**Microbiome of Pukzing Cave in India shows high antimicrobial activity against plant and animal pathogens**

<https://doi.org/10.1016/j.ygeno.2021.10.004>[Get rights and content](https://s100.copyright.com/AppDispatchServlet?publisherName=ELS&contentID=S0888754321003670&orderBeanReset=true)

Under a Creative Commons [license](http://creativecommons.org/licenses/by-nc-nd/4.0/)

*open access*

**Highlights**

* •

Study revealed Pukzing cave ecosystem as a unique source of endemic and moderate thermophilic microorganisms.

* •

Amplicon sequencing of PKSI, PKSII and NRP specific genes revealed presence of AMP genes in the microbial population.

* •

Cave environment harbours unique microbial flora and hypervariable region V4 was found to be more informative.

* •

The study concludes that cave microbial communities could be potential source of future genomic resources.

**Abstract**

Pukzing cave, the largest cave of Mizoram, India was explored for bacterial diversity. Culture dependent method revealed 235 bacterial isolates using three different treatments. Identity of the microbial species was confirmed by 16S rDNA sequencing. The highest bacterial population was recovered from heat treatment (*n* = 97;41.2%) followed by normal (*n* = 79;33.6%) and cold treatment (*n* = 59;25.1%) indicating dominance of moderate [thermophiles](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/thermophile). Antimicrobial potential of isolates showed 20.4% isolates having antimicrobial ability against tested pathogens. [Amplicon sequencing](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/amplicon-sequencing) of PKSI, PKSII and NRP specific genes revealed presence of [AMP](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/antimicrobial-peptides) genes in the microbial population. Six microbial pathogens were selected for screening as they are well known for different disease cause organism in various fields such as agriculture and human health. Cave environment harbors unique [microbial flora](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/microflora) and [hypervariable region](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/hypervariable-region) V4 is more informative. Higher activity of AMP assay against these microbes indicates that cave microbial communities could be potential source of future genomic resources.

* [Previous article in issue](https://www.sciencedirect.com/science/article/pii/S0888754321003724)
* [Next article in issue](https://www.sciencedirect.com/science/article/pii/S0888754321003785)

**Keywords**

Antimicrobial peptides

Biosynthetic genes

Cave microbial diversity

Moderate thermophiles

16s rDNA

**Abbreviations**

ABC

ATP-Binding Cassettes

AICc

Akaike Information Criterion, corrected

AMP

Antimicrobial Peptide

ANOVA

Analysis of Variance

BC

Bray-Curtis

BIC

Bayesian Information Criterion

CHN

Carbon Hydrogen and Nitrogen

COG

Clusters of Orthologous Groups

DNA

Deoxyribonucleic acid

HVR

Hyper Variable Regions

KMG

Koat Maqbari Ghaar

KO

KEGG Orthology

MEGA

Molecular Evolutionary Genetics Analysis

MEGAN

MEtaGenome Analyzer

MG-RAST

Metagenomic Rapid Annotations using Subsystems Technology

MTCC

Microbial Type Culture Collection

NA

Nutrient Agar

NCBI

National Center for Biotechnology Information

NOG

Non-supervised Orthologous Groups

NRPS

Non-ribosomal Peptide Synthetase

OTU

Operational Taxonomic Unit

PCR

Polymerase chain reaction

PKS

Polyketide synthases

SG

Smasse-Rawo Ghaar

TOC

Total Organic Carbon

TRAP

Tripartite ATP-independent periplasmic transporters

QIIME

Quantitative Insights into Microbial Ecology

**1. Introduction**

Cave ecosystem is one of the poorly explored ecosystems on earth and is considered as an extreme environment, not appropriate for the development of life due to extreme abiotic conditions [[35](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0175),[70](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0350)]. Extreme environments of cave have been deemed one of the most promising sources of antimicrobial compounds [[13](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0065)]. Though, majority of caves possess oligotrophic ecosystem with less than 2 mg of [total organic carbon](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/total-organic-carbon) (TOC)/liter, availability of very less light, low temperature and high humidity, the cave environment is considered as specific niche for specialized group of microorganisms including endemic as well as moderate [thermophiles](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/thermophile) [[24](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0120),[63](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0315),[73](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0365)]. Being oligotrophic conditions, the microbial population associated with caves is 106 cells/g of rock [[8](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0040)]. Studies predicted the potential of cave [microbial flora](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/microflora) and their role in geological processes with significant potential to produce bioactive molecules of biotechnological importance [[70](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0350)]. At the same time, the cave associated microbial communities also influence the formation and preservation of cave deposits and serve as primary producers for the complex organisms [[6](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0030)]. [Microbial interactions](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/organismal-interaction) with their environment also play an important role in determining the shape and help in the deposit of wall deposits like stalactite, stalagmite, etc. [[5](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0025)]. Exploration of caves is been gaining interest from the last two decades and several unique microbial strains having [secondary metabolites](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/secondary-metabolite) production potential have been reported [[19](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0095),[56](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0280),[70](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0350)]. Caves are located worldwide, and the available literature suggests the existence of novel genera in the caves located in Mexico [[59](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0295)], Northern Thailand [[48](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0240),[49](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0245)] and Spain [[32](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0160),[33](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0165)].

Previous studies on cave [microbiome](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/microbiome) have reported their important role in understanding cave ecology, mineral formation and ecosystem [bioenergetics](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bioenergy) [[14](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0070),[15](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0075)], still cave microbiology from Mizoram, India and its importance is still scantily reported and has given very less attention [[20](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0100)]. Among the several caves located in Mizoram, Northeast India, Pukzing cave is recorded as the largest cave of the state with stretches of 25 m which is located in Pukzing village near Marpara in the western hills of Aizawl district (<http://www.incredible-northeastindia.com/mizoram/caves.html>). So far, no systematic study is been carried out to understand the microbial population associated with one of the largest caves of Mizoram, Northeast India. This is the 1st report from unique cave of Pukzing, Mezoram with attempt to explore [microbial genomic](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/microbial-genomics) resources having AMP activities from unique ecological niche. The antimicrobial genes are not only screened in these isolates but in vitro AMP assay against six standard microbial pathogen have been done to prove its efficacy of microbicidal activities unlike [[20](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0100)]. This study was carried out for the first time to determine the microbial population of this cave using culture-dependent and culture-independent methods. The present study was planned to have a clear-cut estimation of the bacterial population using the [metagenomics](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/metagenomics) approach as the culture-dependent techniques leads to a high underestimation of microbial biodiversity due to the inability to grow most of the organisms on nutritional media [[68](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0340)]. As we know that the advancement in sequencing technologies and the use of high-throughput sequencing technologies has made the studies of [microbial diversity](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/microbial-diversity) easier and more informative. However, using metagenomics approaches to study the microbial diversity of any location has chances to lose some low abundant taxa as stated earlier [[65](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0325)]. Hence, the present study targeted both culture-based and culture-independent methods to gain the most appropriate representation of the microbial population associated with the studies cave. The bacterial isolates obtained through culture-dependent methods in this study were screened for their [antibacterial activities](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/antibacterial-activity) and the biosynthetic modular genes (PKS I, PKSII and NRPS) were detected in the potential isolates. This study indicated that the use of V4 region in culture-independent method is more appropriate as compared to V3 as the V4 region has given more diverse group of bacterial population and also subsystem annotation of V4 region predicted two functions associated with [amino acid](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/amino-acids) and its derivatives and cofactors, vitamins, prosthetic groups, pigments functions which was not shown in case of V3 region.

**2. Materials and methods**

**2.1. Samples collection and physiochemical analysis**

The fresh cave sediment sample (approximately 150 g) were collected randomly from four different places within the Pukzing cave, situated at Pukzing village near Marpara in the western hills of Mamit District of Mizoram (23°21′44.2”N 92°25′53.6″E). The Pukzing cave which is the largest cave of the state of Mizoram is about 25 m wide and is located at around 2100 m above sea level [[60](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0300),[62](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0310)]. The collected samples were then mixed together to make a composite sample, placed in sterile Himedia Polythene Bags, brought into the laboratory and stored at 4 °C in the refrigerator until use for the culturable study. While for the non-culturable metagenomic [DNA extraction](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/dna-extraction) work the sediment, samples were stored at −80 °C. The pH and temperature of the cave sediment sample was measured using pH meter and thermometers respectively. Further, total carbon, total hydrogen and total nitrogen were determined using CHN analyzer at SAIC, Tezpur University.

**2.2. Culturable bacterial diversity estimation**

2.2.1. Isolation of bacteria from a cave sediment sample

One gram of sediment soil sample was taken and mixed with 10 ml of sterile distilled water. The mixture solution was then separated into normal or pretreated using either cold treatment (15 °C), and hot treatment (55 °C). After treatment, the sample was serially diluted up to 10−1 to 10−5 dilution and were kept for 60 min at respective treatment. 100 μl of each dilution was taken from each treatment and spread on six nutritional media plates. The plates were incubated at 28 °C and 37 °C for 2-3 weeks to observe the colonies of bacteria. The obtained cultures were re-streaked in their respective media and purified cultures were maintained at 4 °C.

2.2.2. Media composition

The nutritional media was used as follows: **1.** Luria Bertani Agar media [LB media] (10 g of [Peptone](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/peptone); 10 g of sodium chloride; 5 g of yeast [extract](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/extract) and 20 g of agar); **2.** Tryptic Soya Agar [TSA media] Media (17 g of [tryptone](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/tryptone); 3 g of soya peptone; 5 g of sodium chloride; 2.5 g of dipotassium hydrogen phosphate; 2.5 g of dextrose and 20 g of agar); **3.** Starch Casein Agar [SCA media] (10 g of starch; 1 g of casein powder; 37 g of seawater and 20 g of agar); **4.** [Tyrosine](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/tyrosine) Agar Media [ISP7 media]; **5.** Tap Water Yeast Extract Agar Media [TWYE media] (5 g of yeast extract; 2 g of dipotassium hydrogen phosphate and 20 g of agar) and **6.** Glycerol [Asparagine](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/asparagine) Agar media [ISP5 media] (1.0 g of L-Asparagine; 1.0 g of [Dipotassium Phosphate](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/dipotassium-phosphate); 0.001 g of Ferrous sulphate heptahydrate; 0.001 g of Manganese chloride tetrahydrate; 0.001 g of [Zinc sulphate](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/zinc-sulfate) heptahydrate and 20.0 g of Agar).

2.2.3. Genomic DNA extraction and PCR amplification using 16S rRNA gene

Total genomic [DNA](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/dna) of the culturable [bacteria isolates](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bacterium-isolate) was extracted by using the Genomic [DNA purification](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/dna-purification) kit (Invitrogen Life technologies, USA). The purity of the obtain DNA (μg/ml) was verified using μ-Drop™ Plate (Thermo Scientific™ Multiskan™ GO Spectrophotometer, USA). PCR amplification of 16S [rRNA gene](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/rna-gene) was performed by using universal primers PA: 5′-AGA GTT [TGA](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/thermogravimetric-analysis) TCC TGG CTC AG-3′) and PH: 5′-AAG GAG GTG ATC CAG CCG CA-3′ [[58](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0290)]. The PCR reaction mixture preparations and its process were carried out as denoted in Passari et al. [[52](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0260)]. The amplified PCR product was run on 1.5% of [agarose](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/agarose) gel and visualized under gel documentation system XR+Bio-Rad (California, United States). The amplified product was purified using Pure-link PCR [Purification](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/copurification) Kit (Invitrogen, USA) and was sequenced commercially at Sci-Genome Labs Pvt. Ltd., India.

2.2.4. Phylogenetic analysis

The obtained sequences were trimmed using Finch TV 1.4.1 version and then compared with the NCBI database using the BlastN search program [[69](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0345)]. After that, the sequences were aligned using the Clustal W software packaged in MEGA 5.05 [[38](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0190),[66](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0330)]. The aligned sequences were used to select a phylogenetic model based on using BIC scores (Bayesian Information Criterion) and AICc value (Akaike Information Criterion, corrected) [[50](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0250)]. A maximum-likelihood [tree](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/tree) was constructed using MEGA 6.0 with Jukes-Cunter model for [actinobacteria](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/actinobacteria); Kimura 3-parameter model for Gram-positive bacteria and Tamura 3-parameter model for Gram-negative bacteria [[37](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0185)]. The robustness of the [phylogenetic tree](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/phylogenetic-tree) was tested by bootstrap analysis (1000 replicates) using *p*-distance model [[25](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0125)].

2.2.5. Antimicrobial peptide assay

Antimicrobial assays are important tools to test and screen the inhibitory effects of myriad compounds against microorganisms before establishing their inhibitory spectra. It shows the potential [microbial genomics](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/microbial-genomics) resources which can be used for discovery of new molecules to be used as AMP (antimicrobial peptide). Antimicrobial screening was performed against [*Staphylococcus aureus*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/staphylococcus-aureus) (MTCC-96), [*Pseudomonas aeruginosa*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/pseudomonas-aeruginosa) (MTCC-2453), *Escherichia coli* (MTCC-739), [*Micrococcus luteus*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/micrococcus-luteus) (MTCC-5262), [*Bacillus subtilis*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bacillus-subtilis) (MTCC-2097) and [*Candida albicans*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/candida-albicans) (MTCC-3017) obtained from Microbial Type Culture Collection (MTCC), Chandigarh, India and maintained in Molecular Microbiology and Systematic Laboratory, Department of Biotechnology, Mizoram University. These six microbial species were selected for screening as they are well known for different disease related to agriculture, domestic animals and human health. The bacterial sample was inoculated in Tryptone yeast extract broth medium (ISP medium1: 5 g of Casein Enzymic hydrolysate; 3 g of Yeast Extract) and incubated at 28 °C, 150 rpm for 7-10 days. Cells were harvested by [centrifugation](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/centrifugation) at 8000 rpm and the supernatant was collected into a fresh tube for testing the [antimicrobial activity](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/antimicrobial-activity) by agar well diffusion method [[61](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0305)]. The test pathogenic microbes (10 − 3 CFUmL-1) were inoculated on a modified nutrient agar plate (5 g of glucose; 5 g of peptone; 3 g of beef extract; 5 g of sodium chloride and 20 g of agar in one liter of sterile distilled water) and wells of 6 mm diameter were prepared by using sterile cork borer. In each of the plates, wells were filled with 50 μl of clear supernatant of various bacterial sample isolates (test) and the plates were incubated at 37 ± 2 °C for 24 h. The antimicrobial data were analyzed in replicates (mean ± standard deviation of mean replicates) using Microsoft Excel XP 2007, while one-way ANOVA was used to determine the difference between antimicrobial activities among the bacterial isolates by using SPSS software version 20.0.

The potential isolates based on antimicrobial activity were selected to detect antimicrobial biosynthetic genes. [Polyketide synthase](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/polyketide-synthase) gene (PKS I) and Non ribosomal peptide synthetase (NRPS) [adenylation](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/adenylylation) domain were amplified using degenerate primers: (K1F:5′-TSAAGTCSAACATCGGBCA-3′and M6R:5′-CGCAGGTTSCSGTACCAGTA-3′ and A3F: 5′-GCSTACSYSAT STACACSTCSGG-3′ and A7R: 5′-SASGTCVCCSGTSGCGTAS- 3′) [[3](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0015)]. Polyketide synthase gene (PKS I) was amplified with a set of degenerate primers (KS∞:5′-TSGCSTGCTTGGAYGCSATC-3′ and KSβ: 5′-TGGAANCCGCCGAABCCTCT-3′) [[46](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0230)]. The reaction was carried out in the Veriti thermal cycler (Applied Biosystems, Singapore) in a final volume of 50 μl containing 50 ng of genomic DNA, 2.0 U of Taq DNA polymerase, 1 mM MgCl2, 0.5 mM of dNTPs, 2.0 μM of each primer and 10% DMSO. All [molecular biology](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/molecular-biology) chemicals and consumables are obtained from In-vitrogen Life technologies, USA. PCR conditions consisted of one [denaturation](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/denaturation) step at 96 °C for 5 min., followed by 35 cycles of denaturation at 96 °C for 60 s, annealing at 59 °C for 60 s, and extension at 72 °C for 2 min. The final extension step was done at 72 °C for 10 min. A negative control reaction mixture without [DNA template](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/dna-template) of actinomycetes was also included with each set of PCR reactions and the PCR products were visualized.

**2.3. Non-culturable bacterial diversity**

2.3.1. Bacterial community profiling using Illumina paired-end sequencing with V3 and V4 variable regions

The total DNA of the composite sample mix collected from Pukzing Cave was obtained by Fast DNA spin kit (QIAGEN, USA) as per the protocol. Out of 1600 bp variable regions (V1-V9), regions V3/V4 were selected as they contain highest heterogeneity offering maximum discriminatory power for metagenomic diversity analysis [[11](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0055)]. Paired-end [Illumina sequencing](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/illumina-dye-sequencing) (250p x 2) of the variable regions V3 and V4 was carried out at SciGenome Pvt. Ltd., Cochin, Kerela, India. Initially pre-processing of pair-end reads in each sample was done with the Fastq-join method [[2](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0010)] to filter out any unpaired reads with uncertain bases. Demultiplexing with [quality Phred score](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/phred-quality-score) of ≥Q20 is performed to remove bases with poor quality sequences. The taxonomic analysis was carried using Quantitative Insights into Microbial Ecology (QIIME) analysis pipeline [[12](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0060)]. UCLUST algorithm [[23](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0115)] was implemented for mapping, processed reads into [operational taxonomic units](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/operational-taxonomic-unit) (OTUs) with a similarity threshold of 97%. Aligning of representative sequences with open reference picking method using the Greengenes reference database [[74](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bib366)] and taxonomic classification with RDP classifier [[18](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0090)] was performed. Based on identified OTUs, phylogenetic tree was obtained using FastTree method [[57](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0285)]. Before farther in-depth study on OTUs, low-abundance OTUs that are OTUs with minimum sample count were removed. In this study, we discarded the OTUs having sample count ≤10.

2.3.2. Taxonomic, statistical and functional analysis

The relative abundance of taxa at each taxon level is calculated using the in-build Perl script. Heat-map at the phylum level was generated by calculating Spearman Correlation [[4](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0020)]. Weighted and unweighted UniFrac distances [[43](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0215)] were calculated for filtered OTUs in all samples. [MedCalc](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/sequest) software was used to perform one-way ANOVA with a post-hoc test ([https://www.medcalc.org](https://www.medcalc.org/), [[45](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0225)]). To study the compositional similarity between samples, the robust Bray-Curtis similarity score was calculated using diversity indices in QIIME [[12](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0060),[71](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0355)] based on comparison of pairwise taxonomic abundances from each sample against other samples followed by Non-metric Multi-Dimensional Scaling using PASTv3.11 software [[27](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0135)]. The phylogenetic tree predicted from FastTree was processed and visualized as phylogenetic cladogram with MEGAN software [[30](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0150),[57](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0285)]. Functional annotation studies including ontology were performed with MG-RAST server [[47](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0235)].

**3. Results**

**3.1. Physiochemical analysis of the cave sediments**

The pH values of the Pukzing cave sediment samples were found to be 7.5 whereas the temperature of the cave was recorded as 27 °C. Moreover, total organic matter and phosphorus content of cave sediment samples were determined 1.5 ± 0.05% and 67.8 ± 0.2 ppm, respectively. Further, total carbon (0.93%) and total hydrogen (0.50%) of cave samples were calculated using CHN analyzer.

**3.2. Isolation of bacteria from the Pukzing cave sediment sample**

In this study, totally 235 [bacterial strains](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bacterial-strain) were selectively isolated based on macroscopic [morphological characteristics](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/morphological-trait). The maximum number of bacterial isolates was obtained from 10−2 dilution (*n* = 81; 46.3%) followed by 10−4 dilution (*n* = 61; 26.8%), 10−3 dilution (*n* = 54; 12.1%) and 10−5 dilution (*n* = 39; 9.7%). The highest bacterial population was recovered from heat treatment (*n* = 97; 41.2%) followed by untreated (*n* = 79; 33.6%) and cold treatment (*n* = 59; 25.1%). Moreover, the maximum number of bacterial isolates was isolated from SCA media (*n* = 92; 39.1%) followed by ISP7 (*n* = 57; 24.2%), ISP5 (*n* = 26; 11.0%), LB (*n* = 24; 10.2%), TSA (*n* = 21; 8.9%) and TH2O media (*n* = 17; 7.2%). Morphologically, most of the bacterial isolates spread out over the plate appeared as flat colonies, with some exhibiting [spores](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bacterial-spore) as rough-surfaced, sticky, smooth, and with pigment production. On incubation, the [bacterial colonies](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bacterium-colony) were observed as white, yellow, orange and pale-yellow colors. Gram test revealed that 142 of the bacterial isolates were Gram-positive (*n* = 142; 60.4%) while 93 of them were Gram-negative (*n* = 93; 39.5).

**3.3. Antimicrobial activity screening of culturable bacterial isolates**

All the 235 isolates were tested for their antimicrobial activities against six microbial pathogens *P. aeruginosa*, *S. aureus*, *E. coli*, *M. luteus*, *B. subtilis* and *C. albicans*. Out of 235 isolates, 133 (56.59%) isolates (58 bacterial and 78 actinobacterial isolates) showed antimicrobial potential against at least five of the tested pathogens. Interestingly, almost all the bacterial isolates exhibited antimicrobial activity against *B. subtilis* (235 isolates) and *S. aureus* (227 isolates) as indicated in [Supplementary Table S1](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "ec0005).

Interestingly, out of these 133 isolates, 48 isolates (29 actinobacterial isolates [60.41%] and 19 bacterial isolates [39.58%]) exhibited broad-spectrum antimicrobial activities against all pathogens. Among these, BPSCV70 (*M. luteus*), BPSCV82 (*Streptomyces* sp.), BPSCV83 (*Streptomyces* sp.), BPSCV84 ([*Micromonospora*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/micromonospora) sp.), BPSCV89 (*Actinobacteria bacterium*), BPSCV102 (*Streptomyces* sp.) and BPSCV120 ([*Actinomycetales*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/actinomycetales)*bacterium*) showed significant [zone of inhibition](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/zone-of-inhibition) accounting for 15 mm against *S. aureus* and *B. subtilis* ([Fig. 1](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "f0005)). Similarly, isolates BPSCV15 (*Virgibacillus* sp.), BPSCV16 ([*Staphylococcus saprophyticus*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/staphylococcus-saprophyticus)), BPSCV46 ([*Kocuria*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/kocuria)*palustris*) and BPSCV99 (*Streptomyces* sp.) exhibited maximum antimicrobial activity against *M. luteus* (10 mm) whereas strain BPSCV102 (*Streptomyces* sp.) showed the highest antimicrobial activity against *C. albicans* (12.5 mm). Given the broad-spectrum antimicrobial nature of 133 bacterial isolates they were all selected as potential candidates for further investigation.

1. [Download : Download high-res image (376KB)](https://ars.els-cdn.com/content/image/1-s2.0-S0888754321003670-gr1_lrg.jpg)
2. [Download : Download full-size image](https://ars.els-cdn.com/content/image/1-s2.0-S0888754321003670-gr1.jpg)

Fig. 1. A) Isolation of bacteria using serial dilution method; B) Mixed [colony bacterial](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bacterium-colony) isolates; C) [Morphological characteristics](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/morphological-trait) of the isolated pure Bacterial isolates; D) [Antibacterial activity](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/antibacterial-activity) of potential isolates where all the isolates showed [zone of inhibition](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/zone-of-inhibition) against pathogens and disk indicates as control. CA: [*Candida albicans*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/candida-albicans); SA: [*Staphylococcus aureus*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/staphylococcus-aureus); BS: [*Bacillus subtilis*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bacillus-subtilis); EC: *Escherichia coli* and PS: [*Pseudomonas aeruginosa*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/pseudomonas-aeruginosa)*.*

**3.4. Biosynthetic gene detection vs antimicrobial activity of the selected isolates**

From the 133 potent isolates selected, the presence of genes encoding PKS type II were detected in 46 isolates (34.5%), while a band of the expected size for PKS type I candidate [amplicons](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/amplicon) was detected in only 37 such isolates (27. 8%) using their respective degenerate primers. Moreover, 54 isolates (40.6%) also detected NRPS gene with the expected band size of 600 bp. In 21 isolates (15.78%) positive amplification products were obtained for all three biosynthetic genes (PKS type I, PKS type II and NRPS) and they were BPSCV3, BPSCV10, BPSCV11, BPSCV12, BPSCV16, BPSCV18, BPSCV21, BPSCV22, BPSCV24, BPSCV35, BPSCV43, BPSCV64, BPSCV66, BPSCV70, BPSCV78, BPSCV80, BPSCV82, BPSCV84, BPSCV89, BPSCV99 and BPSCV102.

**3.5. Identification and phylogenetic analysis of the selected bacterial isolates**

All the potential isolates were identified by 16S-rRNA gene sequencing and amplicon product size was found ~1500 bp. All the isolates showed 98-100% sequence similarity to the reference isolates based on BLAST analysis and the sequences of the isolates were deposited in NCBI GenBank to obtain accession number (MK855189-MK855321). The results showed that the isolates were classified into 13 families belonging to [*Streptomycetaceae*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/streptomycetaceae) (*n* = 66; 49.6%), [*Bacillaceae*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bacillaceae) (*n* = 19; 14.2%), *Micromonosporaceae* (*n* = 8; 6.01%), *Microbacteriaceae* (n = 6; 4.51%), [*Micrococcaceae*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/micrococcaceae) (n = 6; 4.51%), [*Staphylococcaceae*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/staphylococcaceae) (n = 6; 4.51%), *Actinomycetaceae* (*n* = 5; 3.75%), [*Xanthomonadaceae*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/xanthomonadaceae) (n = 5; 3.75%), *Pseudonocardiaceae* (*n* = 3; 2.25%), *Nocardiopsaceae* (n = 3; 2.25%), [*Pseudomonadaceae*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/pseudomonadaceae) (n = 3; 2.25%), *Brevibacteriaceae* (*n* = 2; 1.5%) and *Alcaligenaceae* (n = 1; 0.75%), respectively.

The phylogenetic tree was constructed based on 16S-rRNA gene sequences values using MEGA 6.0. In case of actinobacteria, phylogenetic tree was constructed using the maximum likelihood method with Jukes-Cantor model with lowest BIC (2190.170) and highest AIC (844.878). The gaps were treated by pairwise deletion and the estimated Transition/Transversion bias (R) was 0.50. The tree showed that the entire genus *Streptomyces* group was clustered together in one clade with the bootstrap supported value of 60%. Whereas, the other rare genera like *Micromonospora* were closely clustered with [*Pseudonocardia*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/pseudonocardia) and [*Saccharopolyspora*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/saccharopolyspora) group with bootstrap supported value 55%. Interestingly, genera *Pseudonocardia* and *Saccharopolyspora* were represented under the same family. Further, very rare genera like [*Microbacterium*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/microbacterium)*, Kocuria,*[*Micrococcus*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/micrococcus)*,*[*Nocardiopsis*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/nocardiopsis)*,* and [*Brevibacterium*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/brevibacterium) were clustered together in another clade with the bootstrap supported value of 63% ([Fig. 2](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "f0010)A).

1. [Download : Download high-res image (464KB)](https://ars.els-cdn.com/content/image/1-s2.0-S0888754321003670-gr2_lrg.jpg)
2. [Download : Download full-size image](https://ars.els-cdn.com/content/image/1-s2.0-S0888754321003670-gr2.jpg)

Fig. 2. **A)** Maximum likelihood [phylogenetic tree](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/phylogenetic-tree) with Jukes-Cantor model based on 16S rRNA gene sequences of [actinobacteria](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/actinobacteria) showing the relationship with closest type strain sequences. **B)** Maximum likelihood phylogenetic tree with Kimura 2-parameter model based on 16S rRNA gene sequences of Gram-positive bacteria showing the relationship with closest type strain sequences. **C)** Maximum likelihood phylogenetic tree with Tamura 3-parameter model based on 16S rRNA gene sequences of Gram-negative bacteria showing the relationship with the closest type strain sequences. The number at branches indicate bootstraps value (>50%) from 1000 replicates in all the [trees](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/tree).

In case of Gram-positive bacteria, the phylogenetic tree was also made with maximum likelihood method, but following the Kimura 2-parameter model with the lowest BIC (5088.975) and highest AIC (4558.721) values. Here also, gaps were treated by pairwise deletion while the estimated Transition/Transversion bias (R) was found to be 1.27. The phylogenic tree revealed that all [*Bacillus*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bacilli) genera were clustered together in one clade with the bootstrap supported value of 54% whereas all the *Staphylococcus* genera were clustered separately in another clade with the bootstrap supported value of 88% ([Fig. 2](https://www.sciencedirect.com/science/article/pii/S0888754321003670#f0010)B).

For Gram-negative bacteria, the phylogenetic tree was also built by maximum likelihood method but by using the model of Tamura 3-parameter with lowest BIC (3223.764) and highest AIC (2058.828) values. Gaps again were treated by pairwise deletion while here the estimated Transition/Transversion bias (R) was 1.08. The phylogenetic tree indicated that all isolates of genus [*Stenotrophomonas*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/stenotrophomonas) were closely clustered with genera [*Achromobacter*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/achromobacter) in one clade with the bootstrap supported value of 58% whereas all isolates belonging to genera [*Pseudomonas*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/pseudomonas) was clustered with their type strains *Pseudomonas oryzihabitans* type strain NBRC10219 having bootstrap supported value of 51% ([Fig. 2](https://www.sciencedirect.com/science/article/pii/S0888754321003670#f0010)C).

**3.6. NGS based Non-culturable microbial community profiling of the Pukzing Cave samples and its functional analysis**

3.6.1. Deciphering of the microbial population: Cave sample

The high throughput [Illumina sequencing](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/illumina-dye-sequencing) from total DNA of the cave sediment samples yielded 554,834 and 208,946 paired-end reads with an average Phred scores of 35.93 and 37.43 for the 16S-rRNA variable gene regions V3 and V4, respectively. Based on relative abundance obtained at a phyla level; a total of 38 phyla (BHI80-139; [Chlorobi](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/green-sulfur-bacteria); Elusimicrobia; [Fibrobacteres](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/fibrobacter); Gemmatimonadetes; GN02; MVP-21; OD1; FCPU426; ZB3; Synergistetes; [Euryarchaeota](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/euryarchaeota); NKB19; TM6; GN04; [Thermi](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/thermus); GAL15; WPS-2; Armatimonadetes; [Chlamydiae](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/chlamydiae); AD3; TM7; [Crenarchaeota](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/crenarchaeota); [Fusobacteria](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/fusobacterium); WS2; WS3; [Nitrospirae](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/nitrospirae); [Chloroflexi](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/chloroflexi); Synergistetes; [Acidobacteria](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/acidobacterium); [Cyanobacteria](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/cyanobacteria); [Planctomycetes](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/planctomycetes); Proteobacteria; [Firmicutes](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/firmicutes); [Verrucomicrobia](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/verrucomicrobium); [Bacteroidetes](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bacteroidetes); Actinobacteria & unidentified isolates) was recorded for V3 and V4 region of cave sediment sample ([Fig. 3](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "f0015)A). The phylum Archae group under Euryarchaeota and Crenarchaeota was only found in the V4 region of cave sediment sample ([Fig. 3](https://www.sciencedirect.com/science/article/pii/S0888754321003670#f0015)B).

1. [Download : Download high-res image (422KB)](https://ars.els-cdn.com/content/image/1-s2.0-S0888754321003670-gr3_lrg.jpg)
2. [Download : Download full-size image](https://ars.els-cdn.com/content/image/1-s2.0-S0888754321003670-gr3.jpg)

Fig. 3. Microbial Community Composition including both Bacteria and [Archaea](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/archaeon) at Phylum level (A), only Archaebacteria (B) and only bacteria (C).

The bacterial community analysis exhibited that the phylum Actinobacteria were the most dominant bacteria (93.4%) followed by Proteobacteria (1.82%), Firmicutes (1.34%), Chloroflexi (0.98%) and other (0.89%), in V3 region. Similarly, the most dominant phylum Actinobacteria was recorded (41.4%) followed by Bacteroidetes (11.6%), Verrucomicrobia (10.3%), Firmicutes (9.68%) and other bacteria (21.4%) in V4 resign ([Fig. 3](https://www.sciencedirect.com/science/article/pii/S0888754321003670#f0015)C). In the Pukzing cave, more than 76% of bacterial species were associated with three [Amycolatopsis](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/amycolatopsis), [Streptococcus](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/streptococcus), Solirubrobacteraceae, major phyla including Actinobacteria, Firmicutes, and Proteobacteria ([Supplementary Table S2](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "ec0010)). The detailed identified class under these phylum is given in Supplementary File S3.

3.6.2. Statistical tests, taxonomic differences, and analysis of microbial population: V3 and V4 region

Statistical analyses on taxonomic abundance were performed using a hypergeometric test with an extended error bar plot ([Supplementary Fig. S4](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "ec0020)). It signifies statistical significance (*p*-value >0.05) in species which includes *Saccharopolyspora hirsute* (p-value 8.26e-7), [*Akkermansia muciniphila*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/akkermansia-muciniphila) (1.56e-3), *Nocardioidaceae* (4.57e-3), [*Prevotella*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/prevotella)*copri* (7.14e-3), *Saccharopolyspora* (9.77e-3) and Unassigned taxa (3.59e-6). Other species are showing the difference in proportions including V3 and V4 samples but are not statistical significance.

The similarity in bacterial community structure was inferred from taxonomic data at species level using Bray-Curtis (BC) similarity score and therefore reduced using Non-metric Multidimensional Scaling at 2D space. During processing, taxonomic clades present in at least one sample with a relative OTU count of 100 and above are considered for similarity matrix calculation. Shorter linear distance denotes greater similarity between samples, but in [Fig. 4](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "f0020) it is clearly visible that V3 and V4 samples linear distance is larger and distinct.

1. [Download : Download high-res image (100KB)](https://ars.els-cdn.com/content/image/1-s2.0-S0888754321003670-gr4_lrg.jpg)
2. [Download : Download full-size image](https://ars.els-cdn.com/content/image/1-s2.0-S0888754321003670-gr4.jpg)

Fig. 4. Cave microbial population showed a Nonmetric multidimensional (NMDS) scaling plot of taxonomic similarity (Bray–Curtis): purple ellipse (V3) and green ellipse (V4). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The phylogenetic tree generated using QIIME is compared for V3 and V4 samples and visualized with the MEGAN tool as cladogram ([Supplementary Fig. S5](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "ec0025)) [[30](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0150),[57](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0285)]. MEGAN analysis of Illumina reads showed that Proteobacteria group (alpha, beta, delta and gamma bacteria) were classified together in one group. Similarly, to results from the MEGAN analysis, combination reads of V3 and V4 regions indicated that the PVC group (Planctomycetes and Verrucomicrobia), FCB group (Chlorobi and Bacteriodetes) and Acidobacteria (Acidobacteriaceae and Solibacterales) was also clustered separately. We have found that the larger group of the Terrabacteria (Actinobacteria, [Deinococcus](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/deinococcus), Cyanobacteria, Chloroflexi, Firmicutes, and Armatimonadetes) was clustered together according to the MEGAN analysis of the Illumina reads ([Supplementary Fig. S5](https://www.sciencedirect.com/science/article/pii/S0888754321003670#ec0025)). There is also a group called unassigned which was found as unclassified bacteria in MEGAN analysis.

Rarefaction curve analysis and diversity index are inferred and plotted in [Fig. 5](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "f0025). The alpha diversity analysis demonstrated that the V3 region indicating high diversity (35.0) as compared to V4 region (30.0) ([Fig. 5](https://www.sciencedirect.com/science/article/pii/S0888754321003670#f0025)A**)**. The rarefaction curves showed that a linear relationship between taxa and species richness in the V3 region whereas no linear curve was found between taxa and species richness in the V4 region ([Fig. 5](https://www.sciencedirect.com/science/article/pii/S0888754321003670#f0025)B). Taxonomic and phylogenetic similarity using beta diversity is calculated with Bray-Curtis (B-C) similarity measure and FastUnifrac. Based on the average, we have observed that beta diversity is much higher in both phylogenetic and taxonomic similarity in V3 sample as compared to V4 sample (Supplementary Fig. S6 A–D). The abundance of taxa in a sample computed describes the OTU richness in a sample with more richness in V4 sample as compared to V3 sample (Supplementary Fig. S6 B). At species level microbial communities that are shared and unique to each sample is described in Supplementary Fig. S6 C with 333 species shared in both V3 and V4 sample, 396 being unique to V3 and 129 species unique to V4 sample.

1. [Download : Download high-res image (93KB)](https://ars.els-cdn.com/content/image/1-s2.0-S0888754321003670-gr5_lrg.jpg)
2. [Download : Download full-size image](https://ars.els-cdn.com/content/image/1-s2.0-S0888754321003670-gr5.jpg)

Fig. 5. **A**) Describes the rarefaction curve analysis for the observed OTU's **B)** Alpha diversity was calculated for each community based on taxa distribution at the phylum level using the PAST statistical program.

At the phylum level, heat map analysis is used to understand the taxonomic differences between the two variable regions (V3 & V4). Relative abundance at the phylum level using Spearman correlation (at *P* < 0.05 significance) is generated with a heat map where green signifies strong positive correlation, and pink represents a strong negative correlation ([Fig. 6](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "f0030)). A positive correlation means when one phylum was increased in one variable region (V3), the other variable region increases at the same time with vice-versa in negative correlation. We have observed that a strong positive correlation was much higher in the V4 region as compared to the V3 region.

1. [Download : Download high-res image (631KB)](https://ars.els-cdn.com/content/image/1-s2.0-S0888754321003670-gr6_lrg.jpg)
2. [Download : Download full-size image](https://ars.els-cdn.com/content/image/1-s2.0-S0888754321003670-gr6.jpg)

Fig. 6. Cave microbial population showed a heat map based on Spearman correlation up to species level.

Network interaction analysis on the microbial taxonomic community at the species level is deciphered ([Supplementary Fig. S7](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "ec0035)) using [Cytoscape](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bioinformatics-software) [[17](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0085),[64](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0320)] where red hexagon shape nodes represent the sample, circular nodes in green represent species in samples V3 and V4, Cyan include species in V3 sample and yellow depicts species in V4 sample. Network analysis indicates all the similar species are very closely clustered together in V3 and V4 region that are visualized in green color. The similar species are as follows *Actinomycetales, Pseudonocardia, Rubrobacter, Solirubrobacterales, Streptomyces, Saccharopolyspora, Actinomycetospora, Kibdelosporangium, Pseudonocardiaceae, Amycolatopsis, Streptococcus, Solirubrobacteraceae, JG30-KF-CM45, AKYG1722, iii1-15,*[*Nocardiaceae*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/nocardiaceae)*, Ellin6529,*[*Enterobacteriaceae*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/enterobacteriaceae)*,*[*Clostridiaceae*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/clostridiaceae)*,*[*Ruminococcaceae*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/ruminococcaceae)*,*[*Lachnospiraceae*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/lachnospiraceae)*,*[*Ruminococcus*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/ruminococcus) and *Ellin6075*.

**4. Discussion**

Microorganisms' presence inside the cave systems are commonly oligotrophic or chemolithotrophic in nature and impose precise nutrients for their growth. Hence, it is very difficult to isolate some rare or specific bacteria from these sources [[56](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0280)]. In our study, the maximum number of bacteria was obtained from SCA media (*n* = 92; 39.1%) followed by ISP7 (*n* = 57; 24.2%), ISP5 (*n* = 26; 11.0%), LB (*n* = 24; 10.2%), TSA (*n* = 21; 8.9%) and TH2O media (*n* = 17; 7.2%). This finding was similarly reported by Adam et al. [[1](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0005)] in cave moon milk deposits studies.

In this study, we have found significant antimicrobial activity in cave bacterial isolates against a Gram-positive and Gram-negative bacterial pathogen. In our study, out of 235 isolates, 133 (56.59%) strains showed significant antimicrobial potential at least five of the six tested pathogens whereas 48 (20.4%) isolates exhibited antimicrobial ability against all tested pathogens. Out of 133 isolates, 29 (21.8%) actinobacterial isolates and 19 (14.28%) bacterial isolates showed potential antimicrobial activities against six microbial pathogens. This finding is similar to Tomova et al. [[67](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0335)] in Magura cave of Bulgaria having inhibitory activity against *P. aeruginosa* and the yeast pathogen [*Rhodotorula mucilaginosa*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/rhodotorula-mucilaginosa). Other cave studies were also with similar report of antomicrobial activities like Karstic caves of Turkey ([[36](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0180),[72](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0360)], Yasir [[70](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0350)] -Plz check country), Grotta dei Cervi cave of Italy [[29](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0145)] and Bhullar et al. [[9](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0045)] Lechuguilla cave, New Mexico. Higher activity of AMP assay against various microbes indicates that cave microbial community could be a potential source of future genomic resource discovery especially for wider applications in agricultural, domestic animal and [human](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/human) health ([[42](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0210)]; [[1](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0005),[10](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0050),[22](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0110)]).

As 16S-rRNA gene sequence is the perfect technique to identify the isolates up to species level Clarridge III [[16](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0080)]. In our study, we also found that all the isolates showed 98-100% identity with reference sequences in NCBI GenBank for species identification. All the isolates were divided into 13 families indicates that all isolates are having similar trophic level and surviving in a single microbial community with a strong mutual relationship in ecological niches of the cave. In our study, the phylogenetic tree showed that actinobacteria were the dominant group followed by Proteobacteria and [Firmicutes](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/firmicutes) group. Similarly, Lee et al. [[40](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0200)] reported that the phylogenetic tree of 16S-rRNA gene sequences obtained from 60 caves around the world. Among them, the most abundant groups belong to *Proteobacteria* followed by *Actinobacteria*, and [*Chloroflexi*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/chloroflexi)*.* The similar findings were reported by Yasir [[70](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0350)] who state that the cultured isolates from both caves were divided into the phyla *Proteobacteria*, *Actinobacteria*, and *Firmicutes*. Among them, 13 genera were identified from the culturable study including [*Bacillus*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bacilli), [*Microbacterium*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/microbacterium), [*Pseudomonas*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/pseudomonas) and *Staphylococcus.*

Culture dependent is a commonly useful method to isolate microorganisms from any source, but they do not provide perfect identification of microorganisms. Hence, various molecular techniques have been developed to protect cultural heritage. For example, Pinar et al. [[54](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0270)] has used two PCR based methods for the identification of the cultivable fraction of the halophilic [microflora](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/microflora) that inhabit the Catacombs of Palermo and used molecular approach to develop non-invasive and perfect sampling methods for [DNA extraction](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/dna-extraction) for bacterial and fungal diversity analyses [[53](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0265),[55](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0275)]. The next-generation sequencing technologies were used to analyze bacterial diversity on bricks from barracks in the former Auschwitz IIeBirkenau Museum and to investigate the microbial population colonizing the medieval church of San Leonardo di Siponto (Italy). Interestingly, they have found three dominant bacterial phyla i.e., Proteobacteria Actinobacteria and [Bacteroidetes](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bacteroidetes) [[15](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0075),[26](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0130)]. In our study, total of 38 phyla (BHI80-139; [Chlorobi](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/green-sulfur-bacteria); Elusimicrobia; [Fibrobacteres](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/fibrobacter); Gemmatimonadetes; GN02; MVP-21; OD1; FCPU426; ZB3; Synergistetes; [Euryarchaeota](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/euryarchaeota); NKB19; TM6; GN04; [Thermi](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/thermus); GAL15; WPS-2; Armatimonadetes; [Chlamydiae](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/chlamydiae); AD3; TM7; [Crenarchaeota](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/crenarchaeota); [Fusobacteria](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/fusobacterium); WS2; WS3; [Nitrospirae](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/nitrospirae); Chloroflexi; Synergistetes; [Acidobacteria](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/acidobacterium); [Cyanobacteria](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/cyanobacteria); [Planctomycetes](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/planctomycetes); Proteobacteria; Firmicutes; [Verrucomicrobia](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/verrucomicrobium); Bacteroidetes; Actinobacteria & unassigned other) was recorded in V3 and V4 region of cave sediment sample. Interestingly, we have recorded that the phylum Actinobacteria was the most dominant bacteria in both V3 and V4 regions. Previous researchers reported that the phylum Actinobacteria was found highly dominant in all the cave samples under V4 [hypervariable region](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/hypervariable-region) of 16S-rRNA [[21](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0105),[70](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0350)]. Supplementary File S3 has the details of ones identified under the phylum Actinobacteria. Similarly De-Mandal et al. (2017) reported that cave sediment samples collected from three different caves (Bukpuk [CBP V3], Lamsialpuk [CLP V3] and Reiekpuk [CRP V3]) of Mizoram and found that high abundance of dominant family (>0.01%) under Actinobacteria*.* Moreover, few other genera under Actinobacteria were [*Mycobacterium*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/mycobacterium)*,*[*Corynebacterium*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/corynebacterium)*, Rubrobacter,*[*Actinoplanes*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/actinoplanes)*, Saccharothrix,* and [*Pseudonocardia*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/pseudonocardia). Various studies over the past decade suggested phylum Actinobacteria was highly present in all types of caves because of the favorable condition and environment to sustain inside the caves and are dynamically involved for the formation of crystals in cave walls and [biomineralization](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/biomineralization) process [[7](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0035),[34](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0170)]. Laiz et al. [[39](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0195)] also reported Actinobacteria under the genus *Rubrobacter* was isolated from biodeteriorated monuments that can induce crystal formation in caves and produce biofilm on the limestone. Yasir [[70](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0350)] has isolated number of actinobacteria from two caves i.e. Koat Maqbari Ghaar (KMG) and Smasse-Rawo Ghaar (SG) of North-West region of Pakistan that showed significant antimicrobial activity against various bacterial pathogens.

Interestingly, the genera [*Streptomyces*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/streptomyces) was recorded maximum under the family [*Streptomycetaceae*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/streptomycetaceae) that can synthesize various compounds including alcohols, sugars, [amino acids](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/amino-acids), and aromatic compounds and poses abilities to produce clinically important antibiotics [[44](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0220)]. These findings were similarly reported by various researchers who state that the phylum *Firmicutes* obtained from various regions of caves was highly dominant and identified in more extreme ecosystems that are comparatively more resistant to nutrient stress [[14](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0070),[31](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0155)]. In our study, we have obtained the *Bacillus* group from the Pukzing cave which can form endospores [[31](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0155)]. In addition, the [archaea](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/archaeon) group under mesophilic *Crenarchaeota* and *Euryarchaeota* obtained from Pukzing cave were found in V4 hypervariable region of 16S rRNA. Similarly, both archaea group (mesophilic *Crenarchaeota* and *Euryarchaeota*) was detected in Koat Maqbari Ghaar (KMG) cave of Pakistan and found in the Lechuguilla Cave of United States [[31](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0155)]. Moreover, Legatzki et al. [[41](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0205)] state that the presence of an archaeal community on [calcite](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/calcite) speleothems from Kartchner Caverns, Arizona, USA. Some important genera [*Methylobacterium*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/methylobacterium), *Rhizobium,*[*Kocuria*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/kocuria), [*Acinetobacter*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/acinetobacter), *Renibacterium,* and *Bacillus* were found in our study. Similarly, these genera were reported from another cave and these genera having ability to utilize a carbon substrate as well as play a vital role in [nitrogen fixation](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/nitrogen-assimilation) and calcification [[19](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0095),[31](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0155),[51](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0255),[56](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0280),[63](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0315)]. Numerous numbers of oligotrophic and facultative bacteria like [*Nitrospira*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/nitrospira)*,*[*Sphingomonas*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/sphingomonas)*,*[*Bacillus cereus*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bacillus-cereus)*,*[*Paenibacillus*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/paenibacillus)*, Streptomyces sp.,*[*Brevibacillus*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/brevibacillus)*,* and [*Arthrobacter*](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/arthrobacter) were obtained in our culture-independent study and were previously reported from various oligotrophic environments [[28](https://www.sciencedirect.com/science/article/pii/S0888754321003670" \l "bb0140)]. The phylogenetic analysis of community study indicated that the tree has divided into two major groups i.e. Bacteria and Archaea group. Among them, the bacteria group is the largest group and is divided into three another sub-groups such as FCB group; PVC group and Terrabacteria group. All the [bacterial strains](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/bacterial-strain) were clustered within these groups in our study. All the strains epitomize a novel potential isolate in caves biodiversity that indicates us isolated bacterial population properties to having significant discover a new micro-organism from the cave ecosystem [[34](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0170)]. Moreover, most of the isolated bacteria showed significant antimicrobial potential against bacterial pathogens. Interestingly, few researchers reported that isolated microbiota from cave ecosystem could be a potential source to discover new microorganisms and antimicrobial agents [[48](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0240),[49](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0245),[70](https://www.sciencedirect.com/science/article/pii/S0888754321003670#bb0350)] having wider applicability in health management of crop, animal and human.

**5. Conclusion**

This study reveals that, cave ecosystem is a unique source of endemic and moderate thermophilic microorganisms. Diversity can be explored by both culture-based and culture-independent methods. However, culture dependent methods revealed limited 235 bacterial isolates by temperature treatments viz. heat, cold and normal. Both extremes, cold and heat treatment led to recovery of the highest bacterial population indicating the dominance of moderate thermophiles, Actinobacteria. Our study demonstrates that such microbial communities are having antimicrobial potential with biosynthetic genes like PKS type I, PKS type II and NRPS genes. Culture independent, community-based analysis using the sequencing of both hyper-variable regions V3 and V4 also revealed the dominance of phylum Actinobacteria. In comparative analysis, subsystem annotation using the V4 region was found to be more informative for cave [metagenomes](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/metagenome) functional prediction associated with amino acid and its derivatives like cofactors, vitamins, prosthetic groups and pigments functions. AMP evaluation against six microbial pathogen species revealed that cave microbial communities can be an immensely valuable potential source of microbial genomic resources with higher AMP activity. The microbial panel indicates that these findings are promising for new AMP molecules having highly diverse applications in sectors of agriculture, domestic animals and human health.

v