Phylogenetic profiling of bacterial community from two intimately located sites in Balramgari, North-East coast of India

Arvind Kumar Gupta · Ashraf Yusuf Rangrez · Pankaj Verma · Anil Chatterji · Yogesh S. Shouche

Received: 13 October 2008 / Accepted: 9 January 2009

Abstract Microbial communities in coastal subsurface sediments play an important role in biogeochemical cycles. In this study microbial communities in tidal subsurface sediments of Balramgari in the state of Orissa, India were investigated using a culture independent approach. Two 16S rDNA cloned libraries were prepared from the closely located (100 m along the coast) subsurface sediment samples. Library I sediment samples had higher organic carbon content but lower sand percentage in comparison to Library II. A total of 310 clone sequences were used for DOTUR analysis which revealed 51 unique phylotypes or operational taxonomic units (OTUs) for both libraries. The OTUs were affiliated with 13 major lineages of domain bacteria including *Proteobacteria* (α , β , δ and γ), *Acidobacteria*, *Actinobacteria*, *Cyanobacteria*, *Chloroflexi*, *Firmicutes*, *Verrucomicrobia*, *Bacteroidetes*, *Gemmatimonadetes* and TM7. We encountered few pathogenic bacteria such as *Aeromonas hydrophila* and *Ochrobactrum intermedium*, in sediment from Library I. J-LIBSHUFF comparison depicts that the two libraries were significantly different communities. Most of the OTUs from both libraries possessed $\geq 85\%$ to $< 97\%$ similarity to RDP database sequences depicting

the putative presence of new species, genera and phylum. This work revealed the complex and unique bacterial diversity from coastal habitat of Balramgari and shows that, in coastal habitat a variability of physical and chemical parameter has a prominent impact on the microbial community structure.

Keywords Ecosystem · Microbial diversity · Marine sediment · 16S rDNA

Introduction

For millions of years after emergence of the first life forms, microbial life in the oceans influenced the planet's chemistry, altering the chemical balance of the oceans and atmosphere and introducing gradients of oxidizing and reducing agents. Marine microbial life thrives not only in the surface waters, but also in the lower and abyssal depths from coastal to the offshore regions, and from the general oceanic to the specialized niches such as blue waters of coral reefs to black smokers of hot thermal vents at the sea floor [1]. These microbes harbor a unique metabolic machinery to carry out many steps of the biogeochemical cycles, the smooth functioning of these cycles is necessary for life to continue on earth.

Our perspective on microorganisms in the environment has depended in the past primarily on studies of pure cultures in the laboratory. This view of microbial diversity is limited as it is estimated that more than 99% of organisms seen microscopically are not cultivated by routine techniques [2]. Now with the advent of sequence-based taxonomic framework of molecular phylogeny, one requires

A. K. Gupta¹ · A. Y. Rangrez¹ · P. Verma¹ · A. Chatterji² · Y. S. Shouche¹ (✉)

¹Molecular Biology Unit,
National Centre for Cell Science,
Pune - 411 007, Maharashtra, India

²Institute of Tropical Aquaculture,
University Malaysia Terengganu, Malaysia

E-mail: yogesh@nccs.res.in

the gene sequences for identifying the types of organism occurring in any microbial community. Functioning and health of an ecosystem is determined to be contingent on its biodiversity and community structure. Thus, it is important to study the occurrence of prokaryotic phylotypes and their distributions in natural communities, which can be carried out by sequencing 16S rRNA genes obtained from DNA isolated directly from the environment [3–5].

The northern part of the Bay of Bengal enjoys a tropical climate, governed mainly by the monsoons. The Southwest monsoon brings rainfall to the coastal areas from June to October. The temperature oscillation between summer and winter is less than 5°C. In contrast to the small variations in the temperature, the salinity showed wide fluctuations (22.4% to 33.4%). This was largely due to the riverine flow from two major river systems, the Ganges and Brahmaputra, and a number of smaller rivers such as the Subarnarekha, Budhabalang, Dhamra and Mahanadi [6]. A large number of Indian horseshoe crabs have been reported to migrate towards the sandy shore of Balramgari, a unique beach along north-east coast of India, for breeding purpose throughout the year. The geo-morphological and eco-bio-

logical characteristics of this beach are totally different with respect to any other beaches of India [7]. These features make this environment unique and directs toward exploring the environmentally important diversity.

The aim of this work was to construct and analyse the 16S rDNA clone libraries to elucidate the genetic diversity and differences of microbial communities in subsurface sediments of Balramgari, North-East coast of India.

Materials and methods

Sample collection

Sediment samples utilized for the construction of libraries were collected from two points (for library I and II) from the coast of Balramgari, Orissa (Lat 19°16' N; Long 84°53' E), India (Fig. 1). The collection of the sediment was done in coincidence with the lowest low tide at 1031 hours (tide height ~0.31 m) on 23 July 2005. Four subpoints at a distance of 5 m away from each point at midtide level were selected for collecting the sediment samples. Sediment

Fig. 1 Map showing the place of sample collection, Balramgari, India (Lat 19°16' N; Long 84°53' E)

samples were collected by using a PVC core (depth 5 cm and diameter 5 cm).

Sediment samples for Library I (L-I) construction were collected ~100 m away along the coast towards south from sediment samples collected for Library II (L-II). Samples were immediately transported in sterile bottles in dry ice and stored at -80°C until analysis. The temperature and pH at each sampling site were measured using a centigrade thermometer and a portable pH meter (Philips 4012). The sediment samples were analyzed for their grain size distribution and organic carbon content [8]. The water content of the sediment was estimated immediately after the collection by drying a known weight of it at 100°C for 3 hours followed by the re-weighing of the sample. The loss in weight was expressed as a percentage of water content of the sediment. The organic content of the sediment was estimated using the chromic acid oxidation method [9]. The description of physicochemical characteristics of sediments are summarized in Table 1.

Extraction of total community DNA from sediments

Sediment cores were sliced and upper 0–1 cm layer was removed. Sediments from 1–3 cm sections were mixed and 1 g of the mixture was used for DNA extraction. Total community DNA was extracted from all eight samples using MoBio soil DNA isolation kit (Imperial Bio-Medic) as per the protocol provided by the manufacturer and the integrity of the extracted DNA was checked by 0.8% horizontal agarose gel electrophoresis and ethidium bromide staining. The concentration of extracted DNA was checked with the help of NanoDrop ND-1000 Spectrophotometer (NanoDrop Biotechnologies, USA) and was ranged between 200 and 300 ng/ μl . DNA from replicate samples was pooled and stored at -20°C .

PCR and cloning of 16S rRNA gene

16S rRNA gene amplification was performed from both the pooled extracts in triplicates using universal eubacte-

rial primer set 530F (5' GTC CCA GCM GCC GCG G 3') and 1490R (5' GGT TAC CTT GTT ACG ACT T 3') [10]. Amplification was carried out in a 25 μl reaction mixture containing 200 μM (each) dNTPs, 25 pM each primer, 1 μl (~50 ng) DNA template and 2.5 U of *Taq* DNA polymerase (Bangalore Genei, Bangalore, India) with 1X reaction buffer supplied by the manufacturer. The reaction mixtures were incubated in GeneAmp PCR System 9700 (Applied Biosystems) at 94°C for 5 min (initial denaturation and activation of *Taq* DNA polymerase), followed by 25 cycles at 94°C for 1 min (denaturation), 55°C for 1 min (annealing), and 72°C for 1 min (extension), followed by final extension for 10 min at 72°C . A tube without DNA was taken as negative control. Amplification was done only for 25 cycles to minimize the bias in PCR amplification. 2 μl of amplified DNA was examined by horizontal electrophoresis on 1% agarose gel in TAE buffer (40 mM Tris, 20 mM acetate, 2mM EDTA). To minimize the PCR drift, all the three amplified PCR products from each site were pooled and purified with Qiagen PCR purification kit (Qiagen, USA) as per manufacturers' instruction.

16S rRNA gene library construction and sequencing

Eubacterial libraries were constructed from pooled and purified PCR product by cloning into pGEM-T easy vector (Promega, USA) according to manufacturer's instructions. Transformation was done in chemically competent *Escherichia coli* JM109 cells (Stratagen) with 30 minutes recovery time. Positive colonies from the library were picked out and screened by colony PCR using vector specific primers. The amplification reaction described as above was used and the PCR conditions were as follows: incubation for 7 min at 94°C (mainly for cell rupture), 35 cycles each of 1 min at 94°C (denaturation), at 55°C (primer annealing) and at 72°C (primer extension) and a final elongation step for 10 min at 72°C . A total of 323 clones with proper insert (~960 bp) were purified and sequenced in both directions with an automated ABI-PRISM 3730 system (Applied BioSystems Inc.).

Table 1 Analysis of physicochemical parameters of the sediment samples

No.	Parameters	Library I sediment	Library II sediment
1	Sediment temperature	30.5°C	31.6°C
2	Sediment pH	7.5	7.6
3	Salinity	16.0×10^{-3}	17.2×10^{-3}
4	Organic carbon content	2.12 mg c/g	0.46 mg c/g
5	Sand percentage	78.32%	83.47%
6	Silt and clay	21.68%	16.53%
7	Percentage of water content	5.0%	5.9%
8	Mean grain size	0.129 mm	0.231 mm

Real-time PCR

Qualitative PCR was performed using 7300 Real Time PCR System (Applied Biosystems) and Power SYBR Green PCR master mix supplied by Applied Biosystems. Amplification was carried out in a 25 µl reaction mixture containing 50 ng of DNA, 25 pM each primer and Power SYBR Green PCR master mix. The conditions for the amplification of *ascV* gene (*Aeromonas* specific gene) (ascV-F 5' TAA RCA GAT GAG TAT CGA TGG 3' and ascV-R 5' GAG ACS CGG GTG ACG ATA AT 3'), and *recA* gene (*Ochrobactrum* specific gene) (RecA_F 5' GCG CCG AAA TCG AAG GT 3' and RecA_R 5' GCG AAC CGA ACA TCA CAC C 3') and 16S rRNA gene (as an internal control), were as follows: One denaturation step at 95°C for 10 min, 40 cycles consisting of denaturation at 95°C for 15 sec, a common step for annealing and elongation at 60°C for 60 sec. At the end of the PCR, the samples were subjected to melting curve analysis.

Taxonomic and phylogenetic assessment

The partial DNA sequences obtained with the vector specific primers M13F and M13R were assembled and edited with ChromasPro version 1.33 software. (www.technelysium.com.au/ChromasPro.html). Vector sequences were removed from both the ends. Good quality sequences of approximately 900 nucleotides were used for subsequent analysis. Both libraries were analysed for chimeric sequences and other anomalies using MALLARD software [11] by using pair-wise comparisons within a multiple alignment. Each putative chimera identified by the program was checked with BLASTn [12] and further compared with closest cultured similar rDNA sequences retrieved from the DNA databases. Suspected chimeras were excluded from further analyses. Multiple sequence alignments were performed using ClustalW version 1.8 [13]. Multiple sequence alignment was edited and corrected manually using DAMBE [14] to get unambiguous sequence alignment. Appropriate subsets of 16S rDNA sequences were selected on the basis of initial results and subjected to further phylogenetic analysis using the neighbour-joining method implemented through DNADIST from the PHYLIP version 3.61 [15]. OTUs were generated using furthest neighbour algorithm of DOTUR program [16]. OTUs generated at 0.03 E.D. (Evolutionary Distance) or the OTUs formed by the sequences that present a similarity equal or greater than 97% [17] were used for taxonomic assessment. One representative clone sequence from each OTU was taken for taxonomic assessment and it was carried out using sequence match programme of RDP (www.rdp.cme.msu.edu/

[seqmatch/seqmatch_intro.jsp](#)). Phylogenetic analysis was carried out using bayesian inference method with program MrBayes 3.0 [18]. The analysis for 16S rRNA gene consisted of 3,000,000 generations. An appropriate model of sequence evolution for data set was chosen via Akaike Information Criterion (AIC) using program MrModeltest 2.2. The model selected was (GTR+I+G) for both the data sets. Trees were sampled for every 100 generations. First 3000 trees (10%) were discarded as burnin. Bayesian posterior probabilities were calculated using 50% majority rule consensus and three independent runs were performed for each dataset. Representative sequences have been assigned GenBank accession numbers EF451854-EF451955 and EU236787-EU236933.

Statistical comparison of clone libraries

Good's coverage [19] was calculated using formula, $[1-(n/N)] \times 100$ (where n is the number of single clone OTUs and N is the library size). Rarefaction curves were drawn using the algorithm described by Hurlbert [20] and plotted as the number of OTUs versus the number of clones, assuming that one OTU is formed by the sequences that present a similarity equal or greater than 97%. Two libraries generated were compared using J-LIBSHUFF program version 1.21 [21] which uses a Monte Carlo procedure to calculate the probability that the observed difference between two libraries are due to chance. Biodiversity indices were determined at E.D. 0.03 or a sequence similarity value of 97% for bacterial population using the Shannon Index ($H' = -\sum p_i \ln p_i$) which was used as a measure of diversity including richness and evenness, the Simpson Dominance Index ($SI' = n(n-1)/N(N-1)$) [22] and the ChaoI estimator as an alternative to Shannon diversity.

Results

Analysis of 16S rDNA clone libraries

Two hundred white colonies each from both the libraries were randomly picked and amplified, out of which, 323 positive clones having around 960 bp 16S rDNA insert were sequenced in both direction using vector specific M13F and M13R primers. Among the 323 rDNA sequences from the sediment analyzed, we have found 9 of 171 sequences from L-II to deviate by more than 5% to the *E. coli* K-12 16S rDNA sequence taken as standard [23] in the MALLARD programme, so considered as chimeras. We found only 4 chimeras out of 152 sequences from L-I. All chimeric sequences found, can be split into two fragments which show a reasonably high degree of homology to two very different

bacteria, however none of the other sequences showed this phenomenon. These chimeras were neither submitted to database nor included for further analyses. Thus, chimeras obtained in this study were much less than those reported by Wang and Wang [24]. We found that increasing the replicates (three replicates in this study) for PCR and reducing the reaction cycles to 25 was helpful in reducing PCR drift. Finally we had 148 clones from L-I and 162 clones from L-II for further analysis. Using DOTUR programme we obtained 51 OTUs each from both libraries at 0.03 evolutionary distance. Each OTU represent a phylotypes and may be representative of a bacterial species.

Taxonomic distribution of 16S rDNA clone

The distribution of clones in each OTU, their closest cultured and uncultured homologs with similarity score to sequence match of RDP database for both library were given in detail in Tables 2 and 3. The clones obtained were distributed in various groups of bacteria such as *Proteobacteria* (α , β , δ and γ), *Acidobacteria*, *Actinobacteria*, *Cyanobacteria*, *Chloroflexi*, *Firmicutes*, *Gemmatimonadetes*, *Bacteroidetes*, *Verrucomicrobia* and Candidate division TM7 (Fig. 2). *Proteobacteria* was the dominant class of bacteria identified in both the libraries as also been represented in other studies. L-I was composed of 80.38% whereas L-II contained 88.88% of the total clones of *Proteobacteria*, unevenly distributed within four subdivisions α , β , γ and δ . Among *Proteobacteria*, γ -*proteobacteria* was the dominant subdivision in both the libraries with 59.44% clones in L-I

and 67.67% in L-II. The second most abundant subdivision was α -*proteobacteria* with 16.21% clones in L-I while δ -*proteobacteria* with 11.11% clones in L-II. β -*proteobacteria* were least represented with around 0.67% and 3.7% clones in L-I and L-II respectively. The second dominant taxon in L-I was the high G + C gram-positive bacteria (*Actinobacteria*) with 6.08% clone, while in L-II, it was *Acidobacteria* with 4.3% clones. All the clones affiliated in the class of *Actinobacteria*, *Acidobacteria*, *Bacteroidetes*, *Gemmatimonadetes*, Candidate division TM7 and unclassified bacteria showed homology with uncultured bacterium of soil and sediments with similarity score between 89 and 98% and did not match with any closest cultured relative (Tables 2 and 3). Although, both libraries share common lineages, certain differences were observed, *Gemmatimonadetes* (2.02% clones), *Firmicutes* (2.7% clones) and Candidate division TM7 (1.35% clones) were present in L-I but absent in L-II whereas *Cyanobacteria* (1.23% clones) and *Chloroflexi* (0.61% clones) were present only in L-II (Fig. 2).

Phylogenetic reconstruction of 16S rDNA clones

Phylogenetic reconstruction was carried out using Bayesian inference to determine the relationship of these OTUs to known sequences of database. Phylogenetic reconstruction revealed that all OTUs were clustering within 13 major clades belonging to different division of bacteria. Phylogenetic clustering of all OTUs from both libraries into different clades was shown in Figs. 3 and 4.

Fig. 2 A comparative bar chart indicating the percentage distribution of clones in different taxa identified in Library I (L-I) and Library II (L-II)

Table 2 Summary of taxonomic assessment of bacterial rDNA sequenced clones from Library I using RDP database

OTUs	Clone name	No. of clones	Cultured similarity score	Accession no. of cultured hits	Nearest phylogenetic cultured neighbor	Phylum	Uncultured Similarity Score	Accession no. of uncultured hits	Nearest phylogenetic uncultured neighbor
OTU1	BE-373	1				Acidobacteria	0.941	AF154083	Uncultured hydrocarbon seep bacterium BPC102
OTU2	BE-021	2				Gammaproteobacteria	0.969	DQ396316	Uncultured organism; ctg_NISA196
OTU3	BE-124	2				Deltaproteobacteria	0.928	DQ499326	Uncultured bacterium; CV106
OTU4	BE-306	1				Deltaproteobacteria	0.944	EF516565	Uncultured bacterium; FCPP581
OTU5	BE-338	3				Gemmatimonadetes	0.941	AY913406	Uncultured forest soil bacterium; DUNssu200 (-7B) (OTU#199)
OTU6	BE-383	1				Acidobacteria	0.932	EF125393	Uncultured bacterium; MSB-1B7
OTU7	BE-296	1				Gammaproteobacteria	0.947	AM117932	Gamma proteobacterium HAL40b
OTU8	BE-294	1				Acidobacteria	0.945	AF154083	Uncultured hydrocarbon seep bacterium BPC102
OTU9	BE-045	3	0.823	DQ402051	<i>Rhodobacter azotoformans</i> ; S3;	Alphaproteobacteria	0.826	AM270417	Uncultured organism; 17H9
OTU10	BE-293	1	0.918	AY987846	<i>Rhodovibrio</i> sp. 2Mb1;	Alphaproteobacteria	0.939	DQ917822	Uncultured <i>Ochrobactrum</i> sp.; BME35
OTU11	BE-336	1	0.848	AJ011330	<i>Phyllobacterium myrsinacearum</i> ; HM35;	Alphaproteobacteria	0.883	DQ395494	Uncultured organism; ctg_CGOAB08
OTU12	BE-329	1	0.92	DQ401091	<i>Rhodospirillaceae bacterium</i> CL-UU02;	Alphaproteobacteria	0.977	EF157245	Uncultured bacterium; 101-99
OTU13	BE-341	1	0.89	AY741146	<i>Azospirillum amazonense</i> ; 21R;	Alphaproteobacteria	0.922	AJ567552	Uncultured alpha <i>Proteobacterium</i> ; MBMPE38
OTU14	BE-328	1				Betaproteobacteria	0.977	AB252913	Uncultured beta <i>Proteobacterium</i> ; 242
OTU15	BE-309	1				Gammaproteobacteria	0.95	AY225635	Uncultured gamma <i>Proteobacterium</i> ; AT-s80
OTU16	BE-305	5	0.912	CP000453	<i>Alkalilimnicola ehrlichei</i> MLHE-1;	Gammaproteobacteria			
OTU17	BE-374	1	0.896	AF304195	<i>Methylobacter luteus</i> ; NCIMB 11914;	Gammaproteobacteria	0.908	DQ234131	Uncultured <i>Alcanivorax</i> sp.; DS047
OTU18	BE-384	1	0.883	AY298904	<i>Ectothiorhodospinus mongolicum</i> ; M9;	Gammaproteobacteria	0.905	AY328560	Uncultured bacterium; HOCiCi11
OTU19	BE-080	2	0.931	AF539776	<i>Pseudoalteromonas</i> sp. R6;	Gammaproteobacteria			
OTU20	BE-349	3	0.897	AB246771	<i>Myxobacterium</i> AT1-01;	Deltaproteobacteria	0.945	DQ811828	Uncultured delta <i>Proteobacterium</i> ; MSB-SC5

OTUs name	Clone name	No. of clones	Cultured similarity score	Accession no. of cultured hits	Nearest phylogenetic cultured neighbor	Phylum	Uncultured Similarity Score	Accession no. of uncultured hits	Nearest phylogenetic uncultured neighbor
OTU21	BE-024	1				Firmicutes	0.983	AF371787	Uncultured bacterium; p-2179-s959-3
OTU22	BE-004	2	0.995	AJ491303	<i>Caryophanon tenue</i> (T); type strain: DSM 14152;	Firmicutes			
OTU23	BE-048	3				Actinobacteria	0.946	EF632905	Uncultured bacterium; Par-s-84
OTU24	BE-368	1				Actinobacteria	0.968	DQ395467	Uncultured organism; ctg_CGOAA79
OTU25	BE-308	1				Actinobacteria	0.983	AY913337	Uncultured forest soil bacterium; DUNssu130 (+7A) (OTU#190)
OTU26	BE-331	3				Actinobacteria	0.895	DQ823199	Uncultured bacterium; ORCA-17F18
OTU27	BE-074	1				Actinobacteria	0.945	AF317767	Unidentified bacterium wb1_J07
OTU28	BE-009	2				TM7	0.924	AY930458	Uncultured bacterium; OC28
OTU29	BE-370	1				unclassified bacteria	0.971	DQ300602	Uncultured bacterium; HF130_C5_P1
OTU30	BE-319	1	0.937	DQ868668	<i>Rhodovulum</i> sp. SMB1;	Alphaproteobacteria	0.93	AM176847	Uncultured bacterium; SZB36
OTU31	BE-365	1	0.964	EF186075	<i>Rhodobacter</i> sp. DQ12-45T;	Alphaproteobacteria	0.978	DQ813946	Uncultured bacterium; aab57g01
OTU32	BE-096	1	0.964	AF136850	<i>Ketogulonicigenium robustum</i> ; X6L;	Alphaproteobacteria	0.959	AF007256	Uncultured alpha <i>Proteobacterium</i> ; GAI-5
OTU33	BE-016	3	0.928	AJ401206	<i>Rhodovulum</i> sp. AT2111;	Alphaproteobacteria	0.958	EF125428	Uncultured bacterium; MSB-2C6
OTU34	BE-038	2	0.952	AB258386	<i>Kaistobacter terrae</i> ; KCTC12630;	Alphaproteobacteria	0.969	AF423277	Uncultured soil bacterium; 565-2
OTU35	BE-040	1	0.974	AY258089	<i>Mesorhizobium</i> sp. DG943;	Alphaproteobacteria	0.971	EF125460	Uncultured bacterium; MSB-2G6
OTU36	BE-030	2	0.92	AY654823	<i>Mucus bacterium</i> 86;	Bacteroidetes	0.972	DQ395383	Uncultured organism; ctg_BRRAA64
OTU37	BE-070	1	0.999	DQ480144	<i>Erythrobacter</i> sp. D3043;	Alphaproteobacteria	0.985	DQ396252	Uncultured organism; ctg_NISA320
OTU38	BE-003	2	0.974	EF540469	<i>Sphingopyxis</i> sp. 1/4_C7_32;	Alphaproteobacteria	0.968	AY796041	Uncultured bacterium; 47mm65
OTU39	BE-109	5	0.997	AJ242582	<i>Ochrobactrum intermedium</i> ; OiC8-6;	Alphaproteobacteria	0.997	AY851687	Uncultured <i>Ochrobactrum</i> sp.; p3
OTU40	BE-110	46	0.976	DQ660915	<i>Methylophaga</i> sp. DMS010;	Gammaaproteobacteria	0.968	DQ490031	Uncultured bacterium; ODP-46B-02

OTUs	Clone name	No. of clones	Cultured Similarity Score	Accession no. of cultured hits	Nearest phylogenetic cultured neighbor	Phylum	Uncultured Similarity Score	Accession no. of uncultured hits	Nearest phylogenetic uncultured neighbor
OTU41	BE-343	10	0.964	DQ660915	<i>Methylophaga</i> sp. DMS010;	Gamma proteobacteria	0.959	DQ513013	Uncultured bacterium; FS140-15B-02
OTU42	BE-058	1	0.977	AB167031	<i>Marinobacter</i> sp. NT N115;	Gamma proteobacteria	0.984	AF513454	Alteromonadaceae bacterium LA50
OTU43	BE-372	1	0.999	X74677	<i>Aeromonas hydrophila</i> subsp. <i>hydrophila</i> (T); ATCC 7966T;	Gamma proteobacteria			
OTU44	BE-141	1	0.993	AB021194	<i>Bacillus niacini</i> (T); IFO15566;	Firmicutes	0.995	AY642567	Uncultured low G+C Gram-positive bacterium; LV60-10
OTU45	BE-010	2				Acidobacteria	0.967	DQ378243	Uncultured soil bacterium; M23_Pitesti
OTU46	BE-371	2				Acidobacteria	0.951	DQ395012	Uncultured Acidobacteriaceae bacterium; VHS-B4-48
OTU47	BE-356	1				Verrucomicrobia	0.929	EF516510	Uncultured bacterium; FCPN572
OTU48	BE-088	3	0.983	X87339	<i>Methylophaga thalassica</i> (T); ATCC 33146;	Gamma proteobacteria	0.977	DQ513013	Uncultured bacterium; FS140-15B-02
OTU49	BE-071	1	0.956	AB053124	<i>Alcanivorax</i> sp. Mho1;	Gamma proteobacteria	0.96	DQ234131	Uncultured <i>Alcanivorax</i> sp.; DS047
OTU50	BE-082	12	0.99	AB055205	<i>Alcanivorax</i> sp. K3-3; K3-3 (MBIC 4323);	Gamma proteobacteria	0.978	DQ396108	Uncultured organism; ctg_NISA076
OTU51	BE-334	1	0.944	DQ084461	<i>Pseudomonas</i> sp. FLM05-3;	Gamma proteobacteria	0.944	AY569287	Uncultured <i>Pseudomonas</i> sp.; YJQ-10

***α*-proteobacteria**

Out of 14 OTUs (No. 9, 10, 11, 12, 13, 30, 31, 32, 33, 34, 35, 37, 38 and 39) (27.45%) belonging to *α*-proteobacteria of L-I, except 1 OTU (OTU9), all are clustered with *α*-proteobacteria clade supported by 100 bootstrap value. Three OTUs (No. 10, 12 and 13) of *α*-proteobacteria clustered with *Rhodospirillaceae* bacterium CL-UU02 an isolate from urea-enriched seawater (Cho et al. unpublished, DQ401091). OTU35 of L-I had high 97.4% sequence identity to *Mesorhizobium* sp. DG943 an isolate from the dinoflagellate *Gymnodinium catenatum* known for paralytic shellfish poisoning [25]. OTU37 of L-I had very high 99.9% sequence identity to *Erythrobacter* sp. D3043 an isolate from Qing Dao Coast, an anoxygenic phototroph having large amount of carotenoides, it plays important role in cycling of both organic and inorganic carbon in the ocean (Li et al. Unpublished, DQ480144 and [26]). OTU39 of L-I with 5 clones had very high 99.7% sequence identity to

Ochrobactrum intermedium, an emerging human pathogen among the liver abscess, post-liver transplantation and in the bladder cancer patients, causing presumptive bacteremia [27]. OTU32 of L-I had 96.4% sequence identity to *Ketogulonigenium robustum* which had the ability to produce vitamin C from substrates like L-sorbose [28]. Phylogenetic clustering of above OTUs with their closest homologs organism was supported by 100 bootstrap value. OTU31 clustered with *Rhodovibrio* sp. 2Mb1, an isolate from maras salterns, a hypersaline environment in the peruvian andes [29]. OTU9 of L-I with three clones had only 82.3% sequence identity to *Rhodobacter azotoformans* strain S3, denitrifying phototrophic bacteria holding ability to reduce selenite to red elemental selenium [30]. This OTU had very low sequence similarity with *α*-proteobacteria. Phylogenetically this OTU clustered with *Myxobacterium* (AB246771), a *δ*-proteobacteria with 100 bootstrap confidence value. *α*-proteobacteria group of L-II had 6 OTUs (No. 5, 6, 7, 8, 9 and 10) (11.76%) with 12 clones. All these

Table 3 Summary of taxonomic assessment of bacterial rDNA sequenced clones from Library II using RDP database

OTUs	Clone name	No. of clones	Cultured similarity score	Accession no. of cultured hits	Nearest phylogenetic cultured neighbor	Phylum	Uncultured Similarity Score	Accession no. of uncultured hits	Nearest phylogenetic uncultured neighbor
OTU1	NE-100	1	0.96	AF170424	Sulfur-oxidizing bacterium ND11.1	Gammaproteobacteria	0.961	DQ811847	Uncultured gamma <i>Proteobacterium</i> ; MSB-5C2
OTU2	NE-257	2	0.984	Z67753	<i>Odontella sinensis</i>	Cyanobacteria	0.991	DQ521522	Uncultured bacterium; ANTLV2_F08
OTU3	NE-279	1				Chloroflexi	0.95	DQ154828	Uncultured bacterium; GN01-8.065
OTU4	NE-91	2	0.884	DQ660913	<i>Methylophaga</i> sp. DMS004	Gammaproteobacteria	0.878	DQ490031	Uncultured bacterium; ODP-46B-02
OTU5	NE-166	1	0.948	AY654775	<i>Mucus bacterium</i> 32	Alphaproteobacteria	0.955	DQ153134	Uncultured alpha <i>Proteobacterium</i> ; 06-03-45
OTU6	NE-11	5	0.954	AF513400	<i>Rhodobacter</i> sp.	Alphaproteobacteria	0.968	AY258094	Bacterium DG981
OTU7	NE-211	2	0.958	AM696304	<i>Rhodovulum marinum</i>	Alphaproteobacteria	0.948	AY345479	Bacterium K2-91B
OTU8	NE-366	2	0.934	DQ868668	<i>Rhodovulum</i> sp.	Alphaproteobacteria	0.958	AJ567557	Uncultured alpha <i>Proteobacterium</i> ; MBMPE43
OTU9	NE-150	1	0.962	DQ868668	<i>Rhodovulum</i> sp.	Alphaproteobacteria	0.972	DQ811853	Uncultured alpha <i>Proteobacterium</i> ; MSB-3E4
OTU10	NE-373	1	0.975	AM712634	<i>Andersenella baltica</i> ; type strain: BA141	Alphaproteobacteria	0.998	EF061946	Uncultured alpha <i>Proteobacterium</i> ; XME30
OTU11	NE-53	4	0.97	CP000284	<i>Methylobacillus flagellatus</i> KT	Betaproteobacteria	0.96	AJ582036	Uncultured beta <i>Proteobacterium</i> ; JG36-GS-101
OTU12	NE-302	2	0.941	AB089481	<i>Dexia gummosa</i>	Betaproteobacteria	0.982	AB286331	Uncultured bacterium; 0101
OTU13	NE-303	3				Gammaproteobacteria	0.985	DQ351809	Uncultured gamma <i>Proteobacterium</i> ; Belgica2005/10-ZG-17
OTU14	NE-330	3				Gammaproteobacteria	0.992	EF125435	Uncultured bacterium; MSB-2D6
OTU15	NE-349	2	0.937	EF117913	<i>Thiohalomonas denitrificans</i>	Gammaproteobacteria	0.945	AY225635	Uncultured gamma <i>Proteobacterium</i> ; AT-s80
OTU16	NE-139	1				Gammaproteobacteria	0.979	EF208680	Uncultured bacterium; CI5cm.B02
OTU17	NE-105	1	0.983	AB021367	<i>Marinobacterium stanieri</i> (T); ATCC 27130T	Gammaproteobacteria	0.959	EF190069	Uncultured <i>Marinobacterium</i> sp.; GSX2
OTU18	NE-267	2				Gammaproteobacteria	0.989	EF208690	Uncultured bacterium; CI5cm.F08
OTU19	NE-13	2	0.946	AB053125	<i>Alcanivorax</i> sp.	Gammaproteobacteria	0.936	AJ567576	Uncultured <i>Marinobacter</i> sp.; MBAE14
OTU20	NE-21	1	0.894	AJ237601	<i>Desulfobacterium anilini</i> (T); DSM 4660	Deltaproteobacteria	0.975	DQ463694	Uncultured delta <i>Proteobacterium</i> ; TK-SH10
OTU21	NE-292	4				Deltaproteobacteria	0.984	AJ535245	Uncultured delta <i>Proteobacterium</i>
OTU22	NE-114	1	0.916	EF422413	<i>Desulfurivibrio alkaliphilus</i>	Deltaproteobacteria	0.944	DQ831546	Uncultured delta <i>Proteobacterium</i> ; CBII30
OTU23	NE-43	1	0.871	AY187308	<i>Pelobacter masseliensis</i>	Deltaproteobacteria	0.929	EF999371	Uncultured bacterium; MidBa15

OTUs	Clone name	No. of clones	Cultured Similarity Score	Accession no. of cultured hits	Nearest phylogenetic cultured neighbor	Phylum	Uncultured Similarity Score	Accession no. of uncultured hits	Nearest phylogenetic uncultured neighbor
OTU24	NE-240	1	0.923	X94911	<i>Syntrophobacter</i> sp.	Deltaproteobacteria	0.964	DQ811826	Uncultured delta <i>Proteobacterium</i> ; MSB-5bx5
OTU25	NE-378	3	0.956	AJ620511	<i>Olavius crassitunicatus</i> delta-proteobacterial endosymbiont; d3-P12-1	Deltaproteobacteria	0.991	DQ811824	Uncultured delta <i>Proteobacterium</i> MSB-5B3
OTU26	NE-372	1	0.913	AY493563	Sulfate-reducing bacterium	Deltaproteobacteria	0.923	U81720	Uncultured <i>Eubacterium</i> ; vadinHA60
OTU27	NE-291	1	0.92	CP000482	<i>Pelobacter propionicus</i> DSM 2379	Deltaproteobacteria	0.92	AF529129	Uncultured delta <i>Proteobacterium</i> ; FTLpost101
OTU28	NE-19	1	0.973	AJ271656	<i>Pelobacter</i> sp. A3b3	Deltaproteobacteria	0.974	AJ240980	Uncultured delta <i>Proteobacterium</i> Sva1034
OTU29	NE-364	1				Actinobacteria	0.979	AM259898	Uncultured <i>Actinobacterium</i> ; TAA-10-01
OTU30	NE-251	1				Actinobacteria	0.978	DQ070822	Uncultured <i>Gram-positive bacterium</i> ; JdFBGBact_23
OTU31	NE-82	1				Acidobacteria	0.979	DQ395006	Uncultured <i>Acidobacteria bacterium</i> ; VHS-B4-69
OTU32	NE-37	1	0.905	AJ784892	<i>Haliscomenobacter hydrossis</i> ; DSM 1100	Bacteroidetes	0.916	EF508145	Uncultured <i>Sphingobacterium</i> sp.; MS190-1F
OTU33	NE-376	1				Bacteroidetes	0.939	AJ567581	Uncultured <i>Bacteroidetes bacterium</i> ; MBAE20
OTU34	NE-70	2	0.907	AB331889	<i>Verrucomicrobia bacterium</i> YM27-120	Verrucomicrobia	0.904	AY345492	Unidentified bacterium; W4-B59
OTU35	NE-5	1	0.94	AF165908	<i>Escarpia spicata</i> endosymbiont 'Alvin #2839	Gammaproteobacteria	0.948	AB278144	Uncultured bacterium; Mafs-EB04
OTU36	NE-252	3	0.923	AF328856	<i>Olavius algarvensis</i> sulfur-oxidizing endosymbiont; 1261-9	Gammaproteobacteria	0.938	DQ811837	Uncultured gamma <i>Proteobacterium</i> ; MSB-3A4
OTU37	NE-341	1				Gammaproteobacteria	0.944	DQ351759	Uncultured gamma <i>Proteobacterium</i> ; Belgica2005/10-130-14
OTU38	NE-273	31	0.972	X95460	<i>Methylophaga thalassica</i> ; SM5690	Gammaproteobacteria	0.971	DQ490031	Uncultured bacterium; ODP-46B-02
OTU39	NE-284	9	0.966	DQ660929	<i>Methylophaga</i> sp. DMS044	Gammaproteobacteria	0.965	DQ490031	Uncultured bacterium; ODP-46B-02
OTU40	NE-219	14	0.962	DQ660930	<i>Methylophaga</i> sp. DMS048	Gammaproteobacteria	0.962	DQ490031	Uncultured bacterium; ODP-46B-02
OTU41	NE-140	4	0.988	DQ458821	<i>Marinobacter</i> sp. HS225	Gammaproteobacteria			
OTU42	NE-293	2				Acidobacteria	0.938	EF125465	Uncultured bacterium; MSB-2Y4
OTU43	NE-92	1				Acidobacteria	0.952	AJ241003	Uncultured holophaga/ <i>Acidobacterium</i> Sva0725

OTUs	Clone name	No. of clones	Cultured Similarity Score	Accession no. of cultured hits	Nearest phylogenetic cultured neighbor	Phylum	Uncultured Similarity Score	Accession no. of uncultured Hits	Nearest phylogenetic uncultured neighbor
OTU44	NE-130	1				Acidobacteria	0.952	EF125465	Uncultured bacterium; MSB-2Y4
OTU45	NE-107	1				Acidobacteria	0.965	DQ351815	Uncultured <i>Acidobacteriales</i> bacterium; Belgica2005/10-ZG-24
OTU46	NE-374	1				Acidobacteria	0.968	AB294930	Uncultured <i>Acidobacteria</i> bacterium; pItb-vmat-12
OTU47	NE-75	2				Verrucomicrobia	0.957	AY114336	Uncultured <i>Verrucomicrobia</i> bacterium; LD1-PB9
OTU48	NE-226	28	0.971	X87339	<i>Methylophaga thalassica</i> (T); TCC 33146;	Gammaproteobacteria	0.965	DQ490031	Uncultured bacterium; ODP-46B-02
OTU49	NE-44	1				Deltaproteobacteria	0.938	AY375087	Uncultured bacterium; C10;
OTU50	NE-63	1	0.955	AJ237601	<i>Desulfobacterium anilini</i> (T); DSM 4660	Deltaproteobacteria	0.995	DQ811831	Uncultured delta <i>Proteobacterium</i> ; MSB-5D12;
OTU51	NE-266	2				Deltaproteobacteria	0.986	DQ811801	Uncultured delta <i>Proteobacterium</i> ; MSB-3E9;

OTUs form a separate cluster of α -proteobacteria supported by 100 bootstrap value and grouped with their nearest cultured and uncultured homologs from the database. OTU5 had 94.8% sequence identity to *Mucus bacterium* 32 isolated from the mucus of *Oculina patagonica* (Koren et al. Unpublished, AY654823). OTU6 with 5 clones had 95.4% similarity to *Rhodobacter* sp. 1-5, an isolate from Arctic sea ice, Spitzbergen. OTU7 with 2 clones had 95.8% sequence identity to *Rhodovulum marinum* strain JA242, isolated from different habitats of India [31], it contains bacteriochlorophyll-a and carotenoids in vesicular intracytoplasmic membranes. OTU8 and OTU9 with 2 and 1 clone had 93.4 and 96.2% sequence identity respectively to *Rhodovulum* sp. SMB1 (Choong et al. unpublished, DQ868668). OTU10 had high 97.5% sequence identity to *Andersenella baltica* type strain BA141T, isolated from sediment in the central Baltic Sea characterized by the presence of carotenoids and absence of bacteriochlorophyll a [32].

β -proteobacteria

β -proteobacteria had constituted only 1 OTU from L-I, which had not shown any similarity to cultured sequences. Phylogenetically it forms sister clade with γ -proteobacteria with 97 bootstrap value where as L-II is represented with 2 OTUs (No. 11, 12). OTU11 of L-II with 4 clones had high sequence identity of 97% to *Methylobacillus flagellatus* KT (Copeland et al. unpublished, CP000284). OTU12 of L-II

with 2 clones had 94.1% sequence identity to *Derxia gummosa* strain: IAM14990. *Derxia* is a nitrogen-fixing genus identified with a bacterium isolated from West Bengal soil of India [33].

δ -proteobacteria

δ -proteobacteria with 3 OTUs (No. 3, 4 and 20) (5.88%) containing 6 clones had constituted 4% part of L-I, OTU20 with 3 clones had 89.7% sequence identity to *Myxobacterium* AT1-01 an isolate from hot springs in Japan [34]. δ -proteobacteria group of L-II had 12 OTUs (No. 20, 21, 22, 23, 24, 25, 26, 27, 28, 49, 50 and 51) (23.52%) with 18 clones. OTU20 and OTU50 had 89.4 and 95.5% sequence identity to *Desulfobacterium anilini* strain DSM 4660, sulfate-reducing bacteria capable of aniline and dihydroxybenzenes degradation isolated from marine sediment [35, 36]. OTU22 had 91.6% sequence identity to *Desulfurivibrio alkaliphilus* strain AHT2, halophilic bacteria of reductive sulfur cycle from Soda Lake (Sorokin et al. unpublished, EF422413). OTU24 had 92.3% sequence identity to *Syntrophobacter* sp. which is a syntrophic propionate-oxidizing, sulfate-reducing bacterium from a fluidized bed reactor [37]. OTU25 with 3 clones had 95.6% identity to *Olavius crassitunicatus* clone d3-P12-1, an endosymbiont of a gutless worm (Oligochaeta) from the Peru margin, capable of sulfide oxidation [38]. OTU26 had 91.3% sequence identity to sulfate-reducing bacterium PF2802 capable of

Fig. 3 Unrooted phylogenetic dendrogram for Library I (L-I) dataset was drawn using bayesian inference method with program MrBayes 3.0. The analysis for 16S rRNA gene consisted of 3,000,000 generations. An appropriate model of sequence evolution for data set was chosen via Akaike Information Criterion (AIC) using program MrModeltest 2.2 and the model selected was (GTR+I+G). Trees were sampled for every 100 generations. First 3000 trees (10%) were discarded as burnin. Bayesian posterior probabilities were calculated using 50% majority rule consensus and three independent runs were performed. Text in parentheses indicate the representative clone name.

Fig. 4 Unrooted phylogenetic dendrogram for Library II (L-II) dataset was drawn using bayesian inference method with program MrBayes 3.0. The analysis for 16S rRNA gene consisted of 3,000,000 generations. An appropriate model of sequence evolution for data set was chosen via Akaike Information Criterion (AIC) using program MrModeltest 2.2 and the model selected was (GTR+I+G). Trees were sampled for every 100 generations. First 3000 trees (10%) were discarded as burnin. Bayesian posterior probabilities were calculated using 50% majority rule consensus and three independent runs were performed. Text in parentheses indicate the representative clone name.

degrading n-alkene (Cravo et al. unpublished, AY493563). OTU27 and OTU28 had 92% and 97.3% sequence identity to *Pelobacter propionicus* DSM 2379 and *Pelobacter* sp. clone A3b3 respectively, both found in marine sediments and capable of sulphate reduction (unpublished, CP000482 and [39]). Phylogenetically, all these OTUs clustered with the δ -proteobacteria clade.

γ -proteobacteria

Majority of OTUs (29.4%) belonging to γ -proteobacteria in L-I had shown affiliation to cultured isolates but with very less similarity. Phylogenetically all OTUs clustered with γ -proteobacteria clade showing 100 bootstrap confidence value. OTU16 with 5 clones had 91.2% sequence identity to *Alkalilimnicola ehrlichei* MLHE-1, an anaerobic, facultatively autotrophic arsenite oxidizing bacterium that respire nitrate or nitrite (Copeland et al. unpublished, CP000453 and [40]). OTU17 of L-I had 89.6% sequence identity to *Methylobacter luteus* NCIMB, a methanotrophic bacteria with an ability to consume nitric oxide and produce small amounts of nitrous oxide [41]. OTU18 of L-I had 88.3% sequence identity to *Ectothiorhodospinus mongolicum* strain M9, a purple sulfur bacterium isolated from a Mongolian soda lake having photosynthetic pigments such as bacteriochlorophyll a and carotenoids of the spirilloxanthin series (Gorlenko et al. unpublished, AY298904). OTU40 of L-I with 46 clones and OTU41 of L-I with 10 clones were among the most abundant group, they had high 97.6 and 96.4% sequence identity respectively to *Methylophaga* sp. DMS010, an isolate from marine dimethylsulfide-degrading enrichment. It carries out oxidation of DMS by a route different from those described for *Thiobacillus* species [42]. OTU43 of L-I had very high 99.9% sequence identity to *Aeromonas hydrophila* (ATCC 7966T) which cause infections in invertebrate and vertebrate such as frogs, birds and domestic animals, also an emerging human pathogen irrespective of the host's immune system [43]. OTU48 of L-I had high 98.3% sequence identity to *Methylophaga thalassica* (T); ATCC 33146 it also had an ability to degrade dimethylsulfide. OTU49 and OTU50 of L-I with clones 1 and 12 had high sequence identity of 95.6 to 99% to *Alcanivorax* sp. Mho1 and *Alcanivorax* sp. K3-3 respectively, these are alkane-degrading marine bacterium predominated in petroleum or hydrocarbon contaminated seawater or sediments (Kishira et al. unpublished, AB053124 and [44]). Phylogentic clustering of these OTUs with respective cultured homologs is supported by higher node value. γ -proteobacteria was also the most abundant taxon in Library-II. Most of the OTUs (33.34%) had shown affiliation to cultured representatives. OTU1 had high sequence

identity of 96% to sulfur-oxidizing bacterium NDII.1 an isolate from a shallow water hydrothermal vent (Sievert et al. unpublished, AF17042). 5 OTUs (No. 4, 38, 39, 40 and 48) of L-II with clones 2, 31, 9, 14 and 28 had high 88.4 to 98.8% sequence identity to *Methylophaga* sp. and *Methylophaga thalassica* clustered with this organism with 100 bootstrap value. *Methylophaga* sp. was the most abundant homolog in both the libraries. OTU15 with 2 clones had 93.7% sequence identity to *Thiohalomonas denitrificans* strain HLD 14, a thiodenitrifying halophilic bacteria isolated from sediments of a solar saltern (Tourova et al. unpublished, EF117913). OTU19 of L-II with 2 clones had 94.6% sequence identity to *Alcanivorax* sp. I4 (Kishira et al. unpublished, AB053125), L-I had a rich representation of these sequences with 13 clones. OTU35 of L-II had sequence identity of 94% to *Escarpia spicata* endosymbiont Alvin #2839 an isolate from vestimentiferan tubeworms [45]. OTU36 of L-II with 3 clones had 92.3% identity to *Olavius algarvensis*, a sulfur-oxidizing endosymbiont from an oligochaete worm [46]. OTU41 of L-II with 4 clones had high 98.8% identity to *Marinobacter* sp. HS225 strain HS225, a halophile isolated from the Zhoushan Archipelago of China (Xu et al. unpublished, DQ458821).

Acidobacteria

5 OTUs (No. 1, 6, 8, 45 and 46) (9.8%) from L-I showed close relationships to each other and formed one large cluster having three sister clade with 99 bootstrap confidence value. The closest species from database to the above group are uncultivated bacteria from oil polluted soil of Romania, stromatolites of Hamelin pool in Shark Bay, western Australia, near shore and outer shelf reefs of Australia, Fushan forest soils of Taiwan and hydrocarbon seep sediments. Similarly 6 OTUs (No. 31, 42, 43, 44, 45 and 46) (11.76%) of L-II formed a cluster and were closely related to uncultured *Holophaga/Acidobacteria* from Great Barrier Reef calcareous sediments of Australia, mangrove soil, permeable shelf sediments, organically-enriched fish farm sediments and hydrothermal sediments.

Actinobacteria

5 OTUs (No. 23, 24, 25, 26 and 27) (9.8%) were clustered to *Actinobacteria* from L-I. They formed a separate cluster with 98 bootstrap value and showed homology to a number of uncultured *Actinobacteria* from a variety of places like perennial ice cover of Lake Vida, Antarctica, Nullabar caves, Oregon caves, Hypersaline Gulf of Mexico sediments and deep sea. Only 2 OTUs (No. 29, 30) (3.92%) from L-II were included in *Actinobacteria* and both were closely related to uncultured *Actinobacteria* with 100% node value.

Firmicutes

Firmicutes forms a separate cluster, constituted 3 OTUs (5.88%) with 100 node value from L-I. OTU22 containing 2 clones, had very high 99.5% sequence identity to *Caryophanon tenue* type strain DSM 14152T (Fritze, D. unpublished, AJ491303). OTU44 of L-I had high 99.3% sequence identity to *Bacillus niacini* which produce an ofloxacin ester-enantioselective esterase [47].

Bacteroidetes

Bacteroidetes had almost equal representation among both the libraries. *Bacteroidetes* in L-I is presented by OTU36 with 2 clones grouped with *Mucus bacterium* 86 having 92% sequence identity, it is isolated from the mucus of *Oculina patagonica* a scleractinian coral (Koren et al. Unpublished, AY654823). OTU32 and OTU33 of L-II had shown clustering with *Haliscomenobacter hydrossis* strain DSM 1100, isolated from the pelagic zones of a broad spectrum of freshwater habitats. *H. hydrossis* has one of the smallest widths of any of the filamentous bacteria, 0.5 μm , and filament usually extending to a length of 20–100 μm [6].

Verrucomicrobia

Verrucomicrobia sequence from L-I does not show similarity to any cultured bacteria, and it form a sister clade with the *Firmicutes* with 100 bootstrap value. *Verrucomicrobia* is represented in L-II with 2 OTUs and 4 clones, constituting 2.46% of total clones. OTU34 and OTU47 of L-II had 90.7% similarity to cultured *Verrucomicrobia bacterium* YM27-120 (Matsuo et al. unpublished, AB 331889). These OTUs form a sister clade with *Cyanobacteria*.

Other minor phylogenetic groups

These groups had a minor contribution to the bacterial diversity studied. It includes *Chloroflexi*, *Cyanobacteria*, *Gemmatimonadetes*, Candidate division TM7 and unclassified bacteria. OTUs belonging to *Gemmatimonadetes* and unclassified bacteria from L-I form a cluster with 100 bootstrap confidence value, suggesting that this unclassified OTU may belongs to *Gemmatimonadetes* taxon. Phylogenetically Candidate division TM7 clustere with *Verrucomicrobia* supported by lower bootstrap value. Candidate division TM7 are named after sequences obtained in an environmental study of a peat bog [21], later on partial-length-sequence representatives of this Candidate division were subsequently identified from activated sludges and soil. We found OTU3 clustering with *Chloroflexi* and OTU2 with *Cyanobacteria* in L-II. OTU2 showed clustering to *Odontella sinensis* with 93 bootstrap value, and were

derived from the chloroplast of an alga [48] and were most closely related to the *Cyanobacteria*.

Statistical analyses: species richness

Statistical analyses of biodiversity provide interesting insights (Table 4). 29 singleton OTUs (OTUs having single clone) with Good's coverage 80.41% in L-I and 27 singletons OTUs with Good's coverage 83.34% in L-II suggested that the coverage was fine but it also indicated that any new clone had 19.59% and 16.66% of chance to fall in an unknown species from respective libraries. The higher value of Shannon index (3.1679) in L-II suggest higher diversity, richness and even distribution of abundance, indicating that species are more diverse and evenly distributed in L-II as compared to L-I. The value of Simpson index for library I (0.1100) suggest that probability of finding a new phylo-type is more in L-I than L-II (0.0784). Rarefaction analysis shows a curvilinear plot but it did not plateau with the current coverage. The curve (Fig. 5) indicates that the diversity was sampled with good level of confidence and majority of OTUs in the sample were detected but still there is a need of more comprehensive sampling to cover rest of the undiscovered diversity. Significant $P = 0.05$ decrease in the rate of OTU detection with increasing number of clones examined was observed in rarefaction curves of both libraries. Chao values suggest undiscovered OTUs/phylotypes in the library which could be still explored by more robust method. β -LIBSHUFF with 10,000 randomization found P value (0.001) for comparison, which provided strong evidence that none of the library is a subset of the other and two libraries are considered significantly different communities.

Discussion

Study of marine microbial biodiversity is of vital importance to the understanding of the different processes of the

Table 4 Statistical Indices calculated for two libraries

	Library I	Library II
No. of sequences	148	162
No. of OTUs*	51	51
Singletons	29	27
Shannon Index	3.0918 ± 0.2479	3.1679 ± 0.2123
Simpson Index	0.110039	0.07836
Chao1	87.9091	78
Good's coverage	80.41%	83.34%

* 97% similarity clusters have been considered as operational taxonomic units (OTUs)

Fig. 5 Comparative rarefaction curve for Library I (L-I) and Library II (L-II)

ocean, which may present potent novel microorganisms for screening of bioactive compounds [49]. Phylogenetic characterization of bacteria in tropical marine sediments is a crucial step in our understanding of bacteria in such biomes and to realize the naturally occurring microbial communities. The 16S rDNA sequences obtained in this study shows affiliation to diverse taxons of bacteria, even taxons like *Actinobacteria* and *Firmicutes* that were difficult to lyse were well represented, showing that the DNA was efficiently extracted [50]. A number of studies have emphasized the drawbacks of an increment of PCR cycles during amplification of the 16S rRNA gene as it could bias the incurred clone composition [51]. To exclude this bias we used 25 cycles in the PCR step. Copy number of the 16S rRNA genes [52] and differential PCR amplification efficacy of DNA from heterogeneous templates [53, 54] may also introduce biases in the datasets which can lead to the misinterpretation of the data, as the proportions found in the clone libraries do not always represent the 16S rDNA proportions found in the original samples. However the uncultured phylogenetic studies provide a far less biased picture of community composition [5] than would any single cultivation technique.

We carried out this work to shed light on the bacterial diversity from a tropical coastal habitat in Balramgari, north-east coast of India and to ascertain that microscopic and submicroscopic facts like the presence of gels, particulate matter, pH and other physicochemical parameters affect the bacterial diversity. For this we constructed two 16S rDNA clone libraries from subsurface sediments which lie 100 meters apart. Subset I samples used in constructing Library I had higher content of organic carbon, silt and clay but lower salinity and sand percentage, than subset II (Library

II). The second major difference between two sites selected was the presence and absence of horseshoe crabs. Site I inhabits horseshoe crabs whereas site II was selected in such a way that it completely lacks the crabs. Even though it can not be proved with the present data but we believe that this could be one of the biological factor for different microbial diversity in our study.

All the clones clustered in the class of *Actinobacteria*, *Acidobacteria*, *Bacteroidetes*, *Gemmatimonadetes*, candidate division TM7 and unclassified bacteria showed homology and clustered with only uncultured bacterium of soil and sediments with sequence identity between 89 and 98% and did not match with any closest cultured relative. The homology with uncultured bacteria could be due to the poor representation of cultured bacteria belonging to these classes in databases. Comparative 16S rDNA analysis showed that none of the cloned sequences are identical to any known 16S rDNA sequences in the RDP database, but revealed substantial phylogenetic diversity with little representation among cultured organisms previously studied.

We acquired only one OTU which was monophyletic with cyanobacteria from L-II; it is in accordance to the study which says cyanobacterial population was very poor on sandy shores due to, rough tides, absence of substratum and poor nutrient. Striking observation was presence of pathogenic bacteria such as *A. hydrophila*, *O. intermedium* and *Pseudoalteromonas* and few endosymbionts such as *Mesorhizobium* in L-I, which were totally absent from L-II. The presence of *A. hydrophila* and *O. intermedium* was further confirmed by Real-time PCR using genus specific primers. It was observed that the pathogenic bacteria were identified in the sediment where crabs were present. Therefore there could be host-parasite association between these

bacteria and crab but it can not be confirmed without further studies.

A number of sequences had shown affiliation to sulphate reducing and oxidizing bacteria along with dimethylsulfide-degrading bacteria, these bacteria contribute over 50% of the carbon turnover of coastal marine sediments and take part in the cycling of sulphur compounds in sea water [45]. Few sequences were affiliated to nitrifying bacteria, which oxidize either ammonia to nitrite or nitrite to nitrate and covert nitrogen to a form readily available for other biological processes. This is an extremely important process, since positively charged ammonium ions bind to acidic sediment particles, where they become available for biological processes, more abundant in near shore water than in off shore regions [45]. 62.75% of OTUs (32 OTUs) from L-I and 66.67% of OTUs (34 OTUs) from L-II shown $\geq 87\%$ and $< 97\%$ similarity to database sequences suggesting the putative presence of new genera and species. In few cases where identity was less than or equal to 85%, the rDNA could represent new taxa that may be specific to the sampling site. Thus it is apparent that the diversity of microorganisms in the sediment is extensive and that the phylogenies of many dominant sediment bacteria remain uncharacterized.

In conclusion this work elucidates the existence of complex and unique bacterial profile associated with tropical coastal habitat in Balramgari, north-east coast of India. Detected organisms probably represent novel species or genera, previously not discovered from any environment. It also demonstrates the presence of a myriad number of forces and factors in the coastal habitat which governs the bacterial diversity including physical, chemical and biological factors. Further phylogenetic and functional analyses are required for understanding the microbial diversity and community structure and their interaction with air-water, water-sediment and host microorganism-sediment interfaces. This work provides a framework for further studies of these evidently important habitats.

Acknowledgments This study was supported by a grant from Department of Biotechnology, India. Research fellowships awarded by Council of Scientific and Industrial Research, New Delhi, to A. K. Gupta and P. Verma is gratefully acknowledged. The authors thank Dr. G. Rastogi (South Dakota School of Mines and Technology, USA) and D. Dhotre (National Centre for Cell Science, India) for their helpful suggestions in phylogenetic and statistical analysis.

References

1. Delong EF (1992) Ecology Archaea in coastal marine environments. *Proc Natl Acad Sci USA* 89:5685–5689
2. Hugenholtz P, Goebel BM and Pace NR (1998) Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J Bacteriol* 180:4765–4774
3. Pace NR (1997) A molecular view of microbial diversity and the biosphere. *Science* 276:734–740
4. Qasim SZ (1999) *The Indian Ocean: Images and Realities*, Oxford and IBH, New Delhi. 57–90
5. Amann RI, Ludwig W and Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev* 59:143–169
6. Schauer M and Hahn MW (2005) Diversity and phylogenetic affiliations of morphologically conspicuous large filamentous bacteria occurring in the pelagic zones of a broad spectrum of freshwater habitats. *Appl Environ Microbiol* 71(4):1931–1940
7. Chatterji A (1994) *The Indian Horseshoe Crab – A Living Fossil*. A Project Swarajya Publication. p. 157
8. Chauhan OS (1992) Laminar and grain size measures in beach sediments, east coast beaches. *Ind J Coast Res* 8: 172–182
9. Wakeel SK and Riley JP (1956) The determination of organic carbon in marine muds. *Journal de Conseil permanent international l'exploration de la mer* 22:180–193
10. Weisburg WG, Barns SM, Pelletier DA & Lane DJ (1991) 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* 173:697–703
11. Ashelford KE, Chuzhanova NA, Fry JC, Jones AJ and Weightman AJ (2006) New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras. *Appl Environ Microbiol* 72:5734–5741
12. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search program. *Nucleic Acids Res* 25:3389–3402
13. Thompson JD, Higgins DG & Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acid Res* 22:4673–4680
14. Xie CH & Yokota A (2004) Phylogenetic analyses of the nitrogen-fixing genus *Derrxia*. *J Gen Appl Microbiol* 50: 129–135
15. Felsenstein J (1993) PHYLIP (phylogenetic inference package) version 3.61 University of Washington, Seattle, USA
16. Schloss PD and Handelsman J (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol* 71:1501–1506
17. Stackebrandt E and Goebel BM (1994) Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Bacteriol* 44:846–849
18. Huelsenbeck JP and Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*. 17: 754–755
19. Good IJ (1953) The population frequencies of species and the estimation of the population parameters. *Biometrika* 40: 237–264
20. Hurlbert SH (1971) The non-concept of species diversity: a critique and alternative parameters. *Ecology* 52:577–586

21. Schloss PD, Larget BR and Handelsman J (2004) Integration of microbial ecology and statistics: a test to compare gene libraries. *Appl Environ Microbiol* 70:5485–5492
22. Buckley DH, Graber JR and Schmidt TM (1998) Phylogenetic analysis of nonthermophilic members of the kingdom Crenarchaeota and their diversity and abundance in soils. *Appl Environ Microbiol* 64:4333–4339
23. Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B and Shao Y (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* 277:1453–1474
24. Wang GCY and Wang Y (1997) Frequency of formation of chimeric molecules is a consequence of PCR coamplification of 16S rRNA genes from mixed bacterial genomes. *Appl Environ Microbiol* 63:4645–4650
25. Green DH, Llewellyn LE, Negri AP, Blackburn SI and Bolch CJ (2004) Phylogenetic and functional diversity of the cultivable bacterial community associated with the paralytic shellfish poisoning dinoflagellate *Gymnodinium catenatum*. *FEMS Microbiol Ecol* 47:345–357
26. Yurkov V, Jappe J and Vermeglio A (1996) Tellurite resistance and reduction by obligately aerobic photosynthetic bacteria. *Appl Environ Microbiol* 62:4195–4198
27. Apisarnthanarak A, Kiratisin P and Mundy LM (2005) Evaluation of *Ochrobactrum intermedium* bacteremia in a patient with bladder cancer. *Diagn Microbiol Infect Dis* 53: 153–155
28. Urbance JW, Bratina BJ, Stoddard SF and Schmidt TM (2001) Taxonomic characterization of *Ketogulonigenium vulgare* gen. nov., sp. nov. and *Ketogulonigenium robustum* sp. nov., which oxidize L-sorbose to 2-keto-L-gulonic acid. *Int J Syst Evol Microbiol* 51:1059–1070
29. Maturano L, Santos F, Rossello-Mora R and Anton J (2006) Microbial diversity in maras salterns, a hypersaline environment in the peruvian andes. *Appl Environ Microbiol* 72: 3887–3895
30. Wang DL, Xiao M, Qian W and Han B (2007) Screening and identification of a photosynthetic bacterium reducing selenite to red elemental selenium. *Wei Sheng Wu Xue Bao* 47:44–47
31. Srinivas TNR, Anil-Kumar P, Sasikala C, Ramana CV, Sül-ling J and Imhoff JF (2006) *Rhodovulum marinum* sp. nov., a novel phototrophic purple non-sulfur alphaproteobacterium from marine tides of Visakhapatnam. *India Int J Syst Evol Microbiol* 56:1651–1656
32. Brettar I, Christen R, Botel J, Lunsdorf H and Hofle MG (2007) *Andersenella baltica* gen. nov., sp. nov., a novel marine bacterium of the Alphaproteobacteria isolated from sediment in the central Baltic Sea. *Int J Syst Evol Microbiol* 57:2399–2405
33. Xie CH and Yokota A (2004) Phylogenetic analyses of the nitrogen-fixing genus *Dexia*. *J Gen Appl Microbiol* 50: 129–135
34. Iizuka T, Tokura M, Jojima Y, Hiraishi A, Yamanaka S and Fudou R (2006) Enrichment and phylogenetic analysis of moderately thermophilic myxobacteria from hot springs in Japan. *Microbes Environ* 21:189–199
35. Hippe H, Vainshtein M, Gogotova GI and Stackebrandt E (2003) Reclassification of *Desulfobacterium macestii* as *Desulfomicrobium macestii* comb. Nov *Int J Syst Evol Microbiol* 53:1127–1130
36. Schnell S, Bak F and Pfennig N (1989) Anaerobic degradation of aniline and dihydroxybenzenes by newly isolated sulfate-reducing bacteria and description of *Desulfobacterium aniline*. *Arch Microbiol* 152:556–563
37. Ellner G, Busman A, Rainey FA and Diekmann H (1996) A syntrophic propionate-oxidizing, sulfate-reducing bacterium from a fluidized bed reactor. *Syst Appl Microbiol* 19:414–420
38. Blazejak A, Erseus C, Amann R and Dubilier N (2005) Coexistence of bacterial sulfide oxidizers, sulfate reducers, and spirochetes in a gutless worm (*Oligochaeta*) from the Peru margin. *Appl Environ Microbiol* 71:1553–1561
39. Thamdrup B, Rossello-Mora R and Amann R (2000) Microbial manganese and sulfate reduction in Black Sea shelf sediments. *Appl Environ Microbiol* 66:2888–2897
40. Oremland RS, Hoefl SE, Santini JM, Bano N, Hollibaugh RA and Hollibaugh JT (2002) Anaerobic oxidation of arsenite in monolake water by a facultative arsenite-oxidizing chemoautotroph, strain MLHE-1. *Appl Environ Microbiol* 68:4795–4802
41. Ulledge J, Ahmad A, Steudler PA, Pomerantz WJ and Cavanaugh CM (2001) Family- and genus-level 16S rRNA-targeted oligonucleotide probes for ecological studies of methanotrophic bacteria. *Appl Environ Microbiol* 67:4726–4733
42. Schäfer H (2007) Isolation of *Methylophaga* spp. from Marine dimethylsulfide-degrading enrichment cultures and identification of polypeptides induced during growth on dimethylsulfide. *Appl Environ Microbiol* 73:2580–2591
43. Ruimy R, Breittmayer V, Elbaze P, Lafay B, Boussemart O, Gauthier M and Christen R (1994) Phylogenetic analysis and assessment of the genera *Vibrio*, *Photobacterium*, *Aeromonas*, and *Plesiomonas* deduced from small-subunit rRNA sequences. *Int J Syst Bacteriol* 44:416–426
44. Syutsubo K, Kishira H and Harayama S (2001) Development of specific oligonucleotide probes for the identification and in situ detection of hydrocarbon-degrading *Alcanivorax* strains. *Environ Microbiol* 3:371–379
45. Di Meo CA, Wilbur AE, Holben WE, Feldman RA, Vrijenhoek RC and Cary SC (2000) Genetic variation among endosymbionts of widely distributed vestimentiferan tubeworms. *Appl Environ Microbiol* 66:651–658
46. Dubilier N, Mulders C, Ferdelman T, de Beer D, Pernthaler A, Klein M, Wagner M, Erseus C, Thiermann F, Krieger J, Giere O and Amann R (2001) Endosymbiotic sulphate-reducing and sulphide-oxidizing bacteria in an oligochaete worm. *Nature* 411:298–302
47. Goto K, Omura T, Hara Y and Sadaie Y (2000) Application of the partial 16S rDNA sequence as an index for rapid identification of species in the genus *Bacillus*. *J Gen Appl Microbiol* 46:1–8
48. Kowallik KV, Stoebe B, Schaffran I, Kroth-Pancic P and Freier U (1995) The chloroplast genome of a chlorophyll a+c- containing alga, *Odontella sinensis*. *Plant Mol Biol Rep* 13:336–342
49. Das S, Lyla PS and Khan SA (2006) Marine microbial diversity and ecology: importance and future perspectives. *Cur Science* 90:1325–1335
50. More MI, Herrick JB, Silva MC, Ghiorse WC and Madsen EL (1994) Quantitative cell lysis of indigenous microorganisms

- and rapid extraction of microbial DNA from sediment. *Appl Environ Microbiol* 60:1572–1580
51. Wilson KH and Blichington RB (1996) Human colonic biota studied by ribosomal DNA sequence analysis. *Appl Environ Microbiol* 62:2273–2278
52. Kerkhof L and Speck M (1997) Ribosomal RNA gene dosage in marine bacteria. *Mol Mar Biol Biotechnol* 6: 260–267
53. Chandler DP, Fredrickson JK and Brockman FJ (1997) Effect of PCR template concentration on the composition and distribution of total community 16S rDNA clone libraries. *Mol Ecol* 6:475–482
54. Farrelly V, Rainey FA and Stackebrandt E (1995) Effect of genome size and *rrn* gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. *Appl Environ Microbiol* 61:2798–2801