Research Paper

Profiling Bacterial Diversity of B2 Cave, A Limestone Cave of Baratang, Andaman and Nicobar Islands, India

ARUN KUMAR DE1,2,*, RAMACHANDRAN MUTHIYAN2, JAI SUNDER1, DEBASIS BHATTACHARYA1, A KUNDU1 and S DAM ROY3

1Animal Science Division, ICAR-Central Island Agricultural Research Institute, Port Blair, Andaman and Nicobar Islands, India
2Bioinformatics Centre, ICAR-Central Island Agricultural Research Institute, Port Blair, Andaman and Nicobar Islands, India
3Division of Fisheries Sciences, ICAR-Central Island Agricultural Research Institute, Port Blair, Andaman and Nicobar Islands, India

(Received on 14 March 2018; Revised on 23 August 2018; Accepted on 24 August 2018)

Limestone caves with unique environmental conditions harbor some highly specialized and novel microorganisms. In the present study, bacterial diversity and taxonomic composition of B2 cave, a limestone cave of Baratang, Andaman and Nicobar Islands were investigated by using high throughput illumina sequencing platform. A total of 1,056 operational taxonomic units (OTUs) comprising of 22 bacterial phyla were detected. Proteobacteria dominated the phyla, followed by Actinobacteria, Firmicutes and Nitrospirae. At class level, Gammaproteobacteria was found to be most dominant. Other highly prevalent classes were Alphaproteobacteria, Actinobacteria and Bacilli. Functional analysis of the microbiome revealed high representation of membrane transport, amino acid metabolism, carbohydrate metabolism, replication and repair, energy metabolism, xenobiotics biodegradation and metabolism, lipid metabolism, metabolism of cofactors and vitamins and translation. Interestingly, the cave recorded a high number of unclassified OTUs, which might represent novel species and merit further study to determine their functional significance.

Keywords: Metagenomics; Geochemical Analysis; Limestone Cave; Andaman and Nicobar Islands; Bacterial Diversity; Illumina Sequencing

Introduction

Andaman and Nicobar Islands located at latitude 6°-14° N and longitude 92°-94°E consist of 572 Islands, Islets, rocky outcrops and are stretched to a length of 1912 km. All these islands are biological and geological paradise having indigenous biota of flora, fauna and geological wonders which are yet to be unearthed (De et al., 2014). The endemicity of the flora and fauna is also very high, close to 30% (Kundra et al., 2012). Andaman and Nicobar Islands are amongst the 12 mega biodiversity hotspots of India and are in the confluence of Indo-Myanmarese and Indo-Malaysian biodiversity. However, negligible attention has been paid to the microbial diversity of these pristine islands, especially the extreme habitats like caves.

*Author for Correspondence: E-mail: biotech.cari@gmail.com

Limestone caves are natural cavities in the rock resulting from the seepage of surface waters into limestone structure (Palmer, 1991). Limestone caves have relatively stable temperatures and humidity throughout the year but due to nutrient and energy limitations are considered to be extreme habitats, unfavorable for the development of life (Kumaresan et al., 2014; Tomczyk-Zak and Zielenkiewicz, 2015). Limited access to sunlight hinders photosynthetic activity; thereby making the cave system oligotrophic and the bacteria rely on other energy sources (Ortiz et al., 2014). Some highly specialized microorganisms are expected to be found in the extreme environmental conditions like limestone caves (Schabereiter-Gurtner et al., 2004). In addition, as limestone caves are generally isolated, cave-adapted species tend to show...
a high degree of endemism (Niemiller and Zigler, 2013). Limestone caves have emerged as a model to study biogeography of bacteria. Further, microbes collected from pristine sites like caves that are unexplored and rarely visited by humans are likely to belong to novel taxa or strains with huge bio-prospecting potential. Moreover, the analysis of the microbial diversity of unique ecosystems provides better understanding of the role of microhabitat in shaping of microbial diversity (Ortiz et al., 2013).

For the past few years, attention has been paid to understand the microbial life of extreme environments including caves, revealing unique microbial ecosystems (Jones, 2001; Northup and Lavoie, 2001; Barton et al., 2007). The dominant taxa on cave walls are reported to be Proteobacteria, Acidobacteria, and Actinobacteria (Barton and Jurado, 2007; Pasie et al., 2010; Cuezva et al., 2012). Microbiome analysis from different caves revealed that the bacterial composition varied from cave to cave suggesting the presence of cave specific bacterial lineages (Porca et al., 2012; Wu et al., 2015). Moreover, a number of factors including nutrient inputs and human disturbances influence cave microbial diversity (Wu et al., 2015). In this context, study of the microbial diversity of a limestone cave of Baratang island will be of great importance. However, the heterogeneity of microbial population of the limestone caves of Andaman and Nicobar Islands are poorly understood.

In the present study, the composition and diversity of bacterial community of B2 cave, a limestone cave of Baratang island was analyzed by high throughput illumina sequencing. The cave is small, shallow, and devoid of any human interference. The functional profiling of the microbial communities has also been predicted. Analysis of the community composition of the cave will help to understand the factors influencing microbial diversity and community assemblages.

Materials and Methods

Cave Description and Sample Collection

The B2 cave is situated (Fig. 1A) at Raptus Creek village of the Island of Baratang (Latitude: 12° 05' 726"N, Longitude: 92° 45' 016" E), part of great Andaman archipelago, Bay of Bengal. The cave is located 45 km away from Port Blair, the capital of Andaman and Nicobar Islands, India. Among several limestone caves, this one was selected for the current study due to its negligible human interference. The cave is characterized by high humidity (97.33%) and a mean annual temperature of 22°C. The cave is not open for tourism and has minimal anthropogenic disturbances.

Sampling was done in the month of March, 2017. Cave wall scraps were collected from the cave wall using sterile spatula and placed in sterile polycarbonate tubes (Fig. 1B). The samples were transported to the laboratory in a portable ice box and stored at -20°C until use.

Genomic DNA Isolation and Library Preparation

Genomic DNA was isolated from 5 g of the cave wall scrapings using Nucleospin Soil Kit (Macherey-Nagel, Germany) following the manufacturer’s protocol. 2 µl of isolated DNA sample was resolved on 0.8% Agarose gel at 120 V for 60 min and 1 µl of DNA sample was loaded in NanoDrop for determining A260/A280 ratio. Amplification of the 16S rRNA gene targeting V3-V4 region specific for bacteria was carried out using 10 pmol/µl of each forward (GCCTACGGGNGGCWGCAG) and reverse (ACTACHVGGGTATCTAATCC) primers followed by library preparation using Nextera XT Index Kit (Illumina Inc).

Cluster Generation and Sequencing

High throughput Illumina Mi-seq sequencing was carried out at Eurofins Genomics Ind. Pvt. Ltd, Bangalore, India. Briefly, the amplicons were amplified and amplicon libraries were purified by 1X AMPureXP beads. The sequencing has been performed in duplicate. The libraries were loaded onto MiSeq for cluster generation and sequencing. Paired-end sequencing was done. Raw sequence reads were deposited to the NCBI Sequence Read Archive (SRA) https://www.ncbi.nlm.nih.gov/sra/PRJNA505437 under accession number PRJNA505437.

Pre-processing and Sequence Analysis

Post sequencing, the raw reads were processed using “Quantitative Insights Into Microbial Ecology (QIIME)”, version 1.9.0 (Caporaso et al., 2010). Poor
Bacterial Diversity of a Limestone Cave

quality (quality score < 25) and smaller reads were filtered out using the split_libraries command. Pre-processed sequence reads were clustered into operational taxonomic units (OTU’s) based on their sequence similarity using UCLUST at 97% sequence similarity and each resulting cluster typically represented a species. Since each OTU may be made up of many sequences, a representative sequence for each OTU was picked up for downstream analysis. The representative sequence was used for taxonomic identification of the OTU. This was done by setting assignment method to the RDP classification system (Wang et al., 2007) with a confidence level of 0.8. Greengenes 16S rRNA gene database was used for taxonomic identification of the OTUs. Relative abundance of the bacterial phyla was calculated using QIIME. Alpha diversity plot was generated using QIIME. Genome prediction software PICRUSt (Langille et al., 2013) was used to analyze the functional prediction of the microbiome.

Phylogenetic Tree Construction

Sequences were aligned using QIIME’s default alignment method PyNAST version 1.2.2q (Caporaso et al., 2010). FastTree version 2.1.3q (Price et al., 2009) was used to generate a phylogenetic tree of the aligned OTUs. Trees were visualized with FigTree version 1.4.

Results

Taxonomic Distribution at Phylum Level

After quality filtering, 119,953 high quality reads with length of 57, 218, 959 bases were obtained. A total of 1,056 Operational Taxonomic Units (OTUs) at a 97% cut-off for sequence identity were observed. All the reads were aligned against the Green genes reference database and the phyla-level community taxonomic profiles were derived. A total of 22 phyla (Proteobacteria, Actinobacteria, Firmicutes, Nitrospirae, Planctomycetes, Acidobacteria, Gemmatimonadetes, Chloroflexi, Fusobacteria, Bacteroidetes, Verrucomicrobia, TM7, WS3, Chlorobi, Armatimonadetes, Tenericutes, NC10, SBR1093, GN02, Cyanobacteria, WPS-2 and BRC1) were detected in the sample. The relative abundance of the phyla is presented in Figure 2a. The sample was highly dominated by Proteobacteria (83.43%), while other dominant phyla were the Actinobacteria (8.91%) and Firmicutes (4.94%). A phylogenetic tree highlighting all the bacterial phyla is presented in Fig. 3A. The details of the three most dominant bacterial phyla are given below.

Proteobacteria

Under this phylum, four classes (Gammaproteobacteria, Alphaproteobacteria, Betaproteobacteria and Deltaproteobacteria) were identified (Fig. 2B) among which Gammaproteobacteria was the most dominant (85.51%). Under the class Gammaproteobacteria, abundant orders were Pseudomonadales, Xanthomonadales, Chromatiales and Enterobacteriales (Fig. 2E). Twelve families (Xanthomonadaceae, Pseudomonadaceae, Moraxellaceae, Chromatiaceae, Enterobacteriaceae, Sinobacteraceae, Aeromonadaceae, Piscirickettsiaceae, Pseudoalteromonadaceae, Shewanellaceae, Idiomarinaceae and...
Fig. 2: Taxonomic distribution of microorganisms, (A) Taxonomic distribution at phylum level. The category “others” represents bacterial phyla that were represented in less than 0.1% and include the phyla Verrucomicrobia, TM7, WS3, Chlorobi, Armatimonadetes, Tenericutes, NC10, SBR1093, GN02, Cyanobacteria, WPS-2 and BRC1. (B) Class level distribution of phylum Proteobacteria, (C) Class level distribution of phylum Actinobacteria, (D) Class level distribution of phylum Firmicutes, (E) Order level distribution of class Gammaproteobacteria, (F) Order level distribution of class Alphaproteobacteria, (G) Order level distribution of class Betaproteobacteria and (H) Order level distribution of class Deltaproteobacteria
Succinivibrionaceae) and eight abundant genera (Stenotrophomonas, Pseudomonas, Acinetobacter, Lysobacter, Pseudoxanthomonas, Steroidobacter, Aeromonas and Erwinia) were identified under the class Gammaproteobacteria. The dominant orders under the class Alphaproteobacteria were Rhizobiales, Rhodospirillales, Sphingomonadales and Rhodobacterales (Fig. 2F). The dominant genera under this class were Agrobacterium, Ensifer, Kaistobacter, Rhodoplanes, Hyphomicrobium, Devosia, Pedocribobacterium, Sphingomonas, Sphingobium, Mesorhizobium, Paracoccus, Shinella, Ochrobactrum, Mycoplana, Aminobacter, Rhizobium, Bradyrhizobium, Brevundimonas, Aitiffea and Balneimonas. Under the class Betaproteobacteria, the dominant orders were Burkholderiales and MND1 (Fig. 2G) and dominant genera identified were Achromobacter, Cupriavidus, Comamonas, Xylophilus, Rubrivivax and Variovorax. Deltaproteobacteria was present in low abundance and four order (Myxococcales, Syntrophobacterales, Bdellovibrionales, NB1-j and Desulfobacterales) (Fig. 2H) were identified. A phylogenetic tree based on V3-V4 regions of 16S rRNA gene sequences associated with Preteobacteria phyla is presented in Fig. 3B.

**Actinobacteria**

The identified classes under Actinobacteria were Actinobacteria, Nitriliruptoria, Thermoleophilia, Acidimicrobiia, MB-A2-108, Rubrobacteria and OPB41 (Fig. 2C). The most dominant class under this phylum Actinobacteria was found to be Actinobacteria (85.51) followed by Nitriliruptoria, Thermoleophilia, Acidimicrobiia, MB-A2-108, Rubrobacteria and OPB41. Under Actinobacteria class, dominant families were Pseudonocardiaceae, Streptomycetaceae, Euzeybacaeae, Mycobacteriaceae, Gaiellaceae, Micrococcaceae, Nocardioidaceae, Microbacteriaceae, Promicromonosporaceae and Nocardiaceae and dominant genera were Euzeya, Streptomycyes, Mycobacterium, Arthrobacter, Promicromonospora, Agromyces, Pseudonocardia, Rhodococcus, Nocardioides, Cellulomonas, Microbacterium, Nocardia, Planomonospora, Nonomuraea, Kribbella, Nesterenkonia, Actinomadura, Sporichthya, Janibacter and Amycolatopsis. Phylogenetic tree based on V3-V4 regions of 16S rRNA gene sequences associated with Actinobacteria phyla is presented in Fig. 3C.

**Firmicutes**

Two classes (Clostridia and Bacilli) were identified under this phylum (Fig. 2D). Under the class Bacilli, seven families (Paeniibacillaceae, Bacillaceae, Planococcaceae, Thermoactinomycetaceae, Alcicylobacillaceae, Staphylococcaceae and Exiguobacteraceae) and seven abundant genera (Bacillus, Paenibacillus, Pontibacillus, Brevibacillus, Atmonophilus, Lyisinibacillus and Alcylobacillus) were identified. Clostridia were present in low abundance and five families (Clostridiaceae, Ruminococcaceae, Peptostreptococcaceae, Lachnospiraceae, Sulfobacillaceae) and three abundant genera (Clostridium, Tepidibacter and SMB53) were identified. A phylogenetic tree based on V3-V4 regions of 16S rRNA gene sequences associated with Firmicutes phyla is presented in Fig. 3D.

**Taxonomic Distribution at Class Level**

Top ten abundant bacteria at class level are presented in Fig. 4A. Most abundant class was Gammaproteobacteria followed by Alphaproteobacteria, Actinobacteria, Bacilli, Betaproteobacteria, Nitriliruptoria, Thermoleophilia, Nitrospira, Planctomycetia and Acidimicrobiia.

**Taxonomic Distribution at Order Level**

Top ten abundant bacteria at order level is presented in Figure 4b. Pseudomonadales and Xanthomonadales were the two most abundant bacterial orders found. Rhizobiales, Actinomycetales, Bacillales, Burkholderiales, Chromatiales and Euzeybales were also found at relatively high abundance.

**Taxonomic Distribution at Family Level**

Top ten abundant bacteria at family level are presented in Fig. 4C. The most abundant family was Xanthomonadaceae. Pseudomonadaceae was also found in very high abundance. Other relatively high abundant families were Moraxellaceae, Pseudonocardiaceae, Rhizobiaceae, Paenibacillaceae, Bacillaceae, Alcaligenaceae, Streptomycetaceae and Hyphomicrobiaceae.
Diversity Estimates of the Cave Bacterial Community

The Shannon and Chao1 alpha diversity of the sample was estimated from the distribution of the species level annotations. The Shannon and Chao1 alpha diversity was found to be 5.19 and 2340.67 respectively. The rare fraction curve of the sample is presented in Fig. 5. The vertical axis displays the diversity of community, while the horizontal axis displays the number of sequences considered for the calculation. On the left, the steep slope indicates that a large fraction of species diversity remains to be discovered.

Functional Prediction of the Microorganisms

The predicted functional analysis of the microbiome was carried out PICRUSt. Analysis revealed six functional modules at level 1. Among the five functional modules, metabolism was the most represented (60.73%) followed by Genetic Information Processing (17.00%), Environmental Information Processing (15.77%), Cellular Processes (4.40%), Human Diseases (1.09%) and Organismal Systems (1.01%). At level 2 of functional prediction analysis, 39 KEGG subsystems were present, among which membrane transport (12.24%) was the largest category. High representation of amino acid metabolism (12.04%), carbohydrate metabolism (11.21%), replication and repair (6.91%), energy metabolism (5.98%), xenobiotics degradation and metabolism (5.46%), lipid metabolism (4.80%) and metabolism of cofactors and vitamins (4.41%) were observed (Fig. 6).
Moreover, the functional analysis showed high representation of genes involved in transport and different metabolic pathways. Predicted genes coding for mineral and organic transporters, phosphate and amino acid transporters, oligosaccharide, polyol and lipid transporters, monosaccharide transporters, peptide and nickel transporters, metallic cation, iron-siderophore and vitamin B12 transporters and ABC-2 transporters were detected. In carbon metabolism, genes linked with carbon fixation, methane metabolism, carbohydrate and lipid metabolism, nucleotide and amino acid metabolism and carboxylic acid metabolism were detected. Interestingly, the analysis revealed high representation of genes involved in biosynthesis of secondary metabolites like Monolignol biosynthesis, Flavanone biosynthesis, Flavonoid biosynthesis, Cyanogenic glycoside biosynthesis, Glucosinolate biosynthesis and Paspaline biosynthesis. Predicted genes involved in sulfur metabolism were detected. Genes involved in assimilatory sulfur reduction (PAPSS, Sat, Cys N, Cys D, Cys NC, Cys H, Cys C, Cys JI and Sir) and dissimilatory sulfate reduction and oxidation (Sat, AprAB and DsrAB) were found to be present.

**Discussion**

Caves are considered as potential microbial biodiversity hotspots and as the energy limitation of caves poses a challenging condition for life, it is assumed that the microbes colonizing in cave environment will be highly specialized and novel (Jones et al., 2010). Exploring the diversity of microbes from extreme conditions like caves is an emerging trend with huge potential of finding some novel species of pharmaceutical importance (Tomczyk-Zak and Zielenkiewicz, 2015). Considering the huge biodiversity of Andaman and...
Nicobar islands, studying the unique habitat provides the opportunity to understand global microbial diversity, cave specific bacterial diversity, novel population assemblage and energy dynamics (Ortiz et al., 2013). The present study documents the bacterial community composition and diversity along with the geochemical analysis of a limestone cave of Andaman and Nicobar Islands. At phylum level, Proteobacteria and Actinobacteria were the dominant bacteria owing to their ability to degrade organic matter for production of energy. Interestingly, the cave recorded a high number of unclassified OTUs, which might represent novel species and merit further study to determine their functional significance.

Earlier studies on microbial diversity of extreme habitats were carried out mainly on the cultivation techniques of the microbes. As less than 0.01% of microorganisms were cultivable using standard techniques (Amann et al., 1995; Torsvik and Ovreas, 2002) deciphering the true picture of microbial diversity of extreme habitats was not possible. With the emergence of next generation sequencing technology in recent past, research on the microbiome of extreme environments has been intensified. Despite the increasing number of investigations in the past decade, the microbiology of caves is an important topic for better understanding of the substrate biosphere biodiversity. In the present study a total of 23 bacteria phyla were detected and the three dominant phyla were Proteobacteria, Actinobacteria and Firmicutes (Fig. 2A). At class level, Gammaproteobacteria was found to be most dominant (Fig. 2B). Other highly prevalent classes were Alphaproteobacteria, Actinobacteria and Bacilli (Fig. 4B). The dominant bacterial phylotypes found in the current study are consistent with the species usually associated with nutrient-deficient oligotrophic environments. Proteobacteria has been identified as the most dominant bacterial phyla in several limestone caves like limestone cave of the Western Loess Plateau of China (Wu et al., 2015) and limestone caves of Mizoram, Northeast India (De Mandal et al., 2017). Proteobacteria also has been reported as the largest taxonomic group in extreme environments, such as the sulfurous cave of Grotta Nuova di Rio Garrafo (Jones et al., 2010), Movile cave (Sarbu et al., 2000), microbial mats on lava caves (Northup et al., 2011; Hathawaya et al., 2014), and on the stalactites of the Herrenberg Cave (Rusznyak et al., 2012). The ability of Proteobacteria to degrade a wide range of organic compounds (Tomczyk-ak and Zielenkiewicz, 2015) might be the reason of their successful colonization in extreme conditions. Chemolithoautotrophic microorganisms derive energy from oxidation of reduced metal ions like manganese and iron present in the rocks (Tomczyk-ak & Zielenkiewicz, 2015). It has been found that most iron-oxidizing bacteria fall in the Proteobacteria phylum (Hedrich et al., 2011; Lu et al., 2010). Interestingly, in the present study, the wall scrapings of limestone cave was dominated by Proteobacteria followed by that of Actinobacteria (8.9%). Other authors have also reported the abundance of Actinobacteria in limestone caves (Barton et al., 2007; Miller et al., 2014). Their role in degradation of organic waste and mineral utilization has been well established (Ivanova et al., 2013). They play a vital role in decomposition of organic materials, such as cellulose and chitin; and replenish the supply of nutrients in the soil (Tomczyk-ak & Zielenkiewicz, 2015). Actinobacteria have the ability to degrade all sorts of organic substances such as cellulose, polysaccharides, protein fats and organic acids (Barton et al., 2004). A wide variety of biologically active enzymes are produced by Actinobacteria; they secrete amyloses which helps them to carry out extracellular digestion (Barton et al., 2004). Actinobacteria has the ability to survive under nutrient deficient conditions and they can utilize CaCO3 and substrate phosphate efficiently (Zhuand Dittrich, 2016). Firmicutes and Nitrosporae are also found to be abundant bacterial groups in the current study. These two bacterial groups were also found to be abundant in Frasassi cave (Macalady et al., 2006) and Tilo Bastillo cave (Schabereiter-Gurtner et al., 2004; Pasie et al., 2010). The microbial diversity of the B2 cave was found to be low with shanon index of 5.19 and Chao1 index of 2340.67. The low microbial diversity of the cave wall might be due to the extreme environmental conditions. Ortiz et al. (2013) reported that the microbial diversity shanon index of Kartchen cave surfaces ranged from 4.32 to 6.06. One interesting revelations of the study was that a large proportion of OTUs were unclassified. Further study is required to identify these unclassified OTUs, which may represent novel species with potential of bio prospecting.

PICRUSt software package (Langille et al., 2013) was used to decipher the potential functional roles of bacterial communities identified in the cave.
The predicted functional analysis of metagenomic data revealed the presence of 39 KEGG subsystems at level 2 with Membrane Transport as the largest category. Presence of bacteria with several metabolic potential was identified. The wide range of potential metabolic activities in the cave indicates the presence of complex metabolic pathways, with numerous energy conservation strategies for survival of communities in this nutrient limited cave environment.

Acknowledgement

The study was supported by Department of Biotechnology (DBT), India, under Sub-Distributed Information Centre (Sub-DIC) Project (No: BT/BI/04/066/2004).

References


Miller A Z, Pereira M F C, Calaforra J M, Forti P, Dionísio A and

Niemiller M L and Zigler K S (2013) Patterns of Cave Biodiversity and Endemism in the Appalachians and Interior Plateau of Tennessee-USA PLoS ONE 8 e64177


Palmer A N (1991) Origin and morphology of limestone caves Geol Soc Amer Bull 103 1-21


Torsvik V and Ovreas L (2002) Microbial diversity and function in soil: From genes to ecosystems Curr Opin Microbiol 5 240-245

