins. These metabolic features are crucial for survival in environments with fluctuating nutrient availability, heavy metal presence, wide temperature variation, and high osmotic pressure (51-54). When genomes from any of the three marine provinces were analyzed, using the largest closed circular genome from each dataset as a reference, other genomes from the same dataset were mapped onto it using the BLAST algorithm (with an E-value cutoff of 0.0001). A similar pattern of mapped and unmapped regions was observed, further supporting genomic relatedness among strains within each marine province (Fig. 6).

The two genomes of pure culture isolates *Brevibacterium* sp. BDJS002 and *Brevibacterium* sp. JSBI002, living in the sediment of the Arabian Sea OMZ, diverge in both core genome-based and pan-genome-based phylogenies. Each strain shows highest digital DNA-DNA hybridization (dDDH) value with different type strains (Table 1). However, when only genomes from the Arabian Sea are considered, these two strains share 2584 non-redundant core genes (an increase compared to the Indian_Ocean_Genomes dataset), indicating that the marine habitat plays a pivotal role in shaping genome evolution through horizontal gene transfer to acquire niche-specific metabolic potential. As the geographic scope of marine provinces narrows, a greater number of genes are included in the core genome, regardless of species-level taxonomic identity.

Searching for recently introduced genes within the genomes of marine isolates, through horizontal gene transfer mechanisms in the genome-specific unique gene set, reveals high genomic plasticity, as genes from major lineages such as Proteobacteria, Firmicutes, and Bacteroidetes were found to be incorporated into the genome, encoding various glycosyl transferases, endonucleases, and ion channels. Furthermore, genes of viral origin were also observed to be integrated into the genomes of marine isolates.

CONCLUSION

Core genome and pan-genome-based analysis of *Brevibacterium* strains revealed a high degree of marine province-specific core genome conservation (Figs. 5A and 6). This genomic consistency, observed across various provinces, strongly suggests that the diversification and dispersal of marine *Brevibacterium* populations is tightly coupled with the evolution and expansion of their associated habitats through plate tectonics. Given that oceanic sediment systems are inherently more stable than flowing waterbodies or terrestrial environments of earth's surface, this environment imposes distinct, persistent selective pressures on resident microbes. This environmental stability likely acts to preserve the evolutionary lineage of essential genomic traits. Future potential lies in integrating these findings with larger, taxonomically diverse metagenomic datasets to further resolve the co-evolutionary patterns between marine habitat development and the underlying mechanisms of bacterial genome evolution.

ACKNOWLEDGEMENTS

The server used for data analysis was provided by Professor Wriddhiman Ghosh, Bose Institute, India. I am also thankful to him for his valuable intellectual support.

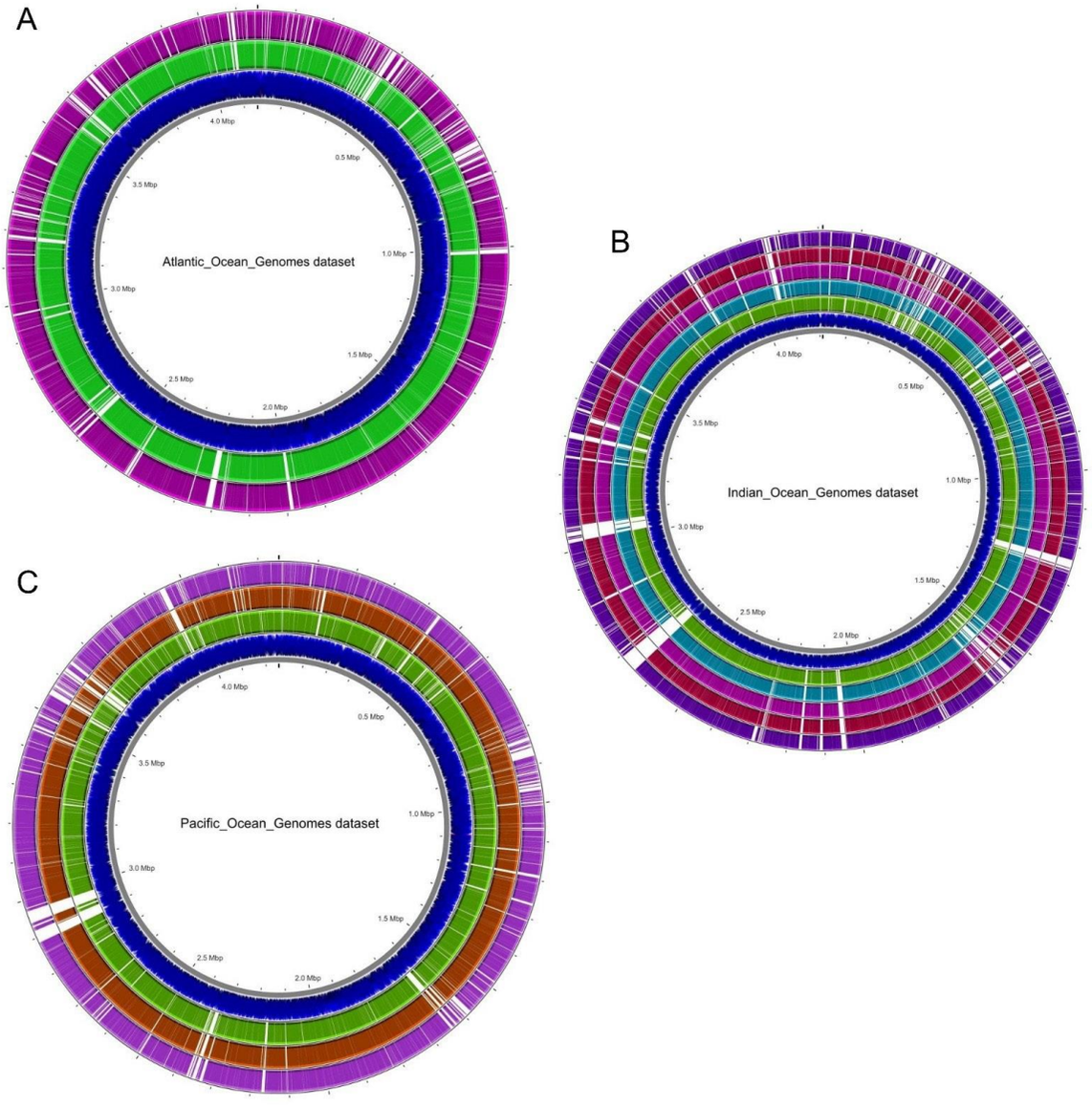


Fig. 6. Circular genome-genome mapping visualization for: (A) All three genomes in the Atlantic_Ocean_Genomes dataset. From inner to outer colored rings; CDSs of *Brevibacterium atlanticum* WO024; mapped region of *Brevibacterium oceani* WW007 on *Brevibacterium atlanticum* WO024; mapped region of *Brevibacterium pigmentatum* YB235 on *Brevibacterium atlanticum* WO024. (B) All six genomes in the Indian_Ocean_Genomes dataset. From inner to outer colored rings; CDSs of *Brevibacterium* sp. BDJS002; mapped region of *Brevibacterium* sp. JSBI002 on *Brevibacterium* sp. BDJS002; mapped region of *Brevibacterium sediminis* COD27 on *Brevibacterium* sp. BDJS002; mapped region of *Brevibacterium sediminis* FXJ8.269 on *Brevibacterium* sp. BDJS002; mapped region of *Brevibacterium sediminis* CGMCC 1.15472 on *Brevibacterium* sp. BDJS002; mapped region of *Brevibacterium oceani* BBH7 on *Brevibacterium* sp. BDJS002. (C) All four genomes in the Pacific_Ocean_Genomes dataset. From inner to outer colored rings; CDSs of *Brevibacterium limosum* o2; mapped region of *Brevibacterium marinum* DSM 18964 on *Brevibacterium limosum* o2; mapped region of *Brevibacterium* sp. Marine on *Brevibacterium limosum* o2; mapped region of *Brevibacterium* sp. CCUG 69071 on *Brevibacterium limosum* o2; Mapping for each dataset was performed using the BLAST algorithm with an E-value cutoff of 0.0001. The coloured rings show regions of high sequence homology illustrating the degree of genomic synteny and conservation among strains from the same province.

REFERENCES

1. Ward AC, Bora N. Diversity and biogeography of marine actinobacteria. *Curr Opin Microbiol* 2006; 9: 279-286.
2. Valverde A, Tuffin M, Cowan DA. Biogeography of bacterial communities in hot springs: a focus on the actinobacteria. *Extremophiles* 2012; 16: 669-679.
3. Mohammadipanah F, Wink J. Actinobacteria from arid and desert Habitats: Diversity and Biological activity. *Front Microbiol* 2016; 6: 1541.
4. Neuenschwander SM, Ghai R, Pernthaler J, Salcher MM. Microdiversification in genome-streamlined ubiquitous freshwater Actinobacteria. *ISME J* 2018; 12: 185-198.
5. Deng T, Lu H, Qian Y, Chen X, Yang X, Guo J, et al. *Brevibacterium rongguiense* sp. nov., isolated from freshwater sediment. *Int J Syst Evol Microbiol* 2020; 70: 5205-5210.
6. Chen P, Zhang L, Wang J, Ruan J, Han X, Huang Y. *Brevibacterium sediminis* sp. nov., isolated from deep-sea sediments from the Carlsberg and Southwest Indian Ridges. *Int J Syst Evol Microbiol* 2016; 66: 5268-5274.
7. Jung MS, Quan XT, Siddiqi MZ, Liu Q, Kim SY, Wee JH, et al. *Brevibacterium anseongense* sp. nov., isolated from soil of ginseng field. *J Microbiol* 2018; 56: 706-712.
8. Belov AA, Cheptsov VS, Vorobyova EA, Manucharova NA, Ezhelev ZS. Stress-tolerance and taxonomy of culturable bacterial communities isolated from a Central Mojave Desert soil sample. *Geosciences* 2019; 9: 166.
9. Pei S, Xie F, Niu S, Ma L, Zhang R, Zhang G. *Brevibacterium profundi* sp. nov., isolated from deep-sea sediment of the Western Pacific Ocean. *Int J Syst Evol Microbiol* 2020; 70: 5818-5823.
10. Zhao J, Shakir Y, Deng Y, Zang Y. Use of modified ichip for the cultivation of thermo-tolerant microorganisms from the hot spring. *BMC Microbiol* 2023; 23: 56.
11. Collins MD. The genus *Brevibacterium*. *Prokaryotes* 2006; 3: 1013-1019.
12. Hoshino T, Doi H, Uramoto GI, Wörmer L, Adhikari RR, Xiao N, et al. Global diversity of microbial communities in marine sediment. *Proc Natl Acad Sci* U S A 2020; 117: 27587-27597.
13. Dmitrijeva M, Tackmann J, Matias Rodrigues JF, Huerta-Cepas J, Coelho LP, von Mering C. A global survey of prokaryotic genomes reveals the eco-evolutionary pressures driving horizontal gene transfer. *Nat Ecol Evol* 2024; 8: 986-998.
14. Arnold BJ, Huang IT, Hanage WP. Horizontal gene transfer and adaptive evolution in bacteria. *Nat Rev Microbiol* 2022; 20: 206-218.
15. Lee SD. *Brevibacterium marinum* sp. nov., isolated from seawater. *Int J Syst Evol Microbiol* 2008; 58: 500-504.
16. Pei S, Niu S, Xie F, Wang W, Zhang S, Zhang G. *Brevibacterium limosum* sp. nov., *Brevibacterium pigmenatum* sp. nov., and *Brevibacterium atlanticum* sp. nov., three novel dye decolorizing actinobacteria isolated from ocean sediments. *J Microbiol* 2021; 59: 898-910.
17. Bhadra B, Raghukumar C, Pindi PK, Shivaji S. *Brevibacterium oceani* sp. nov., isolated from deep-sea sediment of the Chagos Trench, Indian Ocean. *Int J Syst Evol Microbiol* 2008; 58: 57-60.
18. Pham NP, Layec S, Dugat-Bony E, Vidal M, Irlinger F, Monnet C. Comparative genomic analysis of *Brevibacterium* strains: insights into key genetic determinants involved in adaptation to the cheese habitat. *BMC Genomics* 2017; 18: 955.
19. Levesque S, de Melo AG, Labrie SJ, Moineau S. Mobilome of *Brevibacterium aurantiacum* sheds light on its genetic diversity and its adaptation to smear-ripened cheeses. *Front Microbiol* 2019; 10: 1270.
20. Cumsille A, Serna-Cardona N, González V, Claverías F, Undabarrena A, Molina V, et al. Exploring the biosynthetic gene clusters in *Brevibacterium*: a comparative genomic analysis of diversity and distribution. *BMC Genomics* 2023; 24: 622.
21. Sarkar J, Mondal M, Bhattacharya S, Dutta S, Chatterjee S, Mondal N, et al. Extremely oligotrophic and complex-carbon-degrading microaerobic bacteria from Arabian Sea oxygen minimum zone sediments. *Arch Microbiol* 2024; 206: 179.
22. Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 2019; 10: 2182.
23. Chaudhari NM, Gupta VK, Dutta C. BPGA- an ultra-fast pan-genome analysis pipeline. *Sci Rep* 2016; 6: 24373.
24. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010; 26: 2460-2461.
25. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 2004; 5: 113.
26. Vos RA, Caravas J, Hartmann K, Jensen MA, Miller C. BIO::Phylo-phyloinformatic analysis using perl. *BMC Bioinformatics* 2011; 12: 63.
27. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018; 9: 5114.
28. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Mol Biol Evol* 2021; 38: 5825-5829.
29. Liu M, Siezen RJ, Nauta A. In silico prediction of hori-

zontal gene transfer events in *Lactobacillus bulgaricus* and *Streptococcus thermophilus* reveals protocooperation in yogurt manufacturing. *Appl Environ Microbiol* 2009; 75: 4120-4129.
30. Ravenhall M, Škunca N, Lassalle F, Dessimoz C. Inferring horizontal gene transfer. *PLoS Comput Biol* 2015; 11(5): e1004095.
31. D'Hondt S, Spivack AJ, Pockalny R, Ferdelman TG, Fischer JP, Kallmeyer J, et al. Subseafloor sedimentary life in the South Pacific Gyre. *Proc Natl Acad Sci U S A* 2009; 106: 11651-11656.
32. D'hondt S, Inagaki F, Zarikian CA, Abrams LJ, Dubois N, Engelhardt T, et al. Presence of oxygen and aerobic communities from sea floor to basement in deep-sea sediments. *Nat Geosci* 2015; 8: 299-304.
33. Sarkar J, Mondal N, Mandal S, Chatterjee S, Ghosh W (2022). Deep Subsurface Microbiomes of the Marine Realm. Systems Biogeochemistry of Major Marine Biomes. pp109-131.
34. Griffiths JR, Kadin M, Nascimento FJ, Tamelander T, Törnroos A, Bonaglia S, et al. The importance of benthic–pelagic coupling for marine ecosystem functioning in a changing world. *Glob Chang Biol* 2017; 23: 2179-2196.
35. Kallmeyer J, Pockalny R, Adhikari RR, Smith DC, D'Hondt S. Global distribution of microbial abundance and biomass in subseafloor sediment. *Proc Natl Acad Sci U S A* 2012; 109: 16213-16216.
36. Torres-Beltrán M, Vargas-Gastélum L, Magdaleno-Moncayo D, Riquelme M, Herguera-García JC, Prieto-Davó A, et al. The metabolic core of the prokaryotic community from deep-sea sediments of the southern Gulf of Mexico shows different functional signatures between the continental slope and abyssal plain. *PeerJ* 2021; 9: e12474.
37. Batzke A, Engelen B, Sass H, Cypionka H. Phylogenetic and physiological diversity of cultured deep-biosphere bacteria from Equatorial Pacific Ocean and Peru Margin sediments. *Geomicrobiol J* 2007; 24: 261-273.
38. Bhattacharya S, Roy C, Mandal S, Sarkar J, Rameez MJ, Mondal N, et al. Aerobic microbial communities in the sediments of a marine oxygen minimum zone. *FEMS Microbiol Lett* 2020; 367: fnaa157.
39. Shoemaker WR, Jones SE, Muscarella ME, Behringer MG, Lehmkuhl BK, Lennon JT. Microbial population dynamics and evolutionary outcomes under extreme energy limitation. *Proc Natl Acad Sci U S A* 2021; 118(33): e2101691118.
40. Alonso-Sáez L, Gasol JM. Seasonal variations in the contributions of different bacterial groups to the uptake of low-molecular-weight compounds in northwestern Mediterranean coastal waters. *Appl Environ Microbiol* 2007; 73: 3528-3535.
41. Fernandes L, Garg A, Borole DV. Amino acid biogeochemistry and bacterial contribution to sediment organic matter along the western margin of the Bay of Bengal. *Deep-Sea Res I: Oceanogr Res Pap* 2014; 83: 81-92.
42. Choi H, Choi B, Chikaraishi Y, Takano Y, Kim H, Lee K, et al. Microbial alteration in marine sediments: Insights from compound-specific isotopic compositions of amino acids in subseafloor environments. *Front Mar Sci* 2022; 9: 10.3389/fmars.2022.1030669.
43. Yancey PH. Compatible and Counteracting Solutes: Protecting Cells from the Dead Sea to the Deep Sea. *Sci Prog* 2004; 87: 1-24.
44. Wood JM, Bremer E, Csonka LN, Kraemer R, Poolman B, van der Heide T, et al. Osmosensing and osmoregulatory compatible solute accumulation by bacteria. *Comp Biochem Physiol A Mol Integr Physiol* 2001; 130: 437-460.
45. Welsh DT. Ecological significance of compatible solute accumulation by micro-organisms: from single cells to global climate. *FEMS Microbiol Rev* 2000; 24: 263-290.
46. Kammann U, Riggers JC, Theobald N, Steinhart H. Genotoxic potential of marine sediments from the North Sea. *Mutat Res* 2000; 467: 161-168.
47. Dai X, Zhu M. High Osmolarity Modulates Bacterial cell size through reducing initiation volume in *Escherichia coli*. *mSphere* 2018; 3(5): e00430-18.
48. Dupuy P, Sauviac L, Bruand C. Stress-inducible NHEJ in bacteria: function in DNA repair and acquisition of heterologous DNA. *Nucleic Acids Res* 2019; 47: 1335-1349.
49. Mathivanan K, Chandirika JU, Vinothkanna A, Yin H, Liu X, Meng D. Bacterial adaptive strategies to cope with metal toxicity in the contaminated environment – A review. *Ecotoxicol Environ Saf* 2021; 226: 112863.
50. Pedraza-Reyes M, Abundiz-Yañez K, Rangel-Mendoza A, Martínez LE, Barajas-Ornelas RC, Cuéllar-Cruz M, et al. *Bacillus subtilis* stress-associated mutagenesis and developmental DNA repair. *Microbiol Mol Biol Rev* 2024; 88(2): e0015823.
51. Hwang S, Choe D, Yoo M, Cho S, Kim SC, Cho S, et al. Peptide transporter CstA imports pyruvate in *Escherichia coli* K-12. *J Bacteriol* 2018; 200(7): e00771-17.
52. Mohsin H, Shafique M, Zaid M, Rehman Y. Microbial biochemical pathways of arsenic biotransformation and their application for bioremediation. *Folia Microbiol (Praha)* 2023; 68: 507-535.
53. Zhang Y, Gross CA. Cold shock response in Bacteria. *Annu Rev Genet* 2021; 55: 377-400.
54. Mandal S, Bhattacharya S, Roy C, Ramez MJ, Sarkar j, Mapder T, et al. Cryptic roles of tetrathionate in the sulfur cycle of marine sediments: microbial drivers and indicators. *Biogeosciences* 2020; 17: 4611-4631.