<div align="center">

SUPPLEMENTARY MATERIALS

</div>

Metagenomic Mining of Antimicrobial Biosynthetic Gene Clusters from Extreme Environments: A Systematic Review

S1. Full Electronic Search Strategy

S1.1 PubMed/MEDLINE

Search date: December 15, 2025. No date restrictions applied.

> **#1** metagenomic*[tiab] OR metagenome*[tiab] OR "metagenome-assembled genome*"[tiab] OR MAG[tiab]
>
> **#2** extreme*[tiab] OR thermophil*[tiab] OR halophil*[tiab] OR psychrophil*[tiab] OR "deep sea"[tiab] OR "hot spring*"[tiab] OR hypersaline[tiab] OR hydrothermal[tiab] OR Antarctic[tiab] OR Arctic[tiab] OR desert[tiab] OR biocrust[tiab]
>
> **#3** antimicrobial*[tiab] OR antibiotic*[tiab] OR "biosynthetic gene cluster*"[tiab] OR BGC[tiab] OR "secondary metabolite*"[tiab] OR NRPS[tiab] OR PKS[tiab] OR RiPP*[tiab]
>
> **#4** #1 AND #2 AND #3

Results: 156 records

S1.2 Web of Science Core Collection

Search date: December 16, 2025. All years.

> TS=(metagenomic* OR metagenome* OR "metagenome-assembled genome*" OR MAG) AND TS=(extreme* OR thermophil* OR halophil* OR psychrophil* OR "deep sea" OR "hot spring*" OR hypersaline OR hydrothermal OR Antarctic OR Arctic OR desert OR biocrust) AND TS=(antimicrobial* OR antibiotic* OR "biosynthetic gene cluster*" OR BGC OR "secondary metabolite*" OR NRPS OR PKS OR RiPP*)

Results: 143 records

S1.3 Scopus

Search date: December 16, 2025. All years.

> TITLE-ABS-KEY(metagenomic* OR metagenome* OR "metagenome-assembled genome*" OR MAG) AND TITLE-ABS-KEY(extreme* OR thermophil* OR halophil* OR

psychrophil* OR "deep sea" OR "hot spring*" OR hypersaline OR hydrothermal OR Antarctic OR Arctic OR desert OR biocrust) AND TITLE-ABS-KEY(antimicrobial* OR antibiotic* OR "biosynthetic gene cluster*" OR BGC OR "secondary metabolite*" OR NRPS OR PKS OR RiPP*)

Results: 128 records

S1.4 Google Scholar

Search date: December 17, 2025. First 200 results screened.

metagenomics "biosynthetic gene cluster" extreme environment antimicrobial

Results: 60 unique records (after deduplication with above databases)

Total records identified: 487 (PubMed 156 + WoS 143 + Scopus 128 + Google Scholar 60)

S2. PRISMA 2020 Checklist

Completed per Page et al. (2021). *BMJ*, 372, n71.

Section/Topic	#	Checklist Item	Reported On	Done
Title	1	Identify the report as a systematic review	Title page	✓
Abstract – Background	2a	Provide structured summary including background	Abstract	✓
Abstract – Methods	2b	Describe search, selection, synthesis	Abstract	✓
Rationale	3	Describe rationale in context of existing knowledge	Introduction	✓
Objectives	4	Provide explicit statement of objectives	Introduction §5	✓
Eligibility criteria	5	Specify inclusion/exclusion criteria	Methods 2.1	✓
Information sources	6	Describe all information sources	Methods 2.2	✓
Search strategy	7	Present full search strategies	Methods 2.2; Suppl. S1	✓
Selection process	8	Describe process of study selection	Methods 2.3	✓
Data collection process	9	Describe data extraction methods	Methods 2.3	✓

Data items	10	List all variables for which data were sought	Methods 2.3	✓
Study risk of bias	11	Describe methods for assessing risk of bias	Methods 2.4; Suppl. S3	✓
Effect measures	12	Specify effect measures used	N/A (narrative synthesis)	N/A
Synthesis methods	13a-f	Describe synthesis methods	Methods 2.4	✓
Reporting bias assessment	14	Describe methods for assessing reporting bias	Discussion 4.5	✓
Certainty assessment	15	Describe methods for certainty assessment	Not performed	Partial
Study selection	16a-b	Report study selection results with flow diagram	Results 3.1; Figure 1	✓
Study characteristics	17	Report characteristics of each study	Results 3.2; Table 1	✓
Risk of bias in studies	18	Present risk of bias assessments	Results 3.3; Suppl. S3	✓
Results of individual studies	19	Present results for all outcomes	Results 3.4–3.6	✓
Results of syntheses	20a-d	Present synthesis results	Results 3.4–3.6; Figure 2	✓
Reporting biases	21	Report assessment of reporting bias	Discussion 4.5	✓
Certainty of evidence	22	Present certainty assessments	Not formally assessed	Partial
Discussion – interpretation	23a	Provide interpretation of results	Discussion 4.1–4.4	✓
Discussion – limitations	23b	Discuss limitations of evidence and review	Discussion 4.5	✓
Discussion – other	23c-d	Discuss implications, future research	Discussion; Conclusion	✓
Registration and protocol	24a-c	Provide registration info or state not registered	Methods §1	✓
Support	25	Describe sources of financial support	End matter	✓
Competing interests	26	Declare competing interests	End matter	✓
Availability of data	27	Report data availability	End matter	✓

✓ = fully reported; Partial = partially addressed; N/A = not applicable.

S3. Quality Assessment Using Adapted Newcastle-Ottawa Scale

Adaptation rationale: The Newcastle-Ottawa Scale was adapted for metagenomic discovery studies because no validated quality assessment tool exists for this study type. The original NOS domains (selection, comparability, outcome) were replaced with five domains relevant to metagenomic BGC discovery, each scored 0–2 (maximum total: 10). Risk of bias classification: Low (7–10), Moderate (5–6), High (0–4).

Domain definitions:

D1 – Sample representativeness: Adequacy of sampling strategy, geographic/temporal coverage, and environmental characterization (0 = poor documentation; 1 = partial; 2 = comprehensive site characterization with replicates).

D2 – Sequencing depth: Adequacy of sequencing coverage for metagenomic assembly (0 = not reported; 1 = reported but low; 2 = adequate depth with justification).

D3 – Bioinformatic tool currency: Use of validated, current-version tools with appropriate parameters (0 = outdated/unvalidated; 1 = standard tools; 2 = current tools with parameter optimization).

D4 – BGC completeness assessment: Reporting of BGC completeness metrics and fragment vs. complete BGC distinction (0 = not assessed; 1 = partially reported; 2 = systematically assessed with metrics).

D5 – Experimental validation: Functional confirmation of predicted bioactivity (0 = purely in silico; 1 = indirect evidence such as metatranscriptomics demonstrating active BGC expression; 2 = direct bioassay or compound isolation with confirmed bioactivity).

Table S3. Quality Assessment Scores per Study

Study	D1 Sample	D2 Seq.	D3 Bioinfo.	D4 Compl.	D5 Valid.	Total (/10)	Risk of Bias
Crits-Christoph et al. 2018	2	2	1	1	0	**6**	Moderate
Chen et al. 2020	2	1	1	1	0	**5**	Moderate
Sharrar et al. 2020	2	2	1	1	0	**6**	Moderate
Benaud et al. 2021	2	2	2	2	2	**10**	Low
Van Goethem et al. 2021	2	2	2	2	1	**9**	Low
Waschulin et al. 2022	2	2	2	2	0	**8**	Low
Paoli et al. 2022	2	2	2	1	0	**7**	Low

Sánchez-Navarro et al. 2022	1	2	2	2	0	**7**	Low
Bickhart et al. 2022	2	2	2	1	0	**7**	Low
Xu et al. 2022	2	1	1	1	2	**7**	Low
Basili et al. 2023	2	1	2	1	0	**6**	Moderate
Busi et al. 2023	2	1	1	1	0	**5**	Moderate
Medeiros et al. 2024	2	2	2	1	0	**7**	Low
Andreani-Gerard et al. 2024	2	2	2	1	0	**7**	Low
Rego et al. 2020	1	1	1	1	0	**4**	High

Scoring: 0 = not met/not reported; 1 = partially met; 2 = fully met. Risk classification: Low (7–10), Moderate (5–6), High (0–4). D5 scoring: 0 = purely in silico; 1 = indirect expression evidence (metatranscriptomics); 2 = direct bioassay or compound isolation.

S4. List of Studies Excluded at Full-Text Stage with Reasons (n=74)

Studies excluded after full-text review were categorized into five groups. Representative examples are provided for each category; the complete list is available from the corresponding author upon request.

Category 1: Genome mining of cultured isolates only (n=28)

These studies performed BGC analysis on individual microbial isolates rather than metagenomic datasets. While some originated from extreme environments, they did not employ culture-independent metagenomic approaches as required by inclusion criteria. Examples include genome mining studies of thermophilic *Streptomyces* spp. from hot springs, halophilic archaea from salt lakes, and psychrotolerant bacteria from polar regions.

Category 2: Focus on antimicrobial resistance genes without BGC discovery (n=19)

These studies used metagenomic approaches to characterize antimicrobial resistance gene (ARG) profiles in extreme environments but did not include biosynthetic gene cluster identification or analysis. Common themes included resistome characterization of permafrost, deep-sea sediments, and hot spring communities.

Category 3: Non-extreme environment (n=14)

Studies that performed metagenomic BGC mining but sampled from environments not meeting the extreme criteria defined in Section 2.1. Examples include temperate agricultural soils, freshwater lake sediments, human gut microbiome, and marine environments at standard depths (<200 m) without extreme conditions.

Category 4: Review articles (n=8)

Narrative reviews, perspective articles, and commentary pieces discussing metagenomic BGC discovery without presenting primary data. These were used for background context and reference list hand-searching but excluded from the synthesis.

Category 5: Insufficient BGC characterization data (n=5)

Studies that mentioned BGC detection but did not provide sufficient quantitative or qualitative characterization data (e.g., BGC counts, types, novelty assessment, or completeness metrics) to enable meaningful inclusion in the synthesis.

S5. Sensitivity Analysis

To assess the robustness of the primary findings, sensitivity analyses were conducted by excluding studies with arguable inclusion criteria.

Analysis 1: Excluding arguable extreme environments (removing Sánchez-Navarro et al. [2022] and Bickhart et al. [2022])

When these two studies were excluded (n=13 remaining), the three principal findings remained unchanged: (a) long-read sequencing superiority persisted (Waschulin 60%, Van Goethem 24% vs. Medeiros ~2%); (b) BGC novelty rates remained at 60–99% across all environment categories, consistent with the range reported across the full dataset; (c) the validation structure was preserved: 2 of 13 studies (15.4%) achieved direct bioactivity confirmation (Benaud, Xu), 1 study (7.7%) provided indirect expression evidence (Van Goethem), and 10 of 13 studies (76.9%) relied entirely on *in silico* prediction. The exclusion of Sánchez-Navarro et al. removed the highest completeness data point (88%), but the long-read advantage was still clearly demonstrated by remaining studies.

Analysis 2: Excluding Xu et al. (2022) due to culture-dependent component

Xu et al. (2022) combined metagenomic analysis with culture-dependent compound isolation from a deep-sea hydrothermal vent *Streptomyces*. When excluded (n=14 remaining), the validation structure became: 1 study (7.1%) with direct bioactivity confirmation (Benaud), 1 study (7.1%) with indirect expression evidence (Van Goethem), and 12 of 14 studies (85.7%) relying entirely on *in silico* prediction. However, this study's inclusion was justified because it employed metagenomic BGC analysis as the discovery component before targeted isolation, representing a legitimate integrated pipeline approach.

Analysis 3: Excluding Paoli et al. (2022) from qualitative synthesis

Paoli et al. (2022) reported ~40,000 BGCs across all global ocean metagenomes. Given that only an undefined subset met extreme-environment criteria, we assessed findings without this study. The >14,000 cumulative BGC count was already calculated excluding Paoli. The marine/deep-sea category reduced to 3 studies, but the qualitative conclusions regarding marine BGC diversity and novelty remained supported by Xu et al., Basili et al., and Bickhart et al.

Conclusion of sensitivity analyses: All three principal findings (long-read superiority, high novelty [60–99% range], validation gap) were robust to the exclusion of studies with arguable inclusion criteria. The magnitude of specific estimates changed marginally, but the direction and interpretation of all findings remained consistent.