Data in Brief

# First report of bacterial community from a Bat Guano using Illumina next-generation sequencing

Surajit De Mandal, Zothansanga, Amritha Kumari Panda, Satpal Singh Bisht, Nachimuthu Senthil Kumar *

Department of Biotechnology, Mizoram University, Aizawl 796004, Mizoram, India

## ARTICLE INFO

## ABSTRACT

V4 hypervariable region of 16S rDNA was analyzed for identifying the bacterial communities present in Bat Guano from the unexplored cave — Pnahkyndeng, Meghalaya, Northeast India. Metagenome comprised of 585,434 raw Illumina sequences with a 59.59% G+C content. A total of 416,490 preprocessed reads were clustered into 1282 OTUs (operational taxonomical units) comprising of 18 bacterial phyla. The taxonomic profile showed that the guano bacterial community is dominated by Chloroflexi, Actinobacteria and Crenarchaeota which account for 70.73% of all sequence reads and 43.83% of all OTUs. Metagenome sequence data are available at NCBI under the accession no. SRP051094. This study is the first to characterize Bat Guano bacterial community using next-generation sequencing approach.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

## Specifications

| | |
|---|---|
| Organism/cell line/tissue | Metagenome of Bat Guano in Pnahkyndeng Cave |
| Sex | Not applicable |
| Sequencer or array type | Illumina |
| Data format | Analyzed |
| Experimental factors | Environmental sample |
| Experimental features | V4 hypervariable region of 16S rDNA was sequenced using paired end Illumina Mi-Seq technology and the sequence was analyzed using QIIME data analysis package. |
| Consent | Not applicable |
| Sample source location | Bat Guano sample, Pnahkyndeng Cave, Meghalaya, India |

## Direct link to deposited data

http://www.ncbi.nlm.nih.gov/sra/?term=SRP051094

## Experimental design, materials and methods

Bat Guano, an excrement of the cave-dwelling bats forms the basis of the ecology inside the cave by acting as a food source for detritivorous microbes. It contains high content of organic carbon, nitrogen, phosphate, and potassium [1,2]. Bacteria present in Bat Guano were reported to be involved in nitrification process and were also known as potential chitinase producer. A clone library based study in Bat Guano samples has revealed the presence of group 1.1a and 1.1b Crenarchaeota, an efficient ammonia oxidizer [3,4]. Analyzing Bat Guano is also important since they often harbor various pathogens which can be thread for speleologists, and tourists [5]. Although the microbial communities in diverse cave ecosystems have been studied, little is known about the microbial communities of Bat Guano heaps [6,7] and there has been no studies using high throughput sequencing technology.

Meghalaya is known to possess the largest and most diverse Karst caves in the world [8]. Pnahkyndeng Cave located in Ri-Bhoi district of Meghalaya, India is a home of various bats and offering an ideal environment for studying the Bat Guano microbiota without any anthropological influence. Samples were collected on February 2014 from the Bat Guano of Pnahkyndeng Cave (25°57′22.70″N, 91°55′43.10″E), Nongpoh, Ri-Bhoi district, India. Ten composite Guano samples were collected from different places of the cave floor and the soil community DNA was extracted separately using the Fast DNA spin kit for soils (MP Biomedical, Solon, OH, USA). The freshly extracted DNA was purified twice using 0.5% low melting point agarose gel and mixed to prepare a composite sample. Final DNA concentrations were quantified by the using a microplate reader (BMG Labtech, Jena, Germany).

The V4 region of the 16S rRNA gene was amplified using F515/R806 primer combination (5′-GTGCCAGCMGCCGCGGTAA-3; 5′-GGAC TACHVGGGTWTCTAAT-3′). Amplicon was extracted from 2% agarose gels and purified using the QIA quick Gel Extraction kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. Purified amplicon was paired-end sequenced on an Illumina Mi-Seq

* Corresponding author. Tel.: +91 9436352574 (mobile).
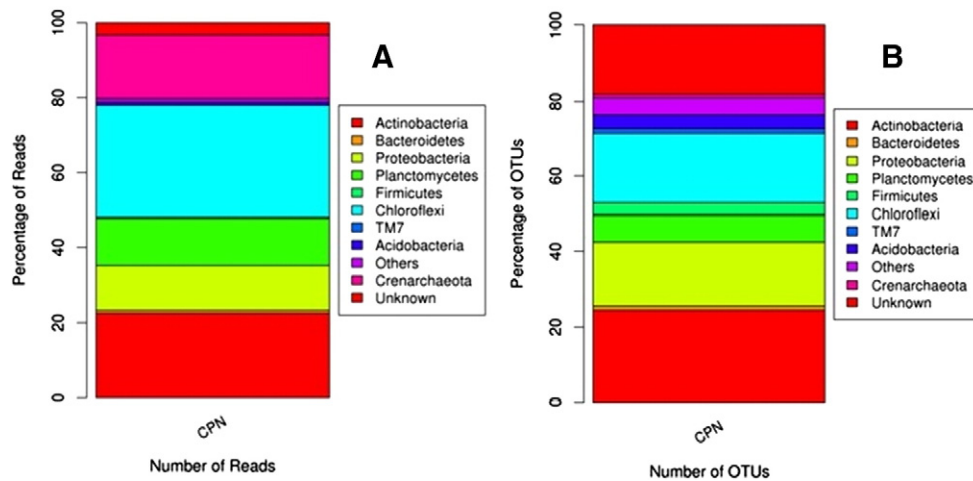E-mail address: nskmzu@gmail.com (N. Senthil Kumar).

**Fig. 1.** Taxonomy classification of reads at phylum level (A), and OTUs at phylum level (B) for the sample. Only top 10 enriched class categories are shown in the figure.

platform. QIIME data analysis package was used for 16S rRNA data analysis [9]. Quality filtering on raw sequences was performed according to base quality score distributions, average base content per read and GC distribution in the reads. Singletons, the unique OTU that did not cluster with other sequences, were removed as it might be a result of sequencing errors and can result in spurious OTUs. Chimeras were also removed using UCHIME and pre-processed consensus V4 sequences were grouped into operational taxonomic units (OTUs) using the clustering program UCLUST at a similarity threshold of 0.97 [10,11]. All the pre-processed reads were used to identify the OTUs using QIIME program for constructing a representative sequence for each OTUs. The representative sequence was finally aligned to the Greengenes core set reference databases using PyNAST program [9, 12]. Representative sequence for each OTU was classified using RDP classifier and Greengenes OTU database. Sequences which are not classified were categorized as unknown.

In our sample, 403,529 reads were classified at the phylum, 282,350 at the order, 188,406 at the family and 2926 sequences were classified at the species levels. Classified OTUs belonged to 18 different phyla dominated by Chloroflexi, Crenarchaeota, Actinobacteria, Bacteroidetes, Proteobacteria, and Planctomycetes (Fig. 1). Analysis of bacterial communities revealed the two most dominant bacteria's — Chloroflexi (29.97%) and Actinobacteria (22.55%), which are known to be a common inhabitant of cave microflora. Other identified phyla include Crenarchaeota (16.96%), Planctomycetes (12.41%) and Proteobacteria (12.03%). Chloroflexi was divided into 11 classes — Thermomicrobia, Planctomycetia, Gitt-GS-136, Ktedonobacteria, Anaerolineae, TK10, TK17, S085, Chloroflexi, Ellin6529, and Gitt-GS-136. The most dominant OTU within this phyla was denovo 317, classified under the class Thermomicrobia (40.52%) followed by denovo 710 under the Thermomicrobia (7.61%), denovo 235 under the genus Thermogemmatisporaceae (7.18%) and denovo 3 under the genus Gemmataceae (2.98%). Actinobacteria was the second most dominant group in our sample containing 314 OTUs. The nutrient rich acidic Bat Guano favors the growth of acidophilic Actinobacteria [13]. The dominant OTU within this phyla was denovo 74 under the genus Mycobacterium (29.39%) followed by denovo 993 (8.27%) and denovo 372 (5.27%) classified under the genus Acidimicrobiales and Actinomycetales, respectively. Only five OTUs were classified up to the species level (*Mycobacterium llatzerense* and *Mycobacterium celatum*). A third dominant phylum in this sample was identified as Planctomycetes comprising of 91 OTUs and 46,063 reads. Identified classes within this phylum were Phycisphaerae, Planctomycetia and BD7-11. Seventeen archeal OTUs were classified under the order NRP-J, Methanomicrobiales and Methanosarcinales. Four of them were identified at the genus

level (Methanosarcina, Haloquadratum, Methanosaeta and Methanocorpusculum). The phylogenetic tree based on the genus level relationships is provided as Supplementary Fig. 1. Previous study on archaeal communities present in Bat Guano identified many ammonia oxidizing bacteria but it was limited with a few number of clones [2].

Our analysis provides in-depth and high throughput identification of the bacterial communities present in Bat Guano. In the present study, we identified 18 bacterial phyla and most of the bacterial genus identified was known to be involved in nitrogen cycling as seen in previous study [2]. A significant portion of OTUs still remains unclassified which indicates the possibility for the presence of novel species in Pnahkyndeng Cave. Further studies like whole metagenome sequencing or functional metagenomics can illustrate the detailed information of this bacterial community.

### Nucleotide sequence accession number

### Competing interests

The authors declare that there are no competing interests.

### Acknowledgments

### References

[1] C. Korine, I. Izahi, Z. Arad, Is the Egyptian fruit bat Rousettus aegyptiacus a pest in Israel? An analysis of the bat's diet and implications for its conservation. Biol. Conserv. 88 (1999) 301–306.
[2] A. Chronakova, A. Horak, D. Elhottova, V. Kristufek, Diverse archaeal community of bat guano pile in Domica Cave (Slovak Karst, Slovakia). Folia Microbiol. 54 (5) (2009) 436–446.
[3] C.A. Francis, K.J. Roberts, J.M. Beman, A.E. Santoro, B.B. Oakley, Ubiquity and diversity of ammonia-oxidizing *Archaea* in water columns and sediments of the ocean. Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 14683–14688.
[4] C. Schleper, G. Jurgens, M. Jonuscheit, Genomic studies of uncultivated Archaea. Nat. Rev. Microbiol. 3 (2005) 479–488.

[5] H. Valdez, R.A. Salata, Bat-associated histoplasmosis in returning travelers: case presentation and description of a cluster. J. Travel Med. 6 (1999) 258–260.

[6] C. Saiz-Jimenez, Microbiological and environmental issues in show caves. World J. Microbiol. Biotechnol. 28 (2012) 2453–2464.

[7] T. Sugita, K. Kikuchi, K. Makimura, K. Urata, T. Someya, K. Kamei, M. Niimi, Y. Uehara, Trichosporon species isolated from guano samples obtained from bat-inhabited caves in Japan. Appl. Environ. Microbiol. 71 (2005) 7626–7629.

[8] S. Baskar, R. Baskar, N. Lee, P.K. Theophilus, Speleothems from Mawsmai and Krem Phyllut caves, Meghalaya, India: some evidences on biogenic activities. Environ. Geol. 57 (2009) 1169–1186.

[9] J.G. Caporaso, J. Kuczynski, J. Stombaugh, K. Bittinger, F.D. Bushman, E.K. Costello, N. Fierer, A.G. Pena, J.K. Goodrich, J.I. Gordon, et al., QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7 (2010) 335–336.

[10] Edgar, et al., UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27 (16) (2011) 2194–2200.

[11] R.C. Edgar, Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26 (2010) 2460–2461.

[12] T.Z. DeSantis, P. Hugenholtz, N. Larsen, M. Rojas, E.L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, G.L. Andersen, Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl. Environ. Microbiol. 72 (2006) 5069–5072.

[13] M. Goodfellow, S.T. Williams, Ecology of actinomycetes. Annu. Rev. Microbiol. 37 (1983) 189–216.