Science of the Total Environment 637–638 (2018) 349–358

Contents lists available at ScienceDirect

Science of the Total Environment

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

Diversity and distribution of Archaea in global estuarine ecosystems

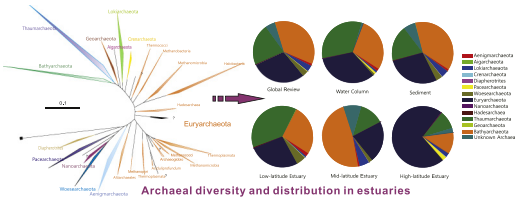
Xiaobo Liu [a], Jie Pan [b], Yang Liu [b], Meng Li [b,*], Ji-Dong Gu [a,c,**]

[a] Laboratory of Environmental Microbiology and Toxicology, School of Biological Sciences, Faculty of Science, The University of Hong Kong, Pokfulam Road, Hong Kong, People's Republic of China
[b] Institute for Advanced Study, Shenzhen University, Shenzhen, Guangdong, People's Republic of China
[c] State Key Laboratory in Marine Pollution, City University of Hong Kong, Kowloon Tong, Hong Kong, People's Republic of China

HIGHLIGHTS

- This is the first work to explore archaeal data in global estuarine ecosystems.
- Archaeal diversity and distribution patterns were systematically investigated.
- Estuarine ecosystem is a large biodiversity pool of Archaea.
- Archaeal distribution demonstrated a geographical differentiation in latitude.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 9 March 2018
Received in revised form 1 May 2018
Accepted 2 May 2018
Available online xxxx

Editor: F Frederic Coulon

Keywords:
Archaea
Diversity
Distribution
Estuary
Latitude

ABSTRACT

Estuarine ecosystem is a unique geographical transitional zone between freshwater and seawater, harboring a wide range of microbial communities including Archaea. Although a large number of Archaea have been detected in such ecosystem, the global patterns in archaeal diversity and distribution are extremely scarce. To bridge this gap, we carried out a comprehensive survey of archaeal communities using ca. 4000 publicly available archaeal 16S rRNA gene sequences (>300 bp) collected from 24 estuaries in different latitude regions. These sequences were divided into 1450 operational taxonomic units (OTUs) at 97% identity, suggesting a high biodiversity that increased gradually from the high- to low-latitude estuaries. Phylogenetic analysis showed that estuarine ecosystem was a large biodiversity pool of Archaea that was mainly composed of 12 phyla. Among them, the predominant groups were Bathyarchaeota, Euryarchaeota and Thaumarchaeota. Interestingly, archaeal distribution demonstrated a geographical differentiation in that Thaumarchaeota was dominated in the low-latitude estuaries, Bathyarchaeota in the mid-latitude estuaries, and Euryarchaeota in the high-latitude estuaries, respectively. Furthermore, the majority of the most abundant 20 OTUs demonstrated an overrepresented or underrepresented distribution pattern in some specific estuaries or latitude regions while a few were evenly distributed throughout the estuaries. This pattern indicates a potential selectivity of geographical distribution. In addition, the analysis of environmental parameters suggested that latitude would be one of the major factors driving the distribution of archaeal communities in estuarine ecosystem. This study profiles a clear framework on the diversity and distribution of Archaea in the global estuarine ecosystem and explores the general environmental factors that influence these patterns. Our findings constitute an important part of the exploration of the global ecology of Archaea.

1. Introduction

Estuary is a mixing zone between continental runoff freshwater and coastal seawater, where a strong physiochemical gradient may exist because of diurnal alterations and changes of many factors including tidal heights, winds, freshwater inputs and anthropogenic interferences (Bernhard and Bollmann, 2010; Vieira et al., 2007). Due to its unique geographical location and characteristics, estuary typically shows a

* Corresponding author.
** Correspondence to: J.-D. Gu, Laboratory of Environmental Microbiology and Toxicology, School of Biological Sciences, Faculty of Science, The University of Hong Kong, Pokfulam Road, Hong Kong, People's Republic of China.
E-mail addresses: limeng848@szu.edu.cn (M. Li), jdgu@hku.hk (J.-D. Gu).

https://doi.org/10.1016/j.scitotenv.2018.05.016
0048-9697/

sharp decrease of nitrogenous compounds and organic matters but an increase of sulfates and chlorides from the estuary head to mouth (Webster et al., 2015). As a large amount of nutrients and continental microorganisms are input and accumulated here, estuarine ecosystem tends to harbor high biodiversity and abundant substrates that support high levels of heterotrophic activities. These activities subsequently result in steep biogeochemical gradients along the vertical depth in sediments (Canfield and Thamdrup, 2009). Such gradients in both salinity and concentrations of organic or inorganic nutrients have already been reported to influence the estuarine microbial community structure (Webster et al., 2015; Xie et al., 2014; Zhou et al., 2017) and subsequently affect the macro-structure and function of estuarine ecosystem (Baird et al., 2004), particularly for the biogeochemical cycles. For example, metabolic reconstructions suggest different ecological roles of archaea in carbon, nitrogen and sulfur cycles in the sediment of White Oak River estuary (Lazar et al., 2017). Carbon metabolisms reveal that heterotrophic archaea may utilize sedimentary organic compounds, despite methanotrophy predominating the carbon cycle in estuarine sediments (Biddle et al., 2006). Furthermore, the large diversity of ammonia-oxidizing archaea (AOA) that belong to Thaumarchaeota marine group I (MGI) might lead to distinct life strategies in the environments from river to sea (Hugoni et al., 2015; Li et al., 2015a, 2015b). Thus, research on biodiversity and ecological roles of estuarine archaea becomes a hot topic of microbial ecology.

Based on the rapid development and the wide application of the high throughput sequencing technique, a great number of archaea have been detected in different estuaries. For example, Crenarchaeota and Euryarchaeota have been already reported to be distributed widely in estuarine sediments (Abreu et al., 2001). Many other archaea have also been proven to play key roles in the biogeochemical cycles in estuarine ecosystem, such as Bathyarchaeota (Lazar et al., 2016), Euryarchaeota (Methanomicrobia) (Kaku et al., 2005), Thaumarchaeota (Marine Group I, MGI) (Francis et al., 2005; Li et al., 2015a, 2015b) and Thorarchaeota (Seitz et al., 2016; Liu et al., 2018). Although a large number of studies have legitimately considered the biogeography of specific archaeal taxa in local estuaries, diversity and distribution patterns of Archaea in global estuarine ecosystem remain largely unknown.

Here we attempt to investigate the diversity and distribution patterns of Archaea in global estuaries based on the present available archaeal 16S rRNA gene sequences in the public databases to (i) uncover the community compositions of Archaea in global estuaries, (ii) explore archaeal geographical distribution patterns, for example, how archaeal species richness changes with the latitudinal gradient, and (iii) determine environmental factors that shape the distribution of archaeal communities. In addition, potential habitats harboring the highest biodiversity are also indicated for the discovery of new archaea.

2. Materials and methods

2.1. Construction of the archaeal 16S rRNA gene dataset

Archaeal 16S rRNA gene sequences were extracted from GenBank database by using Esearch utility to search for records containing the following terms: '16S AND 300:2000[Sequence Length] AND archaea [Organism] AND rrna[Feature key] AND isolation_source[estuarine OR estuary OR river mouth] NOT genome OR chromosome OR plasmid'. Either environmental samples that underwent some modification (e.g., enrichment cultivation) prior to extraction of the DNA used to generate the sequences or studies with <10 archaeal sequences were excluded from the data. As a result, ca. 4000 sequences (no nucleotide ambiguities or chimeras present, checked by QIIME) were obtained from 24 estuaries (by December 2017, Supplementary material 1: Table S1; Supplementary material 2: Dataset S1). Subsequently, these estuaries were designated as 7 estuaries in the low-latitude region (<29 °), 11 in the mid-latitude region (29–50 °) and 6 in the high-latitude region (>50 °) (Supplementary material 1: Fig. S1). Variations in methodologies and sampling efforts among studies were homogenized, and sequences were clustered at a threshold of 97% identity using MOTHUR (Schloss et al., 2009). As a result, 1450 representative sequences (OTUs) were achieved from the 24 estuaries (Supplementary material 1: Table S1). These representative sequences were further analyzed using the following two methods: (i) explicitly phylogenetic analysis and (ii) archaeal taxonomic analysis.

To clearly identify archaeal taxa in estuaries, we referred to the latest archaeal 16S rRNA reference database of the SILVA Release 132 (https://www.arb-silva.de/documentation/release-132/). Our 16S rRNA gene sequence dataset was classified separately into (i) three categories: low-latitude (<29 °), mid-latitude (29–50 °) and high-latitude estuaries (>50 °) (based on the latitude of the sampling site, Supplementary Material 1: Table S1) and (ii) two categories: water column and sediment (based on the isolation source of the samples). Then, a semi-quantitative environmental matrix was compiled for each sample based on the gradient of environmental factors (Supplementary material 1: Table S1): latitude (low to high), temperature (psychrophile or mesophile), oxygen (anoxic or oxic) and salinity (non-saline to hypersaline).

2.2. Phylogenetic analysis

The representative 16S rRNA gene sequences were aligned with the SINA aligner (https://www.arb-silva.de/aligner/) and then imported into the ARB software ((Ludwig et al., 2004); http://www.arb-home.de) that was already loaded with the SILVA database. Highly variable positions were checked and removed using the base frequency filter. After that, sequences were added into the maximum parsimony backbone tree using the 'parsimony quick add marked tool' implemented in the ARB software, thereby maintaining the topology of tree by default (Auguet et al., 2010). The phylogenetic tree of Archaea in estuaries was constructed using the 1450 OTU representative sequences. Archaeal sequence affiliations to each phylum were conducted by comparing with the reference database of the SILVA Release 132. For phylogenetic tree construction, the inference was performed using the RAxML version 7.7.1 to evaluate large phylogenies by the maximum likelihood approach (Stamatakis et al., 2008). The optimal phylogenetic tree was obtained using the GTRCAT model with 1000 bootstrap replicates and drawn with the iTOL (Letunic and Bork, 2016).

2.3. Novelty of the archaeal lineages in estuaries

Sequences that were separated from the known clusters might remain as unaffiliated lineages. We, therefore, named them as unknown archaea in the tree. To identify the possibility of novelty of these lineages, sequences were searched against the reference database of the SILVA Release 132 by BLASTn. The threshold *e*-value of 1×10^{-10} was set and the allowed maximum of target sequence was 6 for each query sequence. Reference sequences that were matched at the highest identity were chosen and the corresponding identity values were collected for further comparison.

2.4. Phylum dominance analysis

To investigate the phylum composition of archaeal communities, a phylum-level abundance table was extracted by summing up the abundance of OTUs designated to the same phylum. Several OTUs that could not be assigned to any phylum were treated as unknown archaea. Abundance of each phylum was compared based on both the latitude (i.e., low, mid and high) regions and the environments (i.e., water column and sediment) of samples. The dominant archaeal phylum was identified by the OTU abundance in the corresponding categories.