Genomics Data 11 (2017) 39–42

Contents lists available at ScienceDirect

Genomics Data

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

Data in brief

Insights into bacterioplankton community structure from Sundarbans mangrove ecoregion using Sanger and Illumina MiSeq sequencing approaches: A comparative analysis

Anwesha Ghosh, Punyasloke Bhadury *

[a] Integrative Taxonomy and Microbial Ecology Research Group, Department of Biological Sciences, Indian Institute of Science Education and Research Kolkata, Mohanpur, 741246 Nadia, West Bengal, India

ARTICLE INFO

Article history:
Received 7 November 2016
Accepted 18 November 2016
Available online 24 November 2016

Keywords:
Bacterioplankton
Mangrove ecosystem
Sanger sequencing
Illumina MiSeq sequencing

ABSTRACT

Next generation sequencing using platforms such as Illumina MiSeq provides a deeper insight into the structure and function of bacterioplankton communities in coastal ecosystems compared to traditional molecular techniques such as clone library approach which incorporates Sanger sequencing. In this study, structure of bacterioplankton communities was investigated from two stations of Sundarbans mangrove ecoregion using both Sanger and Illumina MiSeq sequencing approaches. The Illumina MiSeq data is available under the BioProject ID PRJNA35180 and Sanger sequencing data under accession numbers KX014101-KX014140 (Stn1) and KX014372-KX014410 (Stn3). Proteobacteria-, Firmicutes- and Bacteroidetes-like sequences retrieved from both approaches appeared to be abundant in the studied ecosystem. The Illumina MiSeq data (2.1 GB) provided a deeper insight into the structure of bacterioplankton communities and revealed the presence of bacterial phyla such as Actinobacteria, Cyanobacteria, Tenericutes, Verrucomicrobia which were not recovered based on Sanger sequencing. A comparative analysis of bacterioplankton communities from both stations highlighted the presence of genera that appear in both stations and genera that occur exclusively in either station. However, both the Sanger sequencing and Illumina MiSeq data were coherent at broader taxonomic levels. *Pseudomonas*, *Devosia*, *Hyphomonas* and *Erythrobacter*-like sequences were the abundant bacterial genera found in the studied ecosystem. Both the sequencing methods showed broad coherence although as expected the Illumina MiSeq data helped identify rarer bacterioplankton groups and also showed the presence of unassigned OTUs indicating possible presence of novel bacterioplankton from the studied mangrove ecosystem.

Specifications	
Organism	Sundarbans bacterioplankton metagenome
Sex	Not applicable
Sequencer or array type	Illumina MiSeq
Data format raw data	Fastq file
Experimental factors	Environmental sample
Experimental features	16S rRNA metagenome sequencing
Analysis using	QIIME, MEGAN5
Consent	Not applicable
Sample source location	Water, estuary, Sundarbans, India

Bacterioplankton play key roles in biogeochemical cycling through the microbial loop in marine environment including coastal ecosystems [1]. The composition and distribution patterns of bacterioplankton communities have been surveyed across various coastal ecosystems such as the Columbia estuary [8], Pearl estuary [12] and Delaware Bay [3] to understand their role in ecosystem processes. However, not much is known in terms of bacterioplankton community structure from mangrove ecosystems globally [7,9,11]. Sundarbans, the world's largest contiguous mangrove ecoregion, provides a unique set up to investigate and understand the structure and functional significance of bacterioplankton communities. Seasonal variation in surface water temperature, heavy local precipitation during monsoon, continuous flow of freshwater from Ganga-Brahmaputra-Meghna riverine systems, diurnal intrusion of saline water from Bay of Bengal and dynamicity in dissolved nutrients could act as stressors for bacterioplankton communities of Sundarbans. We analysed the bacterioplankton communities by constructing 16S rRNA clone libraries and subsequent sequencing of individual clones by Sanger sequencing method from extracted environmental DNA from two stations representing the Sundarbans Biological Observatory Time Series (SBOTS). High-throughput sequencing using Illumina MiSeq approach was then undertaken from the same

* Corresponding author.
E-mail address: pbhadury@iiserkol.ac.in (P. Bhadury).

http://dx.doi.org/10.1016/j.gdata.2016.11.017

set of environmental DNA to obtain a deeper insight into bacterioplankton community structure. This study was undertaken in SBOTS located in Sagar Island, the largest island of the Indian Sundarbans. Two spatially separated stations designated as Station 1 (Stn1; 21° 44′ 44.4″ N, 88° 08′ 49.5″ E) and Station 3 (Stn3; 21° 40′ 40.6″ N, 88° 09′ 19.2″ E) as part of SBOTS were selected for this study. One litre of surface water sample was collected from each station in July 2014 following standard published protocol [5]. Collected samples were immediately fixed with molecular grade alcohol and transferred to the laboratory. Biomass was concentrated by filtering water samples through a 0.22 μm nitrocellulose filter paper (Pall, USA) using standard methodology [5]. The filters were immediately stored at −20 °C until further downstream processing. Extraction of environmental DNA (eDNA) pool was undertaken from each filter following published protocol [2]. Clone libraries comprising of forty clones from each library was generated from both the stations as part of an ongoing study spanning from June to December 2014 (Ghosh and Bhadury, 2016, in prep.). In this study, data representing 40 clones from each station only for the month of July 2014 has been discussed. Since maximum heterogeneity in bacterioplankton communities was observed in the month of July 2014 based on clone library data (Sanger sequencing) therefore in order to get a deeper resolution of their community structure, the extracted eDNA from each station for the same month was also sequenced using Illumina MiSeq platform. The V3-V4 hypervariable region of ~460 bp was amplified using Pro340F (5′-CCTACGGGNBGCASCAG-3′) and Pro805R (5′-GACTACNVGGGTATCTAATCC-3′) primers for Illumina

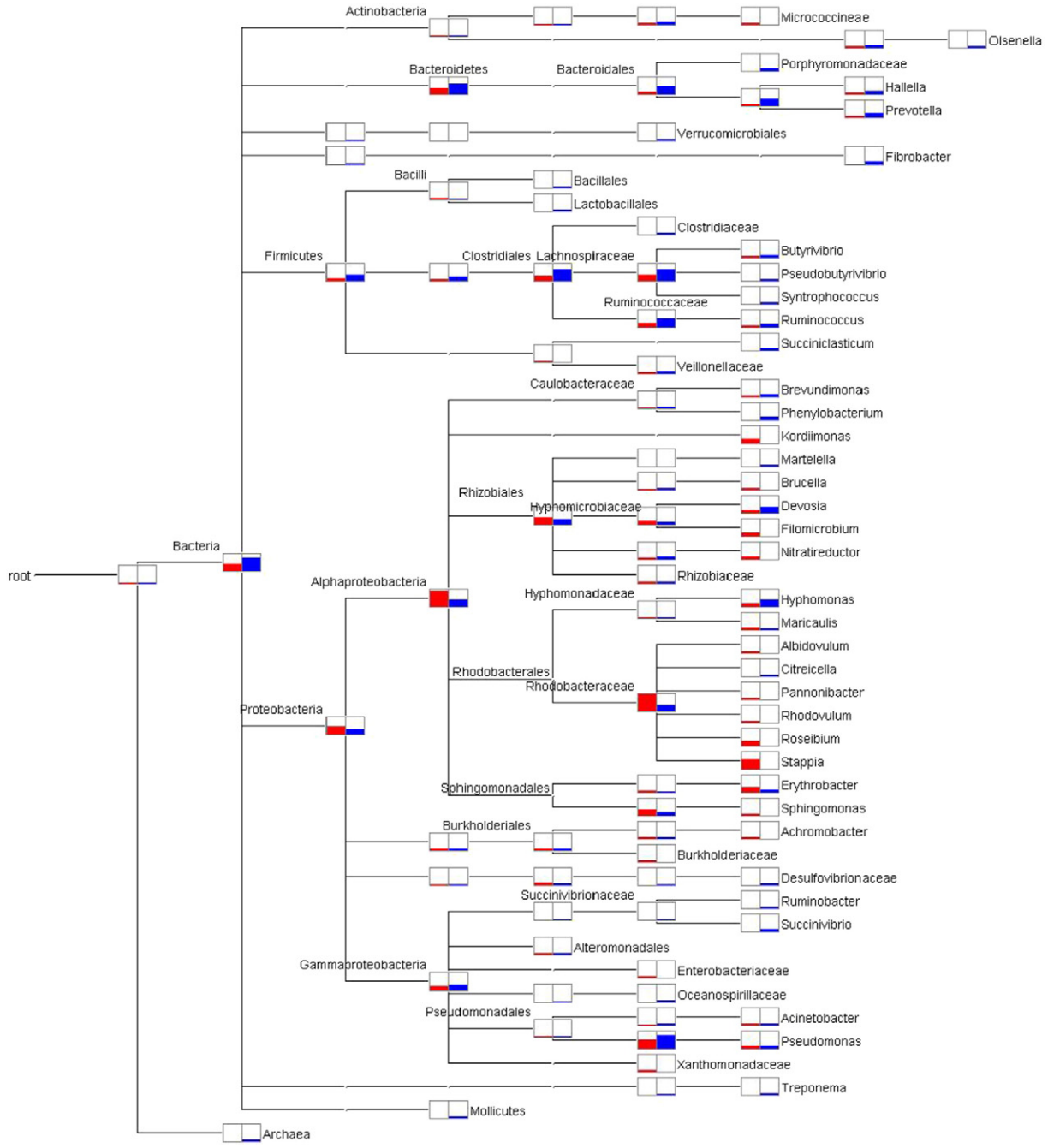

Fig. 1. (a): Phylogram indicating abundance (number of sequences) of 16S rRNA sequences recovered from Stn1 (in red colour) and Stn3 (in blue colour) by Illumina MiSeq sequencing. Taxa showing differential distribution between the two stations have been shown. (b): Phylogram indicating the abundance (number of sequences) of 16S rRNA sequences recovered from Stn1 (in red colour) and Stn3 (in blue colour) from the clone libraries followed by Sanger sequencing. Taxa showing differential distribution between the two stations have been shown.