Marine Genomics 82 (2025) 101204

Contents lists available at ScienceDirect

Marine Genomics

journal homepage: www.elsevier.com/locate/margen

Data Article

Complete genome sequence of *Lysinibacillus* sp. WB86, a potential antimicrobial lanthipeptides producer

Huai-Ying Sun [a], Ji-Xiang Xin [a], Gen Li [a], Zhen-Yi Zhou [a], Xin Li [c], Hong Wang [a], Qun-Jian Yin [b,*], Bin Wei [a,*]

[a] *College of Pharmaceutical Science & Collaborative Innovation Center of Yangtze River Delta Region Green Pharmaceuticals, Zhejiang University of Technology, Hangzhou 310014, China*
[b] *Fourth Institute of Oceanography, Ministry of Natural Resources, Beihai 536015, China*
[c] *CAS and Shandong Province Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences, Nanhai Road 7, Qingdao 266071, China*

ARTICLE INFO

Keywords:
Complete genome
Lysinibacillus
Cold seep
South China Sea
Biosynthetic gene cluster
Lanthipeptides

ABSTRACT

Lysinibacillus sp. WB86 was isolated from a cold seep in the South China Sea, and its complete genome was sequenced using Oxford Nanopore Technologies (ONT). The genome consists of a single circular chromosome spanning 4,537,071 bp, with a G + C content of 37 %. FastANI analysis revealed a 99.38 % average nucleotide identity (ANI) with *Lysinibacillus sphaericus* A1, confirming its classification as a strain of *L. sphaericus*. Genome annotation identified 4397 protein-coding genes, along with 111 tRNAs and 34 rRNAs. Functional categorization using the COG database assigned 84.6 % of the protein-coding genes to specific roles, with the majority involved in transcription, amino acid transport and metabolism, translation and ribosomal biogenesis, and replication, recombination, and repair. Additionally, antiSMASH-based analysis uncovered eight biosynthetic gene clusters (BGCs), only one of which exhibited high similarity (>50 %) to known BGCs. This cluster was predicted to encode novel antimicrobial lanthipeptides, suggesting potential biotechnological applications.

1. Introduction

Deep-sea cold seeps represent unique submarine geological structures primarily distributed along continental margins and plate boundaries. They form through the upward migration or eruption of methane-rich hydrocarbon fluids via sedimentary faults or fissures (Boetius and Wenzhöfer, 2013). The identification of seep carbonates—a byproduct of hydrocarbon seepage—on the northern passive continental margin of the South China Sea in 2004 marked the beginning of cold seep research in this region (Chen et al., 2005). Since then, the South China Sea cold seeps have attracted growing scientific interest and are now regarded as one of the most extensively studied seep systems globally (Feng et al., 2018). Unlike ecosystems dependent on photosynthesis, cold seep ecosystems primarily utilize methane released from the seafloor as their main carbon and energy source, supporting microbial communities and overall productivity through chemosynthesis (Wang et al., 2025). This distinctive energy acquisition mechanism renders cold seeps "deep-sea oases," harboring exceptionally high biodiversity and biomass concentrations (Ruff et al., 2015).

The genus *Lysinibacillus*, a member of the Bacillaceae family, is commonly isolated from marine environments. Although phylogenetically related to *Bacillus*, it differs significantly in cell wall composition. In 2007, several species previously classified under *Bacillus* were reclassified into the new genus *Lysinibacillus* (Ahmed et al., 2007). The genus name reflects both its unique lysine biosynthesis pathway and its evolutionary connection to spore-forming *Bacillus*. *Lysinibacillus* have been reported to have multifarious abilities, such as entomopathogenic, bioremediation, plant growth-promoting, and biocontrol abilities; thus, it was thought to be a good source of biopesticide, plant biostimulant, and bioremediation agent (Ahsan and Shimizu, 2021).

2. Data description

2.1. Strain and separation condition

Lysinibacillus sp. WB86 was isolated from cold seep sediment collected from a cold seep off the southwestern coast of Taiwan in the South China Sea. The isolation was performed using an enrichment

* Corresponding authors.
E-mail address: binwei@zjut.edu.cn (B. Wei).

https://doi.org/10.1016/j.margen.2025.101204
Received 30 June 2025; Accepted 2 August 2025

cultivation approach: sediment samples were first incubated statically at 20 °C for 48 h in 1/10 ISP_3 liquid medium. Subsequently, the cultures were spread onto ISP_3 solid medium plates and further incubated under static conditions at 20 °C. Single colonies were purified through two rounds of streaking to obtain axenic cultures. Phylogenetic analysis based on 16S rRNA gene sequencing confirmed the strain's taxonomic affiliation to the genus *Lysinibacillus*, and it was designated as *Lysinibacillus* sp. WB86.

2.2. Sequencing and data analysis

Lysinibacillus sp. WB86 cells were cultured in 2216E liquid medium at 28 °C for 48 h. After centrifugation at 12,000 rpm for 10 min, genomic DNA was extracted from the cell pellets using a Genomic Kit (Cat#13343 , QIAGEN) following the manufacturer's protocol. Long-read sequencing was performed on the PromethION platform (Oxford Nanopore Technologies, UK) by Minco Biotechnology (Hangzhou, China). The ONT long-read data were de novo assembled using Flye v2.9 (−-nano-raw mode), followed by consensus polishing with Medaka (Kolmogorov et al., 2020). To enhance accuracy, high-quality Illumina short reads were mapped to the draft assembly using Bowtie 2 (Langmead and Salzberg, 2012). The final assembly was evaluated with CheckM2 (Chklovski et al., 2023), with key metrics including total genome size (bp), GC content, N50 value, and contig count reported to assess assembly quality.

Automated genome annotation was performed using Prokka (Seemann, 2014), which predicted coding DNA sequences (CDSs), tRNAs, rRNAs, and genome topology. Functional annotation of protein-coding genes was conducted using the eggNOG (Non-supervised Orthologous Groups) database (Huerta-Cepas et al., 2017). Signal peptides were identified with SignalP v6.0 (Teufel et al., 2022), while transmembrane helices were predicted using the TMHMM Server v2.0 (Server, 2015). The Average Nucleotide Identity (ANI) between *Lysinibacillus* sp. WB86 and all available *Lysinibacillus* genomes in the NCBI database was calculated using FastANI (Jain et al., 2018). Strains exhibiting ANI values >95 % were selected to generate a heatmap matrix, facilitating taxonomic placement of strain WB86. Biosynthetic gene clusters (BGCs) were predicted using antiSMASH v7.0 (Bacterial Version) (Blin et al., 2023). Following genome assembly and BGC annotation, a circular genome map was generated with Proksee in CGView Suite (Grant et al., 2023). For comparative analysis, representative BGCs identified by antiSMASH were visualized and aligned using Clinker (Gilchrist and Chooi, 2021).

2.3. Characterization of the genome of Lysinibacillus sp. WB86

The general features and Minimum Information about a Genome Sequence (MIGS) compliance data of *Lysinibacillus* sp. WB86 are summarized in Table 1. Nanopore sequencing generated 1.6 GB of high-quality clean data. The assembled complete genome comprises a single circular chromosome of 4,537,071 bp with a GC content of 37 %. The genome encodes 4543 genes, including 4397 protein-coding sequences (CDSs), 111 tRNAs, and 34 rRNAs (11 copies each of 23S and 16S rRNA, and 12 copies of 5S rRNA). Functional annotation revealed that 641 (14.58 %), 3720 (84.60 %), and 2262 (51.44 %) genes were assigned to GO, COG, and KEGG databases, respectively. COG classification categorized genes into 19 functional groups, with predominant roles in: transcription; amino acid transport and metabolism; translation, ribosomal structure and biogenesis; as well as replication, recombination and repair.

2.4. FastANI analysis

A heatmap matrix was constructed for *Lysinibacillus* sp. WB86 and 20 closely related strains (ANI >95 %) to evaluate phylogenetic relationships and species delineation (Fig. 1A). The ANI-based clustering revealed distinct phylogenetic patterns: (1) *Lysinibacillus sphaericus* strains formed a highly conserved cluster clearly separated from other species (including L. *tabacifolii*, *L. varians*, and L. *mangiferihumi*), demonstrating robust species boundaries; (2) *Lysinibacillus* sp. WB86 exhibited exceptionally high ANI similarity with multiple L. *sphaericus* strains. Notably, strain WB86 showed the highest ANI value (99.38 %) with L. *sphaericus* A1, confirming its taxonomic classification within this species.

Table 1
General features of *Lysinibacillus* sp. WB86 and MIGS mandatory information.

Items	Description
General feature	
Classification	Domain Bacteria Phylum Firmicutes Class Bacilli Order Caryophanales Family Caryophanaceae Genus *Lysinibacillus* Strain *Lysinibacillus sphaericus* A1
Gram stain	Positive
Cell shape	Rod-shaped
Pigmentation	Beige to pale yellow
Relationship to oxygen	Aerobic
Biotic relationship	Free living
MIGS data	
Sample name	*Lysinibacillus* sp. WB86
Project name	*Lysinibacillus* sp. WB86 genome sequencing
Geographic location	cold seep offshore southwestern Taiwan, South China Sea, China
Depth	−8 cm
Elevation	1159 m
Collection date	2023-10-31
Environment (biome)	ENVO: 01000127
Environment (feature)	ENVO: 01000263
Environment (materlal)	ENVO: 03000033
Sequencing method	Oxford Nanopore
Assembly method	Flye
Coverage	308
Finishing strategy	Complete
Genome features	
Chromosome	1
Genome topology	Circular
Genome size (bp)	4,537,071
G + C content (%)	37
Total genes	4543
Protein coding genes	4397
tRNA genes	111
rRNA operon (16S–5S-23S rRNA)	34
Genes assigned to GO	641
Genes assigned to COGs	3720
Genes assigned to KEGG	2262
Genes with signal peptides	169
Genes with transmembrane helices	5626
Biosynthetic gene clusters of secondary metabolites	8
GenBank accession number	CP194323

2.5. Secondary metabolic potential analysis

The potential of *Lysinibacillus* sp. WB86 to synthesize novel secondary metabolites was analyzed using antiSMASH and BIG-SCAPE. The antiSMASH analysis identified eight biosynthetic gene clusters (BGCs) in the genome of Lysinibacillus sp. WB86 (Fig. 1B & Table 2). Among these, only one displayed high similarity (>50 %) to known BGCs. Comparative analysis using clinker (Fig. 1C) indicated that WB86.region3 contains key biosynthetic genes associated with lanthipeptide production, including a methyltransferase, precursor peptide, class III lanthionine synthetase, and ABC transporters. These genes exhibited 37–78 % sequence similarity to their homologs, particularly in the precursor peptide and modifying enzyme regions. Notably, the characterized

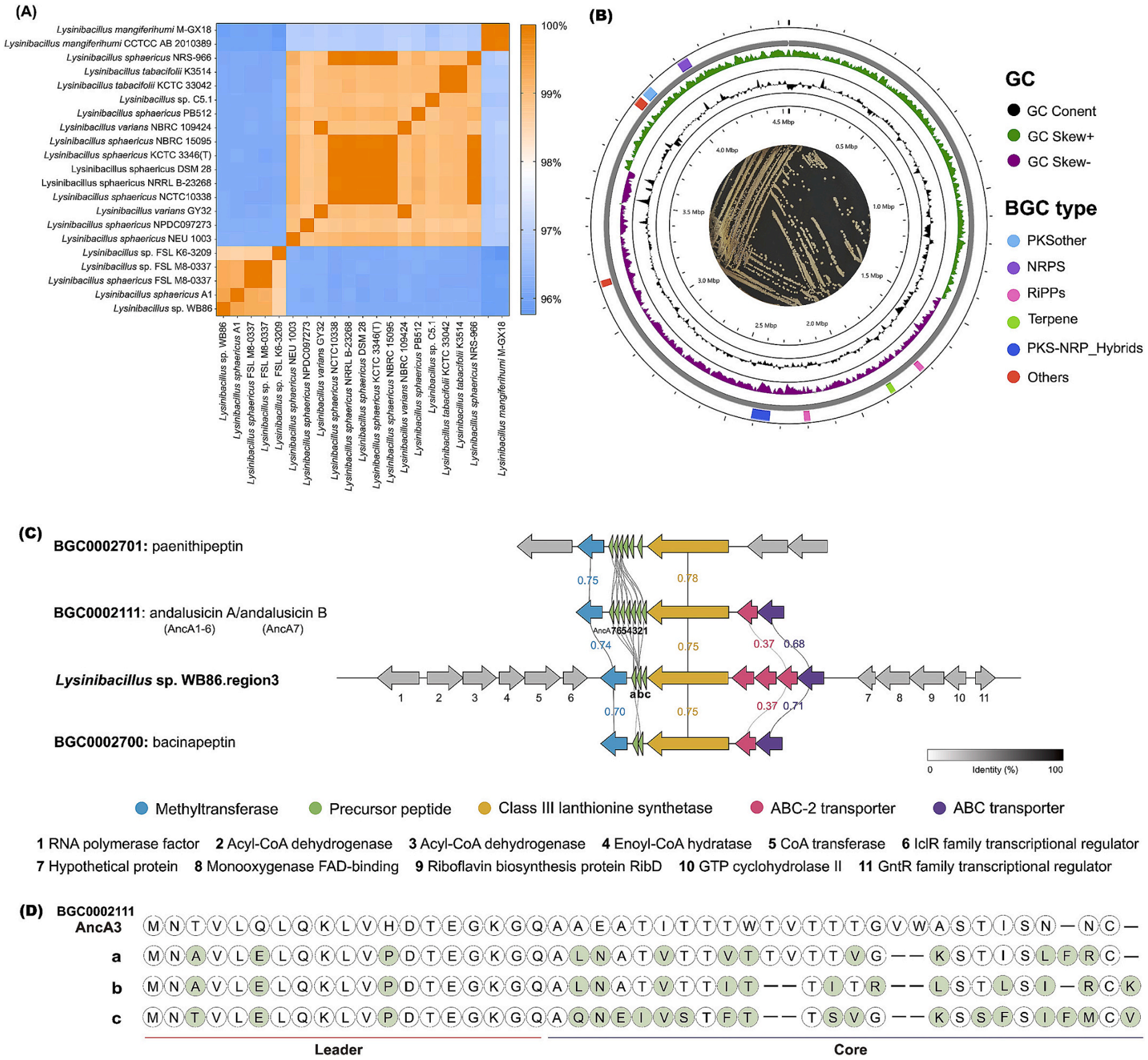

Fig. 1. (**A**) ANI matrix heatmap of *Lysinibacillus* sp. WB86 and 20 phylogenetically related *Lysinibacillus* strains. (**B**) Schematic representation of the genome of *Lysinibacillus* sp. WB86. The inner two rings show GC content and GC skew. The outermost ring shows the positions of predicted biosynthetic gene clusters. (**C**) Comparative architecture of lanthionine BGCs in *Lysinibacillus* sp. WB86 and orthologs. (**D**) Sequence alignment of *Lysinibacillus* sp. WB86.region3 with the AncA3 precursor peptide from BGC000211.

lanthipeptide andalusicin A (encoded by AncA1–6 from BGC000211) has demonstrated inhibitory activity against Gram-positive bacteria (Grigoreva et al., 2021). Sequence alignment of the three precursor peptides from WB86.region3 with the closest homolog, AncA3 (Fig. 1D), revealed significant nucleotide variations, further supporting the potential novelty of these peptides. These findings highlight *Lysinibacillus* sp. WB86 as a promising candidate for the production of novel antimicrobial lanthipeptides.

3. Nucleotide sequence accession number

Strain *Lysinibacillus* sp. WB86 has been deposited in the Guangdong Microbial Culture Collection Center (GDMCC1.5765). The complete genome sequence and its annotation information were submitted to the NCBI database by accession number CP194323 (BioProject ID: PRJNA1265892 and BioSample ID: SAMN48636238).

Funding

This work was supported by the international (regional) cooperation and exchange program jointly initiated by the Guangxi Key Laboratory of Beibu Gulf Marine Resources, Environment and Sustainable Development (no. MRESD-2024-A04), the National Natural Science Programs of China (nos. W2412100 and 42276137), and the National Key Research and Development Programs (nos. 2022YFC2804700 and 2022YFC2804104).

CRediT authorship contribution statement

Huai-Ying Sun: Writing – original draft, Investigation,

Table 2
BGCs encoded by *Lysinibacillus* sp. WB86.

BGCs	BGC type	Most similar known cluster description	Similarity to MIBiG BGCs
Region 1	RiPPs	Unknown	Unknown
Region 2	Terpene	Unknown	Unknown
Region 3	RiPPs	Andalusicin A/andalusicin B	90 %
Region 4	PKS-NRP_Hybrids	Pellasoren/pellasoren B	16 %
Region 5	Others	Fengycin	46 %
Region 6	Others	Schizokinen	20 %
Region 7	PKSother	Bacillibactin/bacillibactin E/ bacillibactin F	30 %
Region 8	NRPS	N-tetradecanoyl tyrosine	6 %

Visualization. **Ji-Xiang Xin:** Validation. **Gen Li:** Validation. **Zhen-Yi Zhou:** Methodology. **Xin Li:** Conceptualization, Project administration. **Hong Wang:** Funding acquisition. **Qun-Jian Yin:** Resources. **Bin Wei:** Conceptualization, Writing – review & editing.

Declaration of competing interest

The authors declare that there are no conflicts of interest.

Data availability

Data will be made available on request.

References

Ahmed, I., Yokota, A., Yamazoe, A., Fujiwara, T., 2007. Proposal of Lysinibacillus boronitolerans gen. Nov. sp. nov., and transfer of Bacillus fusiformis to Lysinibacillus fusiformis comb. nov. and Bacillus sphaericus to Lysinibacillus sphaericus comb. nov. Int. J. Syst. Evol. Microbiol. 57, 1117–1125.

Ahsan, N., Shimizu, M., 2021. *Lysinibacillus* species: their potential as effective bioremediation, biostimulant, and biocontrol agents. Rev Agric Sci. 9, 103–116.

Blin, K., Shaw, S., Augustijn, H.E., Reitz, Z.L., Biermann, F., Alanjary, M., Fetter, A., Terlouw, B.R., Metcalf, W.W., Helfrich, E.J., Van Wezel, G.P., 2023. AntiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Res. 51, W46–W50.

Boetius, A., Wenzhöfer, F., 2013. Seafloor oxygen consumption fuelled by methane from cold seeps. Nat. Geosci. 6, 725–734.

Chen, D.F., Huang, Y.Y., Yuan, X.L., Cathles, L.M., 2005. Seep carbonates and preserved methane oxidizing archaea and sulfate reducing bacteria fossils suggest recent gas venting on the seafloor in the northeastern South China Sea. Mar. Pet. Geol. 22, 613–621.

Chklovski, A., Parks, D.H., Woodcroft, B.J., Tyson, G.W., 2023. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. Nat. Methods 20, 1203–1212.

Feng, D., Qiu, J.W., Hu, Y., Peckmann, J., Guan, H., Tong, H., Chen, C., Chen, J., Gong, S., Li, N., Chen, D., 2018. Cold seep systems in the South China Sea: an overview. J. Asian Earth Sci. 168, 3–16.

Gilchrist, C.L.M., Chooi, Y.H., 2021. Clinker & clustermap.Js: automatic generation of gene cluster comparison figures. Bioinformatics 37, 2473–2475.

Grant, J.R., Enns, E., Marinier, E., Mandal, A., Herman, E.K., Chen, C., Graham, M., Van Domselaar, G., Stothard, P., 2023. Proksee: in-depth characterization and visualization of bacterial genomes. Nucleic Acids Res. 51, W484–W492.

Grigoreva, A., Andreeva, J., Bikmetov, D., Rusanova, A., Serebryakova, M., Garcia, A.H., Slonova, D., Nair, S.K., Lippens, G., Severinov, K., Dubiley, S., 2021. Identification and characterization of andalusicin: N-terminally dimethylated class III lantibiotic from Bacillus thuringiensis sv. andalousiensis. iScience 24, 102480.

Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., Von Mering, C., Bork, P., 2017. Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. Mol. Biol. Evol. 34, 2115–2122.

Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T., Aluru, S., 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat. Commun. 9, 5114.

Kolmogorov, M., Bickhart, D.M., Behsaz, B., Gurevich, A., Rayko, M., Shin, S.B., Kuhn, K., Yuan, J., Polevikov, E., Smith, T.P.L., Pevzner, P.A., 2020. metaFlye: scalable long-read metagenome assembly using repeat graphs. Nat. Methods 17, 1103–1110.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with bowtie 2. Nat. Methods 9, 357–359.

Ruff, S.E., Biddle, J.F., Teske, A.P., Knittel, K., Boetius, A., Ramette, A., 2015. Global dispersion and local diversification of the methane seep microbiome. Proc. Natl. Acad. Sci. USA 112, 4015–4020.

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069.

Server, T.M.H.M.M., 2015. v. 2.0. Tmhmm Server v2.0. http://www.cbs.dtu.dk/services/tmhmm.

Wang, B., Chen, X., Xie, Y., Wang, P., Feng, J.C., Zhang, S., 2025. Hydrate formation in porous media with upward-migrating methane and its implications for the evolution of deep-sea cold seep ecosystems. Sci. Total Environ. 959, 178299.