Abstract

The unsafe disposal of hospital effluents contributes to gross contamination of water bodies with antibiotic residues, antibiotic resistance genes and antibiotic resistance bacteria. This study reports the microbial community profile of hospital wastes collected from various regions of West Bengal, India, using 16S rRNA gene amplicon sequencing. The data set Liquid Sludge (LS) contains 15,372,973 reads with an average length of 301 bps with average 52 ± 5% GC content. The data set Solid Sludge (SS) contains 16,071,594 reads with an average length of 301 bps with average 53 ± 4% GC content. Data of this study are available at NCBI BioProject (PRJNA360379). In sample LS, an abundance of 19.3% for the members of Bacteroidetes was observed. In sample SS, an abundance of 19.7% for the members of Euryarchaeota was observed.

© 2019 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/).

1. Data

This data reports the occurrence of microbial abundance in hospital effluents collected from Purulia (LS) and Kolkata (SS), West Bengal, India.
The data set LS contains 15,372,973 reads with an average length of 301 bps as shown in Fig. 1(a). Of the sequences that passed QC, 28,095 sequences (93%) contain ribosomal RNA genes, 1,104 sequences (3.65%) contain predicted proteins with known functions, and 1,069 sequences (3.53%) contain predicted proteins with unknown function.

The data set SS contains 16,071,594 reads with an average length of 301 bps represented by Fig. 1(a). Of the sequences that passed QC, 12,171 sequences (59%) contain ribosomal RNA genes, 1,885 sequences (9.21%) contain predicted proteins with known functions, and 6,403 sequences (31.30%) contain predicted proteins with unknown function.

In sample LS, the community study revealed an abundance of 19.3 % for the members of Bacteroidetes. In sample SS, the community study revealed an abundance of 19.7 % for the members of Euryarchaeota.

As shown in Fig. 1(b), at the genus level, Prevotella was the most dominant microbial member with abundance of 17.47% in LS. For SS, Methanosaeta was the dominant genus with abundance of 16.47%.

The comparative genus diversity LS and SS sample sets is represented by Fig. 1(c) with Prevotella, Bacillus, Clostridium etc. as the common genera.

2. Experimental design, materials and methods

The untreated hospital wastewater and sludge were aseptically collected in sterile containers from rural and urban regions of West Bengal from the main drainage systems of the hospitals aseptically. The effluents were transported within 10 hours of collection. The samples were stored at −20 °C until further processing.

The DNA isolation from each sample was done following the protocol by Bonet et al., 2012 [2]. The quantification of DNA was performed using Qubit dsDNA HS Assay kit (Life Tech). The concentration was determined by Qubit® Fluorometer by taking 1 μl of each sample. The microbial genomic DNA from hospital effluents was normalized to concentration <10 ng/μl. The PCR library preparation of amplicons
was carried out using Nextera XT Idex Kit (Illumina, Inc.). The 16S Metagenomic Sequencing Library preparation protocol was followed. Primers for the amplification of the V3–V4 hyper-variable region [V3 Forward Oligo: CCTACGGGNBGCASCAG and V4 Reverse Oligo: GACTACNVGGGTATCTAATCC] of 16S rDNA gene of bacteria and Archaea were used. The amplification of amplicons with the Illumina adaptors were performed by i5 and i7 primers that add multiplexing index sequences as well as common adapters required for cluster generation (P5 and P7) as per the standard Illumina protocol. The purification of amplicon libraries was done by 1/2 AMpureXP beads and checked on Agilent DNA 1000 chip on Bioanalyzer 2100 and quantified on fluorometer by Qubit dsDNA HS Assay kit (Life Technologies). The library size of Sample LS and Sample SS were 2 million reads each. The sequencing of the libraries was done using the Illumina sequencing chemistry to generate ~150 Mb of data per sample. After obtaining the Qubit concentration for the library and the mean peak size from Bioanalyzer profile, library was loaded onto Illumina Platform at appropriate concentration (10–20 pM) for cluster generation and sequencing. The PCR amplicons were tagged with complementary adapter oligos on paired-end flow cell using the kit reagents. The designing of adapters was done as such to allow selective cleavage of the forward strands after re-synthesis of the reverse strand during sequencing. The copied reverse strand was then used to perform sequencing from the opposite end of the fragments.

2.1. Bioinformatics analysis

The quality control of raw reads was performed by using FASTQC toolkit (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) [3]. The processed paired end reads were clustered into OTU’s (Operational Taxonomic Units) by using QIIME software (qiime.org) in order to identify the microbial community profile [4]. The OTU’s were further processed for taxonomic assignment (Greengenes database), phylogenetic and diversity analysis [5]. A large plethora of microbial communities were identified from this study.

Metagenome sequence data from this study are available at the NCBI Sequence Read Archive (SRA) and BioSamples under accession numbers: SAMN11571463 and SAMN11571474.
Funding

The work is supported by the funds provided to Dr. Mahashweta Mitra Ghosh and Dr. Sayak Ganguli by West Bengal DST project 236 (Sanc./ST/P&T/1G-39/2017).

Acknowledgement

The authors hereby acknowledge the Department of Microbiology, St. Xavier’s College, Kolkata for providing the necessary facilities for conducting this study.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] K. Ghosh, S. Mandal, Antibiotic resistant bacteria in consumable fishes from Digha coast, West Bengal, India, Proc. Zool. Soc. 63 (1) (2010) 13–20. https://doi.org/10.1007/s12595-010-0002-8.

[2] N. Garcias-Bonet, J.M. Arrieta, C.N. de Santana, C.M. Duarte, N. Marbà, Endophytic bacterial community of a Mediterranean marine angiosperm (Posidonia oceanica), Front. Microbiol. 3 (2012) 342. https://doi.org/10.3389/fmicb.2012.00342.

[3] S. Andrews, FastQC: a quality control tool for high throughput sequence data, Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc, 2010.

[4] J.G. Caporaso, J. Kuczynski, J. Stombaugh, K. Bittinger, F.D. Bushman, N. Fierer, A.G. Pêa, J.K. Goodrich, J.I. Gordon, G.A. Huttley, S.T. Kelley, D. Knights, J.E. Koenig, R.E. Ley, C.A. Lozupone, D. McDonald, B.D. Muegge, M. Pirrung, J. Reeder, J.R. Sevinsky, P.J. Turnbaugh, W.A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld, R. Knight, QIIME allows analysis of high-throughput community sequencing data, Nat. Methods 7 (5) (2010) 35–36. https://doi.org/10.1038/nmeth.f.303.

[5] D. McDonald, M.N. Price, J. Goodrich, E.P. Nawrocki, T.Z. DeSantis, A. Probst, et al., An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea, ISME J. 6 (2012) 610–618. https://doi.org/10.1038/ismej.2011.139.