Electronic Journal of Biotechnology 78 (2025) 46–52

Contents lists available at ScienceDirect

Electronic Journal of Biotechnology

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

Research article

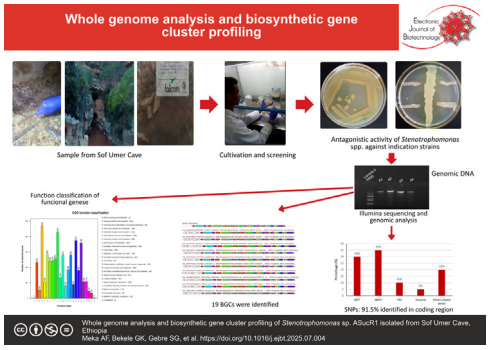
Whole-genome analysis and biosynthetic gene cluster profiling of *Stenotrophomonas* sp. ASucR1 isolated from Sof Umer Cave, Ethiopia ☆

Abu Feyisa Meka [a,b,*], Gessesse Kebede Bekele [a,c], Selfu Girma Gebre [a,d], Musin Kelel Abas [a,c], Mesfin Tafesse Gemeda [a,c,*]

[a] *Department of Biotechnology, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia*
[b] *Department of Biology, Bule Hora University, Bule Hora, Ethiopia*
[c] *Biotechnology and Bioprocess Centre of Excellence, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia*
[d] *Armauer Hansen Research Institute, Addis Ababa, Ethiopia*

GRAPHICAL ABSTRACT

Whole genome analysis and biosynthetic gene cluster profiling of *Stenotrophomonas* sp. ASucR1 isolated from Sof Umer Cave, Ethiopia



ARTICLE INFO

Article history:
Received 29 April 2025
Accepted 14 July 2025
Available online 16 September 2025

Keywords:
Biosynthetic gene clusters
Secondary metabolites
Stenotrophomonas sp.
Whole-genome sequencing

ABSTRACT

Background: Sof Umer Cave is a unique habitat that hosts industrially significant microbes. In this study, *Stenotrophomonas* sp. ASucR1 was isolated from the cave rock and screened for antimicrobial activity. High-molecular-weight genomic DNA was extracted and subjected to whole-genome sequencing using the Illumina NovaSeq platform. Comprehensive genomic and biosynthetic gene cluster (BGC) profiling was conducted.

Results: In vitro tests revealed that *Stenotrophomonas* sp. ASucR1 exhibited a broad spectrum of antagonistic activity. Functional genome annotation identified diverse biosynthetic gene clusters (BGCs) and metabolic pathways, including genes involved in the synthesis of secondary metabolites. A total of 19 BGCs were identified, several of which showed no matches in the minimum information about a biosynthetic gene cluster (MiBIG) database, indicating the presence of previously uncharacterized bioactive compounds. Single-nucleotide polymorphism (SNP) analysis showed that 91.5% of variants were identified within coding regions, with 85.84% being synonymous. Classification of SNPs and insertion-deletion mutations through clusters of orthologous groups (COG) analysis highlighted their association with key biological functions.

☆ Audio abstract available in Supplementary material.

Peer review under responsibility of Pontificia Universidad Católica de Valparaíso.

* Corresponding authors.
E-mail addresses: abufeyisa21@gmail.com (A.F. Meka), mesfin.tafesse@aastu.edu.et (M.T. Gemeda).

https://doi.org/10.1016/j.ejbt.2025.07.004

Conclusions: This study highlights the metabolic versatility and biosynthetic potential of *Stenotrophomonas* sp. ASucR1, a promising candidate for antimicrobial development and biotechnological applications. The identification of various biosynthetic gene clusters paves the way for exploring bioactive compounds with pharmaceutical significance.

How to cite: Meka AF, Bekele GK, Gebre SG, et al. Whole genome analysis and biosynthetic gene cluster profiling of *Stenotrophomonas* sp. ASucR1 isolated from Sof Umer Cave, Ethiopia. Electron J Biotechnol 2025;78. https://doi.org/10.1016/j.ejbt.2025.07.004.

1. Introduction

Microorganisms that inhabit extreme and underexplored environments have gained increasing attention for their potential to produce novel bioactive compounds [1,2]. Caves, in particular, represent a rich yet unexplored ecosystem [3]. Environments are characterized by darkness, high humidity, and low nutrient availability, which provide unique selective pressures that promote the evolution of microbial communities with specialized metabolic capabilities [4]. Recent metagenomic and culture-based studies have shown that microbial isolates from cave ecosystems possess diverse biosynthetic gene clusters (BGCs) responsible for the production of antimicrobial, antifungal, anticancer, and antioxidant agents [4,5].

The genus *Stenotrophomonas* includes Gram-negative, aerobic, nonfermentative bacteria that are typically found in soil, water, and environments associated with plants [6,7]. Some species, especially *Stenotrophomonas maltophilia*, are recognized as opportunistic pathogens in humans [7]. However, many environmental isolates of *Stenotrophomonas* are valued for their biotechnological potential, playing important roles in bioremediation, promoting plant growth, and synthesizing bioactive secondary metabolites [8]. Despite these benefits, the genus remains underexplored regarding the biosynthesis of secondary metabolites, from unique ecological ecosystem like caves [9].

Whole-genome sequencing (WGS) has transformed microbial genomics by allowing for a thorough analysis of genetic content, encompassing metabolic pathways, resistance genes, and biosynthetic gene clusters (BGCs) [10]. Tools such as antibiotics and secondary metabolite analysis shell (antiSMASH) pipeline [11] and Prokka [12] help in the in-silico identification and functional annotation of BGCs. This provides insights into the biosynthetic potential of microorganisms without the need for initial metabolite extraction or purification. Additionally, variant calling and comparative genomics enable researchers to explore the adaptive evolution of strains living in extreme environments [13].

In Ethiopia, Sof Umer Cave is one of the largest and most ecologically diverse cave systems, yet its microbial biodiversity remains largely unexplored. Recognizing the ecological uniqueness of this habitat and the potential of cave microbiomes as sources of new biomolecules, we aimed to isolate and characterize a bacterial strain from this site with promising biotechnological applications. In this study, we report the isolation, genomic sequencing, and biosynthetic gene cluster profiling of *Stenotrophomonas* sp. ASucR1. Our objectives were to explore the organism's genetic potential for secondary metabolite production, understand its evolutionary adaptations through single-nucleotide polymorphisms (SNPs) and insertion-deletions (InDels) analysis, and evaluate its suitability as a candidate for antimicrobial drug discovery.

2. Materials and methods

2.1. Sample collection and isolation

Rock samples were aseptically collected from the interior walls of Sof Umer Cave using sterile spatulas and transported to the laboratory in sterile containers under refrigerated conditions. Samples were processed within 48 h of collection. Serial dilution and spread plating were performed on nutrient agar and soil extract agar to isolate distinct colonies. Plates were incubated at 28°C for 72 h. Colonies showing unique morphology were sub-cultured repeatedly to obtain pure isolates (Fig. S1A).

2.2. Preliminary antimicrobial screening

The pure isolate of ASucR1 was tested for antimicrobial activity using the perpendicular streak method against standard strains of *Escherichia coli* ATCC 25922, *Pseudomonas aeruginosa* ATCC 27853, *Staphylococcus aureus* ATCC 25923, and *Candida albicans* ATCC 10231. After a 24-h incubation at 37°C, the zones of inhibition were measured (Fig. S1C).

2.3. Genomic DNA extraction and quality assessment

Bacterial cells were cultured in LB broth at 28°C with shaking at 150 rpm for 48 h. Genomic DNA was extracted using the Qiagen DNeasy Blood & Tissue Kit, following the manufacturer's instructions. DNA purity and concentration were determined using a Thermo Scientific NanoDrop 3300 Fluorospectrometer. DNA integrity was evaluated through 1% agarose gel electrophoresis (Fig. S2).

2.4. Library preparation and whole-genome sequencing

DNA libraries were prepared using the Nextera DNA Flex Library Preparation Kit. The libraries were then validated with an Agilent Bioanalyzer and sequenced on an Illumina NovaSeq 6000 platform, producing 150 bp paired-end reads. Read quality metrics were evaluated using FastQC 0.12.0 [14]. For quality filtering and adapter trimming, Trimmomatic [15] was used (Fig. S3).

2.5. Genome assembly and quality metrics

Clean reads were assembled using SPAdes v3.15.2 [16]. The quality of the assembly was assessed with QUAST [17] to obtain metrics including N50, the number of contigs, and genome size. Genome completeness and contamination were evaluated using CheckM [18].

2.6. Read alignment and coverage analysis

Filtered reads were mapped to the assembled genome using BWA-MEM [19]. SAMtools [20] was used to calculate mapping statistics, including read depth. Coverage across the genome was determined using BEDTools [21].

2.7. Variant calling and annotation

Single-nucleotide polymorphisms (SNPs) and insertion-deletions (InDels) were identified using the genome analysis

toolkit (GATK) [22]. SnpEff was used to annotate the variants and predict their functional consequences [23]. Variants were categorized based on coding effect (synonymous, non-synonymous, stop-gain, stop-loss) [24].

2.8. Biosynthetic gene cluster identification

Antibiotics and secondary metabolite analysis shell (antiSMASH) v7.0 [11] was used to predict and annotate biosynthetic gene clusters (BGCs) based on the assembled genome. Cluster types were categorized as polyketide synthase (PKS), nonribosomal peptide synthetase (NRPS), ribosomally synthesized and post-translationally modified peptides (RiPPs), and terpene, and homology searches were performed against the minimum information about a biosynthetic gene cluster (MIBiG) database to assess novelty [25].

2.9. Functional gene annotation

Genome annotation was conducted using Prokka [12]. Functional categorization was performed using Blast2GO [26], KEGG Mapper [27], and eggNOG-mapper [28] to assign genes to known metabolic and regulatory pathways. Cluster of orthologous groups (COG) classification was used to determine the distribution of genes with SNPs and InDels across functional categories [29].

2.10. Data visualization and statistical analysis

OriginPro [30] was used for plotting SNP distributions, cluster of orthologous groups (COG) functional categories, and BGC counts. Descriptive statistics were computed to summarize assembly metrics and variant impacts.

3. Results and discussion

3.1. Antagonistic activity

Stenotrophomonas sp. ASucR1 exhibited significant antagonistic activity against four clinical isolates. Inhibition zones measured for *Escherichia coli* (15.3 mm), *Staphylococcus aureus* (15.5 mm), *Pseudomonas aeruginosa* (12.7 mm), and *Candida albicans* (10.2 mm) (Fig. S1B–C). These findings suggest that *Stenotrophomonas* sp. ASucR1 produces bioactive compounds with broad-spectrum antimicrobial activity, likely due to its adaptation to the nutrient-poor cave environment, which promotes secondary metabolite biosynthesis [31,32].

3.2. Genomic DNA quality

The extracted genomic DNA showed high molecular weight and purity, with A260/A280 ratios between 1.80 and 2.03 (Table S1). The presence of intact bands on the agarose gel confirmed the DNA's suitability for downstream sequencing applications [33].

3.3. Sequencing output and assembly statistics

Whole-genome sequencing generated 8,204,852 raw paired-end reads. After quality trimming, 98.59% of reads were retained, with a Q20 score of 98.03% (Table 1 and Table S2), indicating high sequence quality. The genome assembly produced an N50 of 150,000 bp and a total assembly length of approximately 4.7 Mb, consistent with other Stenotrophomonas species [34,35]. Read mapping achieved an average depth of 112.87X and 85.2% genome coverage (Table 2). CheckM analysis revealed 98.5% genome completeness with 1.2% contamination, supporting the reliability of the dataset for further functional analysis [36].

3.4. SNP and InDel annotation

A total of 17,238 SNPs were detected, of which 91.5% occurred in coding sequences (Fig. 1A). Among these, 85.84% were synonymous, and 13.7% were non-synonymous, potentially impacting protein structure and function (Fig. 1B). A small proportion (0.05%) represented stop-gain and stop-loss mutations. These variants were predominantly associated with genes involved in cell wall biosynthesis, membrane transport, and oxidative stress responses [37]. InDel mutations, though fewer, were found to affect genes related to signal transduction and transcriptional regulation. The prevalence of synonymous mutations suggests evolutionary pressure to maintain protein function while permitting limited adaptive changes [38].

3.5. Functional annotation and cluster of orthologous groups classification

Cluster of orthologous group (COG)-based functional categorization of SNP and InDel-affected genes showed enrichment in

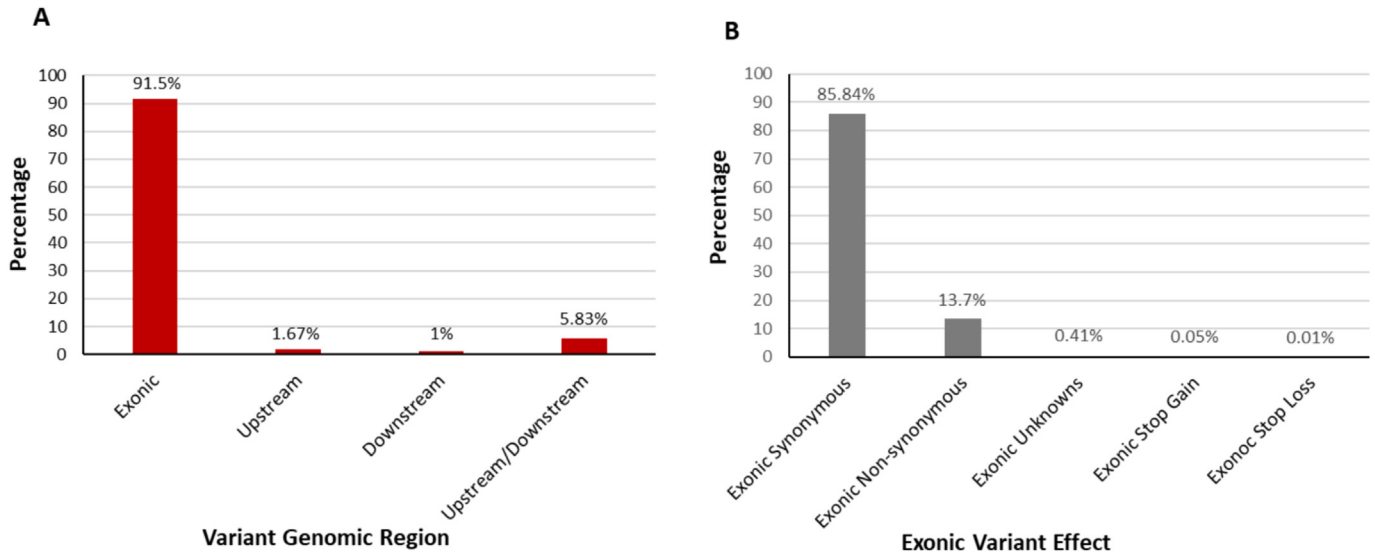


Fig. 1. SNP distribution and functional **impact in the *Stenotrophomonas* sp. ASucR1 genome.** (A) SNP distribution across genomic regions is as follows: upstream (1.67%), downstream (1%), upstream/downstream (5.83%), and exonic regions (91.5%). (B) Functional classification of exonic SNPs reveals that the majority are synonymous (85.84%), followed by non-synonymous mutations (13.7%). Stop-gain mutations account for 0.05%, stop-loss for 0.01%, and unknown exonic variants for 0.41%.

Table 1
Statistics of sequencing quality metrics for Stenotrophomonas sp. strain ASucR1.

Code	Raw reads	Clean data (G)	Effective (%)	Q20 (%)	GC (%)
ASucR1	8,204,852	1.2	98.59	98.03	61.39

Table 2
Mapping and coverage statistics for *Stenotrophomonas* sp. strain ASucR1.

Sample	Mapped reads	Total reads	Mapping rate (%)	Average depth(X)	Coverage at least 1X (%)
ASucR1	4,460,684	8,089,342	55.14	112.87	85.20

metabolic pathways (22%), transport and catabolism (18%), and secondary metabolite biosynthesis (15%) (Fig. 2). Mutations in genes associated with replication, recombination, and repair further suggest an active genome capable of maintaining integrity under stress conditions [38]. These results highlight the strain's adaptability to nutrient-poor cave environments.

3.6. Biosynthetic gene cluster analysis

Antibiotics and secondary metabolite analysis shell (antiSMASH) analysis predicted 19 biosynthetic gene clusters (BGCs), including non-ribosomal peptide synthetases (NRPS) 35%, ribosomally synthesized and post-translationally modified peptides (RiPPs) 30 %, polyketide synthases (PKS) of types I and II (10%), terpenes (5%), and unknown clusters (20%). Notably, some of these BGCs showed high similarity to known clusters that encode antimicrobial agents, such as bacillibactin and enterobactin (Table S3, Table S4). In contrast, two hybrid clusters, among others, displayed no close matches in the MIBiG database (Fig. 3), indicating a novel biosynthetic potential [39]. These uncharacterized clusters may produce unique secondary metabolites, which warrant future structural and functional validation [40].

Comparative analysis of homologous BGCs revealed over 82% conservation with reference strains, underscoring the evolutionary preservation of core secondary metabolite production of *Stenotrophomonas* sp. AScuR1 [41,42]. The query sequence exhibits 84% gene similarity with nine out of the ten target sequences, demonstrating strong conservation of the BGC across *Stenotrophomonas* strains. However, the final sequence, *Stenotrophomonas maltophilia* KJ (NC_015947) and other strains, shows 84% gene similarity, suggesting minor variations in gene content or arrangement (Fig. 4, Fig. S4). The high sequence conservation of certain BGCs across different *Stenotrophomonas* strains indicates their evolutionary significance, potentially enhancing the ecological fitness of these organisms [9,43].

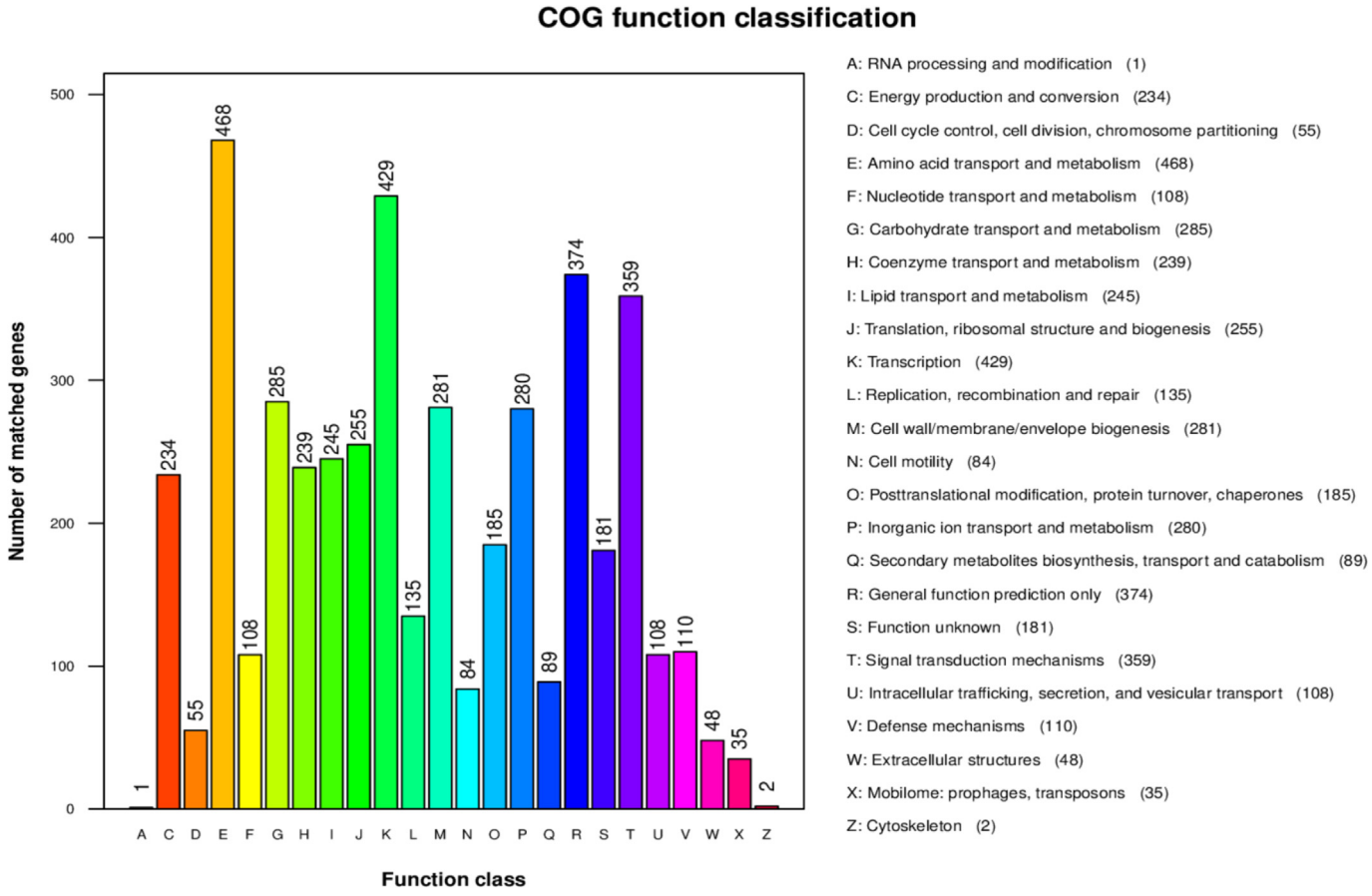


Fig. 2. Cluster of Orthologous Groups Function Classification of genes and annotation of SNP gene.

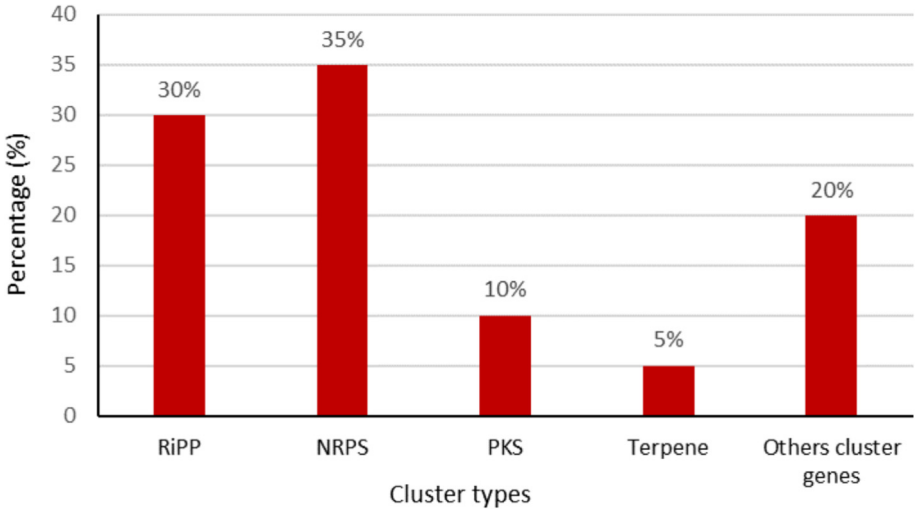


Fig. 3. Distribution of Biosynthetic Gene Cluster (BGC) categories in *Stenotrophomonas* ASucR1 isolate. The graph shows the proportions of various biosynthetic gene cluster types identified in the analyzed *Stenotrophomonas* ASucR1 isolate.

Fig. 4. Putative region of Biosynthetic Gene Cluster in *Stenotrophomonas* sp. ASucR1 isolate. The alignment, conducted via Cluster-BLAST, revealed gene clusters that are related to the queried gene cluster. Genes with significant homology (BLAST; highest sequence identity of 84%; BLAST alignment covering 100% of the sequence).

3.7. Comparative genomics insight

When compared to other Stenotrophomonas genomes, ASucR1 showed a higher number of predicted BGCs, possibly due to its cave origin. Cave environments exert selective pressures that may drive the diversification of biosynthetic genes. High GC content (61.39%) further correlates with genome stability and the presence of regulatory elements critical for metabolic control [44,45]. The genomic profile supports ASucR1's classification as a metabolically versatile environmental strain.

3.8. Ecological and biotechnological implications

The discovery of multiple and novel BGCs, coupled with antimicrobial activity, underscores ASucR1's potential in natural product discovery. The strain's ability to adapt to the cave's oligotrophic

conditions suggests a rich enzymatic repertoire for survival and competition. As a non-pathogenic species with biosynthetic diversity, ASucR1 is a strong candidate for further exploration in pharmaceutical and industrial biotechnology [45].

4. Conclusions

This study presents a genomic analysis of *Stenotrophomonas* sp. ASucR1, a bacterial strain isolated from the oligotrophic environment of Sof Umer Cave, Ethiopia. The organism demonstrated strong antimicrobial activity against a broad range of pathogenic microbes, indicating its potential as a source of novel bioactive compounds. High-throughput sequencing and bioinformatics analyses revealed a diverse array of biosynthetic gene clusters, including non-ribosomal peptide synthetases, polyketide synthases, and several uncharacterized clusters with no homology in current databases, underscoring its potential to produce previously unknown secondary metabolites. The identification of extensive SNP and InDel variants, particularly those affecting functional genes involved in metabolism, stress response, and signal transduction, suggests that ASucR1 has undergone genomic adaptations enabling survival in the nutrient-limited cave environment. Functional annotation and COG classification further highlight the organism's metabolic versatility and evolutionary resilience. These findings indicate that ASucR1 is a metabolically robust and excellent candidate for further functional studies. Its rich biosynthetic repertoire and genomic stability make it an excellent candidate for further functional studies, including metabolite purification, structural characterization, and heterologous expression of cryptic gene clusters. Furthermore, this work contributes to the field of cave microbiome exploration, emphasizing the importance of understudied environments in the search for next-generation antibiotics and industrially relevant enzymes.

CRediT authorship contribution statement

Abu Feyisa Meka: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Data curation, Conceptualization. **Gessesse Kebede Bekele:** Writing – review & editing, Visualization, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Selfu Girma Gebre:** Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Formal analysis, Conceptualization. **Musin Kelel Abas:** Writing – review & editing, Visualization, Validation, Supervision, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Mesfin Tafesse Gemeda:** Writing – review & editing, Visualization, Validation, Supervision, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Financial support

The authors extend their sincere gratitude to Bule Hora University and Addis Ababa Science and Technology University for their generous financial support, which enabled the completion of this research.

Declaration of competing interest

The authors declare that this research was conducted without any financial affiliations that could be perceived as potential conflicts of interest.

Supplementary material

https://doi.org/10.1016/j.ejbt.2025.07.004.

Data availability

The data that have been used are confidential. The sequence data supporting the findings of this study have been submitted to the NCBI Sequence Read Archive (SRA) under accession number PRJNA1198506 (https://www.ncbi.nlm.nih.gov/sra/PRJNA1198506).

References

[1] Cyske Z, Jaroszewicz W, Żabińska M, et al. Unexplored potential: Biologically active compounds produced by microorganisms from hard-to-reach environments and their applications. Acta Biochim Pol 2021;68(4):565–74. https://doi.org/10.18388/abp.2020_5887. PMid: 34536268.
[2] Meka AF, Bekele GK, Abas MK, et al. Exploring microbial diversity and functional gene dynamics associated with the microbiome of Sof Umer cave, Ethiopia. Discov Appl Sci 2024;6:400. https://doi.org/10.1007/s42452-024-06110-x.
[3] Khatri A, Kumar K, Thakur IS. Microbiome of caves for bioprospecting: A critical review. Syst Microbiol Biomanuf 2025;5:550–66. https://doi.org/10.1007/s43393-024-00322-3.
[4] Samanta B, Sharma S, Budhwar R. Metagenome analysis of speleothem microbiome from subterranean cave reveals insight into community structure, metabolic potential, and BGCs diversity. Curr Microbiol 2023;80:317. https://doi.org/10.1007/s00284-023-03431-9. PMid: 37561193.
[5] Meka AF, Bekele GK, Abas MK, et al. Exploring bioactive compound origins: Profiling gene cluster signatures related to biosynthesis in microbiomes of Sof Umer Cave, Ethiopia. PLoS ONE 2025;20(3):e0315536. https://doi.org/10.1371/journal.pone.0315536. PMid: 40048434.
[6] Mukherjee P, Roy P. Genomic potential of *Stenotrophomonas maltophilia* in bioremediation with an assessment of its multifaceted role in our environment. Front Microbiol 2016;7:967. https://doi.org/10.3389/fmicb.2016.00967.
[7] Brooke JS. Advances in the microbiology of *Stenotrophomonas maltophilia*. Clin Microbiol Rev 2021;34(3):e00030–e119. https://doi.org/10.1128/CMR.00030-19. PMid: 34043457.
[8] Farooq A, Kanwal R, Bashir K, et al. Multi-omics analysis of *Stenotrophomonas maltophilia* S-11 reveals its potential for Pb^{2+} bioremediation in contaminated soil. J Hazard Mater 2025;495:138867. https://doi.org/10.1016/j.jhazmat.2025.138867. PMid: 40499424.
[9] Ryan R, Monchy S, Cardinale M, et al. The versatility and adaptation of bacteria from the genus *Stenotrophomonas*. Nat Rev Microbiol 2009;7:514–25. https://doi.org/10.1038/nrmicro2163. PMid: 19528958.
[10] Pandey A, Israr J, Pandey J, et al. Current approaches and implications in discovery of novel bioactive products from microbial sources. Curr Microbiol 2025;82:258. https://doi.org/10.1007/s00284-025-04237-7. PMid: 40263159.
[11] Blin K, Shaw S, Augustijn HE, et al. antiSMASH 7: New features for the secondary metabolite genome mining pipeline. Nucleic Acids Res 2023;51(W1):W46–50. https://doi.org/10.1093/nar/gkad344. PMid: 37140036.
[12] Seemann T. Prokka: Rapid prokaryotic genome annotation. Bioinformatics 2014;30(14):2068–9. https://doi.org/10.1093/bioinformatics/btu153. PMid: 24642063.
[13] Lin H, Yu M, Wang X, et al. Comparative genomic analysis reveals the evolution and environmental adaptation strategies of vibrios. BMC Genomics 2018;19:135. https://doi.org/10.1186/s12864-018-4531-2. PMid: 29433445.
[14] Utturkar S, Dassanayake A, Nagaraju S. Bacterial differential expression analysis methods. In: Himmel M, Bomble Y, editors. Metabolic Pathway Engineering. Methods Mol Biol. p. 135–51. https://doi.org/10.1007/978-1-0716-0195-2_8. PMid: 32720149.
[15] Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 2014;30(15):2114–20. https://doi.org/10.1093/bioinformatics/btu170 PMid: 24695404.
[16] Prjibelski A, Antipov D, Meleshko D, et al. Using SPAdes De Novo Assembler. Curr Protoc Bioinform 2020;70(1):e102. https://doi.org/10.1002/cpbi.102. PMid: 32559359.
[17] Gurevich A, Saveliev V, Vyahhi N, et al. QUAST: Quality assessment tool for genome assemblies. Bioinformatics 2013;29(8):1072–5. https://doi.org/10.1093/bioinformatics/btt086. PMid: 23422339.
[18] Parks DH, Imelfort M, Skennerton CT, et al. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 2015;25:1043–55. https://doi.org/10.1101/gr.186072.114. PMid: 25977477.
[19] Yao Z, You FM, N'Diaye A, et al. Evaluation of variant calling tools for large plant genome re-sequencing. BMC Bioinformatics 2020;21:360. https://doi.org/10.1186/s12859-020-03704-1. PMid: 32807073.
[20] Danecek P, Bonfield JK, Liddle J, et al. Twelve years of SAMtools and BCFtools. GigaScience 2021;10(2):giab008. https://doi.org/10.1093/gigascience/giab008. PMid: 33590861.
[21] Quinlan AR. BEDTools: The Swiss-army tool for genome feature analysis. Curr Protoc Bioinform 2014;47(1):11.12.1–11.12.34. https://doi.org/10.1002/0471250953.bi1112s47. PMid: 25199790.
[22] Saleh AA, Xue L, Zhao Y. Screening *Indels* from the whole genome to identify the candidates and their association with economic traits in several goat breeds. Funct Integr Genomics 2023;23:58. https://doi.org/10.1007/s10142-023-00981-w. PMid: 36757519.

[23] Cingolani P. Variant annotation and functional prediction: SnpEff. In: Ng C, Piscuoglio S, editors. Variant Calling. Methods Mol Biol. p. 289–307. https://doi.org/10.1007/978-1-0716-2293-3_19. PMid: 35751823.
[24] Karunakaran C, Niranjan V, Setlur AS, et al. Exploring the role of non-synonymous and deleterious variants identified in colorectal cancer: A multi-dimensional computational scrutiny of exomes. Curr Genomics 2024;25 (1):41–64. https://doi.org/10.2174/0113892029285310231227105503. PMid: 38544823.
[25] Zada S, Sajjad W, Rafiq M, et al. Cave microbes as a potential source of drugs development in the modern era. Microb Ecol 2022;84:676–87. https://doi.org/10.1007/s00248-021-01889-3. PMid: 34693460.
[26] Conesa A, Götz S, García-Gómez JM, et al. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 2005;21(18):3674–6. https://doi.org/10.1093/bioinformatics/bti610. PMid: 16081474.
[27] Kanehisa M, Sato Y. KEGG Mapper for inferring cellular functions from protein sequences. Protein Sci 2020;29(1):28–35. https://doi.org/10.1002/pro.3711. PMid: 31423653.
[28] Cantalapiedra CP, Hernández-Plaza A, Letunic I, et al. eggNOG-mapper v2: Functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol Biol Evol 2021;38(12):5825–9. https://doi.org/10.1093/molbev/msab293. PMid: 34597405.
[29] Galperin MY, Wolf YI, Makarova KS, et al. COG database update: Focus on microbial diversity, model organisms, and widespread pathogens. Nucleic Acids Res 2021;49(D1):D274–81. https://doi.org/10.1093/nar/gkaa1018. PMid: 33167031.
[30] Stevenson KJ. Review of OriginPro 8.5. J Am Chem Soc 2011;133(14):5621. https://doi.org/10.1021/ja202216h.
[31] Gatinho P, Salvador C, Gutierrez-Patricio S, et al. From cultural and natural heritage to reservoir of biomedicine: Prospection of bioactive compounds produced by bacterial isolates from caves. Int Biodeterior Biodegradation 2024;190:105773. https://doi.org/10.1016/j.ibiod.2024.105773.
[32] Pipite A, Lockhart PJ, McLenachan PA, et al. Isolation, antibacterial screening, and identification of bioactive cave dwelling bacteria in Fiji. Front Microbiol 2022;13:1012867. https://doi.org/10.3389/fmicb.2022.1012867. PMid: 36605510.
[33] Obrador-Sanchez JA, Tzec-Sima M, Higuera-Ciapara I, et al. A quick and effective in-house method of DNA purification from agarose gel, suitable for sequencing. 3 Biotech 2017;7:180. https://doi.org/10.1007/s13205-017-0851-1. PMid: 28664367.
[34] Narzisi G, Mishra B. Comparing de novo genome assembly: The long and short of it. PLoS ONE 2011;6(4):e19175. https://doi.org/10.1371/journal.pone.0019175. PMid: 21559467.
[35] Chklovski A, Parks DH, Woodcroft BJ, et al. CheckM2: A rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. Nat Methods 2023;20:1203–12. https://doi.org/10.1038/s41592-023-01940-w. PMid: 37500759.
[36] Krishnamurthy HK, Rajavelu I, Pereira M, et al. Inside the genome: Understanding genetic influences on oxidative stress. Front Genet 2024;15:1397352. https://doi.org/10.3389/fgene.2024.1397352. PMid: 38983269.
[37] Zhang J, Qian W. Functional synonymous mutations and their evolutionary consequences. Nat Rev Genet 2025. https://doi.org/10.1038/s41576-025-00850-1.
[38] Terlouw BR, Blin K, Navarro-Muñoz JC, et al. MIBiG 3.0: A community-driven effort to annotate experimentally validated biosynthetic gene clusters. Nucleic Acids Res 2023;51(D1):D603–10. https://doi.org/10.1093/nar/gkac1049. PMid: 36399496.
[39] Blin K, Kim HU, Medema MH, et al. Recent development of antiSMASH and other computational approaches to mine secondary metabolite biosynthetic gene clusters. Brief Bioinform 2019;20(4):1103–13. https://doi.org/10.1093/bib/bbx146. PMid: 29112695.
[40] Zhao Y, Ding WJ, Xu L, et al. A comprehensive comparative genomic analysis revealed that plant growth promoting traits are ubiquitous in strains of *Stenotrophomonas*. Front Microbiol 2024;15:1395477. https://doi.org/10.3389/fmicb.2024.1395477. PMid: 38817968.
[41] Ulrich K, Kube M, Becker R, et al. Genomic analysis of the endophytic *Stenotrophomonas* strain 169 reveals features related to plant-growth promotion and stress tolerance. Front Microbiol 2021;12:687463. https://doi.org/10.3389/fmicb.2021.687463. PMid: 34220780.
[42] Ercole TG, Kava VM, Petters-Vandresen DAL, et al. Unveiling agricultural biotechnological prospects: The draft genome sequence of *Stenotrophomonas geniculata* LGMB417. Curr Microbiol 2024;81:247. https://doi.org/10.1007/s00284-024-03784-9. PMid: 38951210.
[43] Galià-Camps C, Pegueroles C, Turon X, et al. Genome composition and GC content influence loci distribution in reduced representation genomic studies. BMC Genomics 2024;25:410. https://doi.org/10.1186/s12864-024-10312-3. PMid: 38664648.
[44] Hu EZ, Lan XR, Liu ZL, et al. A positive correlation between GC content and growth temperature in prokaryotes. BMC Genomics 2022;23:110. https://doi.org/10.1186/s12864-022-08353-7. PMid: 35139824.
[45] Sarkar D, Nanda S, Poddar K, et al. Isolation and characterization of an antibacterial compound producing *Stenotrophomonas* strain from sewage water, production optimization, and its antibiotic potential evaluation. Environ Qual Manag 2022;31(4):51–62. https://doi.org/10.1002/tqem.21764.