Marine Genomics 1 (2008) 55–61

Contents lists available at ScienceDirect

Marine Genomics

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

Evidence of biogeography in surface ocean bacterioplankton assemblages

Luisa I. Falcón *, Ana M. Noguez, Laura Espinosa-Asuar, Luis E. Eguiarte, Valeria Souza

Laboratorio de Evolución Molecular y Experimental, Instituto de Ecología, UNAM, Mexico

ARTICLE INFO

Article history:
Received 18 June 2008
Accepted 30 June 2008

Keywords:
Bacterial assemblages
Community structure
Beta diversity
Environmental heterogeneity

ABSTRACT

Regardless of the importance of bacterial assemblages as essential components of ecosystems, little is known about how their populations are structured. We analyzed the composition and turnover rates, based on 16S r*DNA* sequences, of surface water oceanic bacterial assemblages of the fraction between 0.1 and 0.8 μm along a latitudinal gradient (45°6′42′′N in the North Atlantic to 15°8′37′′S in the South Pacific) including geographic distance, temperature, chlorophyll *a* and salinity. Here we show that oceanic bacterial assemblages between 0.1 and 0.8 μm, can be structured by a variety of environmental interactions that include separation by distance and chlorophyll *a* concentration in temperate North Atlantic coastal samples and temperature in tropical Atlantic and Pacific coastal and open ocean samples. Bacterial phyla composition diverged between temperate and tropical regions. This study suggests that some bacterial assemblages could be structured both by environmental and spatial factors, while others by environmental factors alone.

1. Introduction

Despite the importance of prokaryotes in biogeochemical cycles we know almost nothing about how their populations are being structured and about the processes behind their patterns of distribution (Dolan, 2005). The traditional point of view posed by Beijerink (1913) and later by Baas-Becking (1934) stating "everything is everywhere and the environment selects" is still a major paradigm among microbial ecologists. The previous statement implies a low rate of species turnover with spatial distance (beta diversity) and as a function of environmental gradients (Finlay and Fenchel, 1999; Fierer et al., 2007; Ramette and Tiedje, 2007a,b). The studies conducted on this topic are not conclusive and the literature shows a wide spectrum of results, from those where patterns of distribution are found (Cho and Tiedje, 2000; Papke et al., 2003; Whitaker et al., 2003; Horner-Devine et al., 2004a,b; Noguez et al., 2005; Green and Bohannan, 2006; Hughes-Martiny et al., 2006; Horner-Devine et al., 2007; Pommier et al., 2007, to name a few), to those that suggest cosmopolitan distributions for some groups of prokaryotes (Purdy et al., 2002; Brinkmeyer et al., 2003; Selje et al., 2004; Baldwin et al., 2005).

Detailed information regarding patterns of distribution of prokaryote species is limited (Papke et al., 2003; Horner-Devine et al., 2004a, b; Noguez et al., 2005) and it has been hindered due to various causes that include cultivation complexity and the ambiguous definition of species (Dolan, 2005; Ramette and Tiedje, 2007a). Fortunately, in the last 15 years, new molecular techniques have allowed us to explore the enormous diversity of the microbial world where the analysis of regions of the 16S r*DNA* gene has been extensively used (Woese et al., 1990). The recent sequencing of communities by shotgun (Venter et al., 2004) and the subsequent massive effort to describe bacterial assemblages in surface ocean waters, such as the Global Ocean Sampling (GOS) (Rusch et al., 2007) and others (Pommier et al., 2007; Yutin et al., 2007), have opened a whole new perspective on bacterial diversity not only at the taxonomic level but also at the protein level (Yooseph et al., 2007), giving us a finer glimpse of the diversity of prokaryotes and their potential function in the oceans.

Beta diversity or species turnover provides information about distributional ranges of species and can be calculated: a) regarding geographic distance, which indicates limitations of species dispersal capacity; and b) regarding habitat variables, then suggesting the degree of habitat constraint over species distribution (McKnight et al., 2007; Soininen et al., 2007). Recent work has shown strong beta diversity in prokaryote assemblages in soil (Horner-Devine et al., 2004b; Noguez et al., 2005; Fierer and Jackson, 2006; Rammette and Tiedje, 2007a,b) and streams (Fierer et al., 2007). New attempts to discuss the theoretical bases of prokaryote ecology in the light of their patterns and processes (Green and Bohannan, 2006; Hughes-Martiny et al., 2006) have addressed the following main questions: What happens with microorganisms with different capacity for dispersal? Do their small sizes and high abundances prone them to low extinction and low speciation rates? Or their lack of dispersal barriers prevents them from geographic isolation? The challenge as Green and Bohannan (2006) suggest is to look for "… a spatial scale, a degree of sampling effort and a level of taxonomic resolution at which microbial biodiversity scaling relationships approach those of macroorganisms".

* Corresponding author. Laboratorio de Evolución Molecular y Experimental, Instituto de Ecología, UNAM. Circ. Ext. Cd Universitaria, CP 04510 México DF, Mexico. Tel.: +52 55 5622 9060; fax: 52 55 5622 8990.
E-mail address: falcon@ecologia.unam.mx (L.I. Falcón).

1874-7787/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.margen.2008.06.005

The Global Ocean Sampling expedition (GOS) (Rusch et al., 2007) constitutes the first in-depth attempt for deciphering prokaryote assemblage diversity present in surface oceanic waters along the globe using a metagenomic approach. This study showed, among other things, that genomic diversity is much greater than previously estimated, and recognized three major clusters based on metagenomic sequence similarity pertaining to temperate Atlantic, tropical Atlantic and tropical Pacific Oceans, including both coastal and open ocean sites. Here we present a detailed analysis aimed at identifying patterns of beta diversity based on 16S r*DNA* sequences for surface water column oceanic bacterial assemblages of the small bacterioplankton fraction between 0.1 and 0.8 μm along a latitudinal gradient from 45°6′42″N in the temperate North Atlantic to 15°8′37″S in the tropical South Pacific. We test for biogeographic patterns of prokaryotes taking into account geographic distances, and environmental variables that include temperature, chlorophyll *a* concentration and salinity.

2. Materials and methods

2.1. Regionalization

Sampling locations were approximately 200 nautical miles (~320 km) apart starting in the Atlantic coast of Canada, along the coast to Florida (USA), Gulf of Mexico, Caribbean, across the Panama Canal and into the Pacific Ocean, including coastal and open ocean sites (Fig. 1). Three regions were considered based on similarity of 16S r*DNA* sequences that were consistent with those defined comparing shotgun datasets (Fig. 10 in Rusch et al., 2007). These regions were temperate North Atlantic (region I), tropical Atlantic (region II) and tropical Pacific (region III). Region I had its northernmost location in Bay of Fundy, Nova Scotia, Canada (45°6′42″N; 64°56′48″W) and its southernmost location south of Charleston, NJ, USA (32°30′25″N; 79°15′50″W), including stations 2–14. This region consisted of coastal localities. Region II included coastal and open ocean stations 15–19 from Key West, FL (USA) (24°29′18″N; 83°4′12″W) to northeast of Colón in Panama (10°42′59″N; 80°15′16″W). Region III was located in the tropical Pacific Ocean, including stations 22, 23, 25–29, 31, 34–37, 47 and 51, spanning from 6°29′34″N; 82°54′14″W to 15°8′37″S; 147°26′6″W.

2.2. Data

Analyses were based on 16S r*DNA* sequences obtained as part of GOS (Rusch et al., 2007). The available 16S r*DNA* dataset included 4125 fragments from the small bacterioplankton fraction (0.1–0.8 μm). The metagenomic dataset used was created with shotgun sequencing, which generates different size fragments from all genomic regions of the total *DNA* for the sampled community thus, each resulting ribotype sequence was of different size and corresponded to a different region of the 16S r*DNA* gene. To avoid overestimations that would be generated when considering fragments of the same organism from a different 16S r*DNA* region, hence of site-specific richness and patterns of replacement (β diversity) of the bacterial community, we selected only those fragments that aligned in the same region resulting in a dataset of 810 sequences (*E. coli* 16S r*DNA* gene position 103 to 597). Further, in order to include only individual entities, sequences that shared 98% similarity were grouped using DOTUR version 1.53 (http://www.plantpath.wisc.edu/fac/joh/dotur.html), which were defined as Organizational Taxonomic Units (OTUs). The 98% cutoff value to define OTUs generated 292 groups distributed in the three geographic regions analyzed.

Sequences were downloaded from the Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis (CAMERA) web page (http://camera.calit2.net/), and were aligned in Sequencher version 4.1.4 (Gene Codes Corporation, Ann Arbor, MI). Genetic similarity between each pair of locations per region was calculated with the classic Jaccard index (*J*) using estimates (Chao et al., 2005). The classic Jaccard index (*J*) ranges from 0 to 1 and is calculated as: $J=A/A+B+C$, where A = the number of species shared by two assemblages, B and C = the number of species unique to each assemblage.

2.3. Pair wise comparisons

Environmental parameters included temperature (°C), chlorophyll *a* concentration (mg m^{-3}), salinity (ppt) and geographic location (latitude, longitude) and were used to build pair wise comparisons between stations. Regressions between the natural logarithm of *J* ($\ln J+1$, due to presence of Jaccard values equal to zero), and pair wise values for each environmental variable were conducted, where *J* was

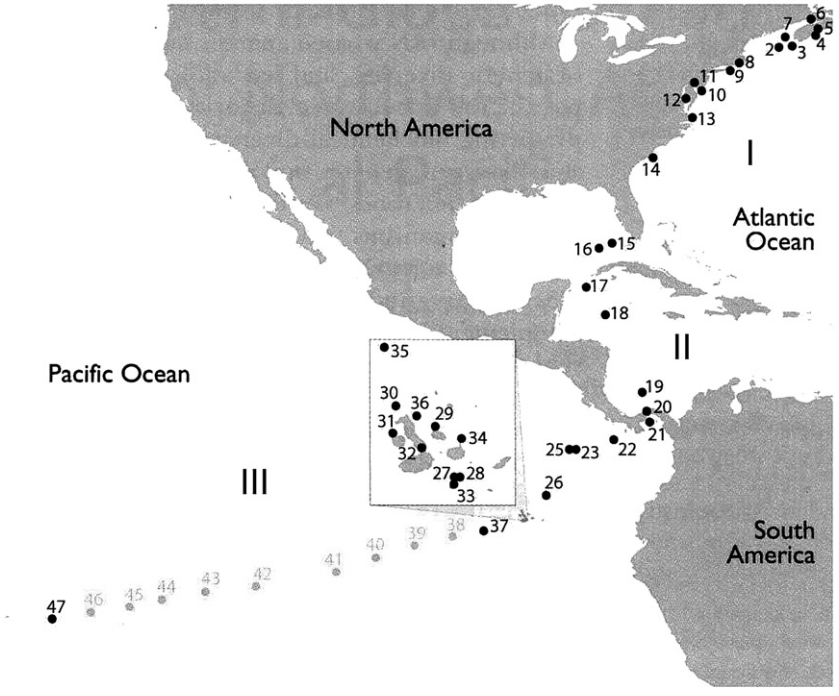

Fig. 1. Location of sampling sites indicating the three geographical regions analyzed. Figure modified from Rusch et al. (2007).