Microbial diversity analysis in the oxygen minimum zones of the Arabian Sea using metagenomics approach

Mandar S. Paingankar[1,2], Kedar Ahire[1], Pawan Mishra[1], Shriram Rajpathak[1] and Deepti D. Deobagkar[1,3,]*

[1]Molecular Biology Research Laboratory, Centre for Advanced Studies, Department of Zoology, Savitribai Phule Pune University, Pune 411 007, India
[2]Present address: Department of Zoology, Government Science College, Chamorshi Road, Gadchiroli 442 605, India
[3]ISRO Chair Professor, ISRO Space and Technology Cell, Savitribai Phule Pune University, Pune 411 007, India

Large oxygen-depleted areas known as oxygen minimum zones (OMZs) have been reported from the Arabian Sea, and recent reports indicate that these areas are expanding at an alarming rate. In marine waters, oxygen depletion may also be related to global warming and temperature rise. The acidification and deoxygenation due to OMZs can lead to major consequences wherein the plants, fish and other biota will struggle to survive in the ecosystem. The present study has identified the microbial community structure using next generation sequencing-based metagenomics analysis in water samples collected at different depths from the OMZs and non-OMZs of the Arabian Sea. Environmental variables such as depth, site of collection and oxygen concentration might influence species richness and evenness among microbial communities in these locations. Our observations suggest that population dynamics of microbes consisting of nitrate reducers accompanied by sulphate reducers and sulphur oxidizers influences the interconnected geochemical cycles of OMZs. In addition to providing baseline data related to the diversity and microbial community dynamics in waters in the OMZs; such analysis can provide insight into processes regulating productivity and ecological community structure of the ocean.

Keywords: Bacterial diversity, metagenomics, microbial communities, oxygen minimum zones.

THE oxygen minimum zones (OMZs) in the Arabian Sea are the second-most intense areas amongst tropical oceans in the world[1,2] with a near-total depletion of oxygen at depths from 200 to 1000 m (ref. 3). In these locations, suboxic levels ($\Box 5$ μmol O_2/kg) of oxygen are present over vast areas at different depths and denitrification occurs in the upper portion[4]. Geochemical observations indicate that oxygen minimum zones have expanded over the past decades[5], and they could expand further in response to ocean warming and increased stratification

associated with climate change[6,7]. It has been suggested that the biological consumption of oxygen is most intense below the region of highest productivity in the western Arabian Sea[8–10]. The total volume of OMZs in the ocean is growing at an alarming rate, their upper boundaries are vertically shoaling, and the degree of anoxia is intensifying within the cores of the OMZs[5,11]. The expansion of OMZs in the Arabian Sea has become a major concern because of its impact on the marine ecosystem. The expansion of OMZs due to climate change and its impact on the ecosystem and the atmosphere is multi-dimensional and requires intense studies.

An OMZ is characterized by high nitrite accumulation and very low or undetectable oxygen concentration[12]. Nitrous oxide (N_2O) concentration in OMZs has been reported to vary inversely with nitrite concentration[13]. Often as the oxygen levels diminish, the ecosystem cannot sustain normal biotic inhabitants and macrofauna. As a result, OMZs are often associated with coastal and equatorial upwelling regions, and the increased primary production rates determine the high levels of altered microbial metabolism[11,14]. Importantly, nitrogen (N) cycling plays a crucial role in nitrate reduction to N_2 (denitrification), and anaerobic ammonia oxidation (anammox) along with nitrate reduction to ammonia[15]. Moreover, nitrification has been shown to be an important source of oxidized N at the OMZ boundaries[16–18].

Interestingly, various metagenomic studies on OMZs have revealed that complex communities (such as nitrifiers) play an important role in N cycle in the OMZs[18]. Members of the Planctomycetes, Thaumarchaeota and Nitrospinae phyla have been observed to perform majority of anammox, ammonia oxidation and nitrite oxidation, and play an important role in OMZ dynamics[12,18–24]. Although some reports are available, denitrification[25–27] and heterotrophic denitrification via a complete sequential reduction of nitrate (NO_3) to N_2 has not been fully explored in the OMZs[28,29]. A few studies have been carried out to understand the microbial diversity in OMZs of the Arabian Sea[30–34]. The special growth requirements of

Table 1. Information on sampling sites

Place	Code	Latitude	Longitude	Depth (m)	Date	Time	Oxygen (μmol kg^{-1})	OMZ/non-OMZ
Goa	GAS1	15.755	72.819	1001	22 May 2015	05.30 am	15.75	OMZ
Goa	GAS2	15.755	72.819	503	22 May 2015	05.30 am	1.94	OMZ
Goa	GAS3	15.707	73.009	205	22 May 2015	10.45 am	3.48	OMZ
Goa	GAS4	15.616	73.228	93	22 May 2015	03.45 pm	73.26	Non-OMZ
Mangaluru	MGS5	12.826	74.339	102	24 May 2015	10.00 am	30.6	Non-OMZ
Mangaluru	MGS6	13.058	74.071	176	24 May 2015	01.30 pm	6.32	OMZ
Mangaluru	MGS7	13.005	74.012	983	24 May 2015	06.15 pm	21.39	OMZ
Mangaluru	MGS8	13.005	74.012	200	24 May 2015	06.15 pm	6.21	OMZ
Calicut	CLS9	11.377	74.712	1000	25 May 2015	06.30 am	24.82	Non-OMZ
Calicut	CLS10	11.621	74.903	189	25 May 2015	11.25 am	12.55	OMZ
Calicut	CLS11	11.370	75.128	96	25 May 2015	03.00 pm	23.08	Non-OMZ

Figure 1. Location of sampling sites in the Arabian Sea. (Distance from the coast is provided in the Supplementary Table 1).

these microbes and abundance of uncultured organisms (over 99%) make next generation sequencing (NGS)-based metagenomics, the method of choice in order to unravel the complexities of microbial communities, their dynamics and ecological significance.

In the present study, water samples collected from different depths of the Arabian Sea (100 to 1000 m across the transect) from Goa, Mangaluru and Calicut (Table 1) were processed for high-throughput NGS-based metagenomics (based on 16S rRNA gene sequencing). The microbial diversity and predicted metabolic activities associated with these microbial communities in the OMZs and non-OMZs of the Arabian Sea provide valuable insight into the nature of biogeochemical processes.

Material and methods

Sample collection and processing

Water samples at different depths were collected during the *Sagar Sampada* cruise (Cruise Number 340, 16 May to 8 June 2015) from Goa (GAS1, GAS2, GAS3 and GAS4; distance from the coast ranging from 51 to

90 km), Mangaluru (MGS5, MGS6, MGS7 and MGS8; distance from the coast ranging from 52 to 84 km) and Calicut (CLS9, CLS10 and CLS11; distance from the coast ranging from 66 to 109 km (Table 1 and Figure 1)). A conductivity–temperature–depth (CTD) system equipped with attached oxygen and turbidity measurement sensors was deployed to record the physical properties of water (Table 2) and the samples were grouped into OMZ and non-OMZ.

Using the on-board system of the *Sagar Sampada*, 2000 ml water samples were collected from each sampling site in screw-top sterile coliform water-sample bottles. The samples were then filtered through 0.22 μm filter (Millex, Merck Millipore, USA) in biosafety cabinet (Thermo Fisher, USA) and filter paper was used for DNA extraction.

DNA extraction

Water (1000 ml) was collected from each sampling site and organisms collected by filtering water through 0.22 μm filter (Millex, Merck Millipore, USA) were utilized for DNA isolation using Power water DNA isolation kit (MoBio Laboratories Inc., Carlsbad, CA, USA). Isolation was carried out on the ship to avoid degradation of DNA. Next, DNA concentration was measured using the Quantus fluorimeter (Promega, USA).

Amplification primers and sequence analysis

16s rRNA (corresponding to V3 and V4 regions) was amplified from total genomic DNA isolated (16S amplicon PCR forward primer 5′TCGTCGGCAGCGTCAGATGTGT-ATAAGAGACAGCCTACGGGNGGCWGCAG3′; 16S amplicon PCR reverse primer 5′GTCTCGTGGGCTCGGA-GATGTGTATAAGAGACAGGACTACHVGGGTATCTA-ATCC3′) with appropriate sample bar-coding index sequences and Illumina adapters. AMpure XP beads were employed to remove unused primers and other unwanted nucleic acid fragments, and the purified PCR amplicons

Table 2. Physico-chemical parameters of samples collected from the Arabian Sea

Place	Code	Depth (m)	Oxygen (μmol kg^{-1})	Temperature (°C)	Conductivity (S/m)	Turbidity ECO (NTU)	Salinity	Density (σ–θ, kg/m^3)
Goa	GAS1	1001	15.75	8.12	3.69	0.2858	35.23	27.45
Goa	GAS2	503	1.94	11.86	4.06	0.2611	35.54	27.03
Goa	GAS3	205	3.48	14.86	4.32	0.2877	35.36	26.28
Goa	GAS4	93	73.26	23.89	5.35	0.2728	36.22	24.6
Mangaluru	MGS5	102	30.6	21.54	5.02	0.7585	35.56	24.77
Mangaluru	MGS6	176	6.32	15.16	4.34	0.2435	35.23	26.12
Mangaluru	MGS7	983	21.39	14.33	4.26	0.2489	35.24	26.31
Mangaluru	MGS8	200	6.21	8.003	3.67	0.2897	35.13	27.39
Calicut	CLS9	1000	24.82	7.954	3.67	0.3002	35.12	27.39
Calicut	CLS10	189	12.55	14.82	4.29	0.313	35.14	26.12
Calicut	CLS11	96	23.08	20.28	4.88	0.0332	35.44	25.03

were quantified, normalized and an equimolar pool of all the samples was made. This multiplexed library was further subjected to quality control (Agilent Bioanalyzer DNA chip, USA). The sequencing libraries generated from V3 and V4 amplicons from all the samples were sequenced using an Illumina paired-end overlapping sequencing. Sequence reads were binned according to index sequences and QC of the raw sequence data was performed by custom scripts. Low-quality reads were filtered out and trimmed based on observed quality pattern in the dataset. Read pairs with high sequence quality and overlapping regions were fused together to obtain a single read traversing the full length of V3 and V4 regions.

Bioinformatics analysis

The sequences which were less than 300 base pairs and those with less-than-average quality score (25 or less) were removed from the library. The taxonomic assignment of unassembled clean metagenomic sequences was performed using Ez-Taxone database[35] and BLASTX. Table 3 provides information related to the metagenomics reads of the samples.

Statistical analysis

Dominance, Simpson, Shannon, evenness, Brillouin, Menhinick, Margalef, equitability, Fisher_alpha, Berger-Parker, Chao-1, Whittaker, Harrison, Cody, Routledge, Wilson–Shmida, Mourelle, Harrison 2 and Williams indices of clonal and beta diversity were estimated using the PAST3 programs available from the University of Oslo (Norway) website. Relationship between chemical composition and (i) species diversity unifrac distances, (ii) species alpha diversity indices, and (iii) species beta diversity indices was determined by Mantel test. *P* values were calculated using 9999 permutations on rows and columns of dissimilarity matrices. Principal coordinate analysis (PCoA), canonical correlation analysis (CCorA), permutational analysis of variance (PERMANOVA) and

analysis of similarity (ANOSIM) were performed using the Past 3 software. For predictive functional analyses, the PICRUSt software package[36] was used to identify predicted gene families and associated pathways.

Analysis of predicated functional profiles for the identified microbial communities

The 16S rRNA sequencing datasets were analysed using PICRUSt script (normalize_by_copy_number.py script) for copy-number normalization[36]. Functional predictions were assigned up to KO tier-3 and categories including metabolism, genetic information processing, environmental information processing and cellular processes were analysed further. KEGG pathway analysis was carried out employing functions .py PICRUSt scripts followed by STAMP (Statistical Analysis of Metagenomic Profiles) software[37], with Welche's *t*-test and *P* value cut-off of 0.05 to reject null hypotheses. This identification of functional features of the genes and metabolic pathways has relevance in understanding metabolic processes in the context of the ecosystem.

Identification of bacterial markers by LDA effect size analysis

Linear discriminant analysis (LDA) effect size (LEfSe) analysis was utilized for identification of unique microbial communities present in different samples[38]. This analysis with LDA score threshold of 2 using on-line Galaxy version 1.0 was used to identify variations in bacterial diversity at specific locations and depths.

Results

Species diversity in the Arabian Sea

Water samples (total 11) collected from the Arabian Sea at specific locations and depths (Table 1 and Supplementary Table 1) were subjected to metagenomic

Table 3. Metagenomics reads information and taxonomic affiliations of bacteria present in samples collected from the Arabian Sea

Place	Site	Valid reads	OTUs	Average read length	Goods library coverage	Bacteria			
						Order	Family	Genus	Species
Goa	GAS1	48317	39457	423.53	0.23528	278	594	973	1390
Goa	GAS2	47364	38991	422.3	0.224305	334	673	1003	1338
Goa	GAS3	47953	38157	419.8	0.259608	332	712	1081	1460
Mangaluru	GAS4	44960	34992	416.53	0.283029	285	563	897	1256
Mangaluru	MGS5	43168	36154	421.56	0.208024	319	668	1045	1421
Mangaluru	MGS6	47834	38732	421.25	0.242735	314	653	1052	1459
Mangaluru	MGS7	48669	40730	423.61	0.213277	291	569	923	1288
Calicut	MGS8	48702	39181	421.18	0.249805	265	565	931	1333
Calicut	CLS9	30013	27721	438.47	0.087662	235	451	637	786
Calicut	CLS10	45356	38356	423.02	0.198585	318	633	961	1318
Calicut	CLS11	45726	37224	421.58	0.236824	294	608	965	1380

analysis using NGS technology of amplified rDNA libraries. A total of 498,062 (45,278.36 ± 5369.15 per sample) high-quality sequences with 3551 (1311.72 ± 186.45 per sample) distinct bacterial species were recorded (Table 3 and Supplementary Table 1). The data obtained after NGS were analysed extensively to obtain similarities and differences in the microbial flora in the OMZs and the non-OMZs. In all three sampling sites in the OMZs, 1371 species were common, while 777 species were found common at all depths (100, 200, 500 and 1000 m) across different sampling sites (Figure 2 a and b).

Community composition of the Arabian Sea

A large number of uncultured and novel microbes were abundant at these locations (Supplementary Table 1). Proteobacteria and SAR406 were common, while Firmicutes, Spirochaetes, Chloroflexi and Verrucomicrobia were present in relatively lower numbers. Alpha proteobacteria (20.43–35.51%, Figure 2 c) Deltaproteobacteria (11.03–15.93%) and Gamma proteobacteria (9.98–32.18%) were abundant in significant numbers in all the OMZ samples analysed. At the family level, SAR11-2_f (5.38–12.49%), Bacteria_uc_f (5.84–19.38%), Ruthia_f (0.69–7.56%), Arenicella_f (1.44–4.38%), Nitrospinaceae (1.86–5.89%) and Erythrobacteraceae (0.45–4.80%) were present across all samples (Figure 2 b). Bacterial orders such as SAR11 (10.73–23.81%), Bacteria_uc_o (5.84–19.39%), Ruthia (0.70–7.91%), Alteromonadales (0.34–14.17%) and Nitrospinaceae (1.96–6.11%) showed high abundance in all the samples. At 1000 m depth, SAR324_f (7.90–10.67%), Bacteria_uc_f (6.69–19.34%) and Erythrobacteraceae (0.94–4.80%) were predominant. Bacterial families such as Homogoneae and Thoreales were affiliated only with GAS4 sample, whereas Synarophyceae, Ceramiales, Euglenida and Cloacamonas were exclusively present in the CLS11 sample. Vaucheriales, Crenarchaeota, Pedinophyceae, Zetaproteobacteria and synergista were specific to the MGS5 sample. Nitrospi-

reae, Methanomicrobia and Bryopsida were exclusive to 200 m depth. SAR11-2_f (7.81–12.30%) and SAR11-1_f (6.31–12.32%) were predominant at 100 and 200 m depths, while Prochlorococcaceae (1.81–3.19%) was predominantly present at 100 m depth and SAR406_o_uc (1.17–2.69%) was abundant at 200 m depth.

Genera such as Bacteria_uc_g (5.85–19.38%), Pelagibacter (3.44–9.89%), SAR324_g (2.60–7.20%) were ubiquitous. *Croceicoccus* (1.25–2.14%) was predominantly present in samples from Goa compared to other samples. Correspondence analysis revealed that at 1000 m, *Methylopila*, *Mycoplasma*, *Asticcacaulis*, *Cellulomonas* and *Phalaenapsis* were exclusive to the MGS7 sample, whereas *Spirochaeta*, *Chroococcidiopsis*, *Thysira* and *Leeuwenhoekiella* were selectively present in the CLS9 sample. Water sample at 1000 m depth from Mangaluru (MGS5) revealed the presence of *Terasakiella*, *Chlorodendrales*, *Vaucheria*, *Congregibacter*, *Planktotelea* and *Pseudoflavinifactor*, whereas *Spirobacillus*, *Moraxella*, *Tiobacter*, *Roseburia*, *Marinoscillum*, *Thiohalophilus*, *Akkermansia*, *Caedobacter*, *Oceanicaulis*, *Epibacterium* and *Ditylium* were exclusively seen in Goa (GAS4). A more detailed analysis of data based at the species level revealed that Bacteria_uc_s, SAR406_f_uc_s, Ruthia_f_uc_s, Arenicella_f_uc_s, Nitrospinaceae_uc_s, Oceanospirillaceae_uc_s and Rhodospirillaceae_uc_s were present in high numbers in all samples that were analysed in the study (Figure 2 d and Supplementary Table 1).

Linear discriminant analysis effect size analysis

In order to determine the unique and predominant bacteria present at a particular location, a comparative assessment of the biodiversity was carried out. This resulted in the identification of specific marker families for different locations as well as for various sampling depths. Bacterial families, including Erysipelotrichi_uc_f, SAR11_uc, EU335161_o_uc, Pseudoalteromonadaceae
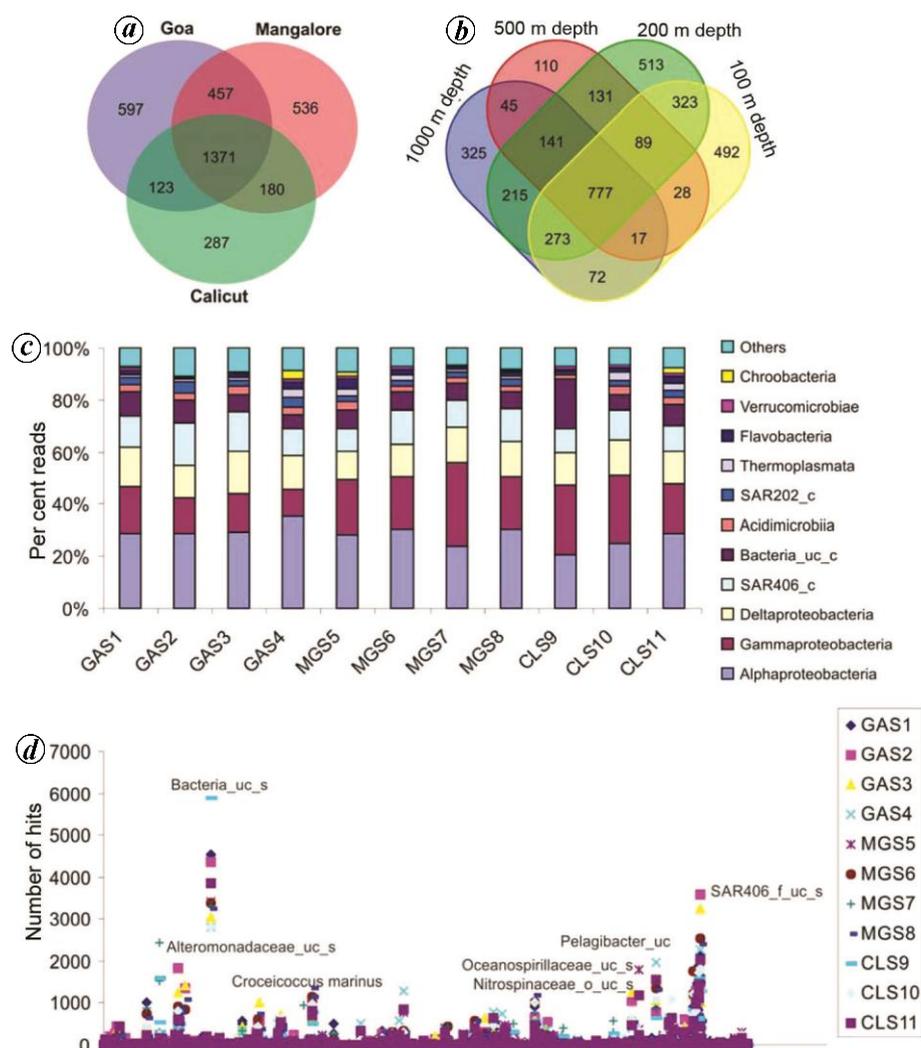
Figure 2. Venn diagrams showing the number of unique and shared species between (**a**) three sampling sites and (**b**) different depth of Arabian Sea. Distribution of predominant bacterial classes in the sample are based on 16S rRNA gene sequencing. **c**, Phyla distribution are displayed as stacked bar charts for individual mangrove samples (x-axis) against taxa abundance (y-axis). **d**, Abundance of species displayed as a scatter plot for individual samples (x-axis) against species abundance (y-axis).

and Alteromonadales_uc were specific to Calicut while family FJ444691_c_uc_f was seen in Mangaluru. Water samples from Goa showed significant enrichment of Salinisphaeraceae, EU686587_f, and Dehalococcoidales_uc. The analysis with respect to depth showed enrichment of bacterial families (total 66) such as Brumimicrobiaceae, Bacteriovoracaceae, Dinophysiaceae, Spirochaetaceae, and Chaetocerotaceae at a shallower depth (100 m), while families (total 22) such as Methylobacteriaceae, Halomonadaceae and Rhizobiaceae were found to be enriched at a depth of 500–1000 m.

Alpha and beta diversity of samples

Alpha diversity analysis highlighted the rich taxonomic diversity in the sea samples (Supplementary Table 2).

Simpson index of all samples close to 1 indicated the presence of highly diverse microbial communities. Shannon's index varied from 4.29 to 5.21, indicating high species richness in bacterial diversity in these sea samples. Evenness index ranged from 0.093 to 0.138, while Margalef richness index was also high, emphasizing the richness of bacterial species in the sea area. Chao-1 analysis predicted the number of bacterial species in each sample to be between 1106 and 2122 (Supplementary Table 2). No significant difference was observed in alpha diversity indices when pairwise comparison was carried out between Goa, Mangaluru and Calicut sampling sites (ANOVA $P > 0.05$; Mann–Whitney U test $P > 0.05$ for each comparison). Beta diversity indices of these sea samples are provided in Supplementary Table 3. At the species level, high beta diversity was observed in all the sea samples (Supplementary Table 3). This extensive
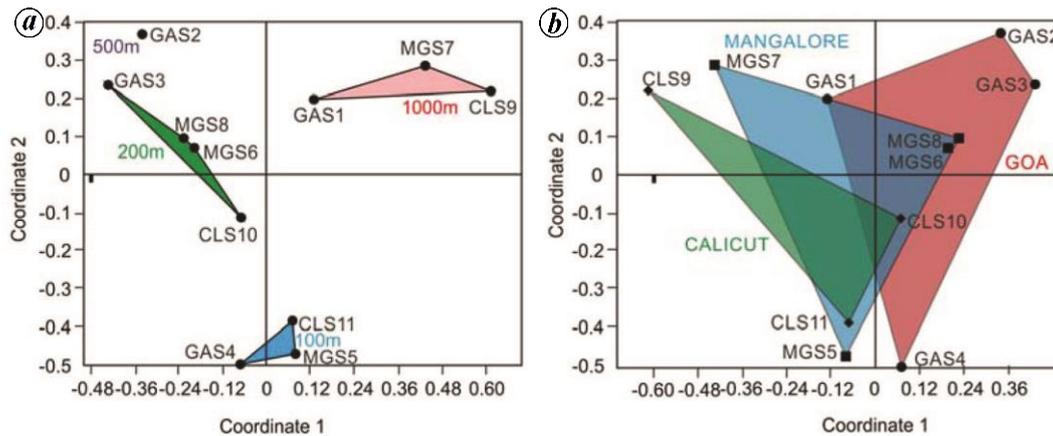
Figure 3. Principal coordinates analysis representation of the similarity matrix generated by cluster analysis. Samples from (*a*) depth and (*b*) collection site are represented by different shapes, and the distance between dots represents relatedness obtained from the similarity matrix.
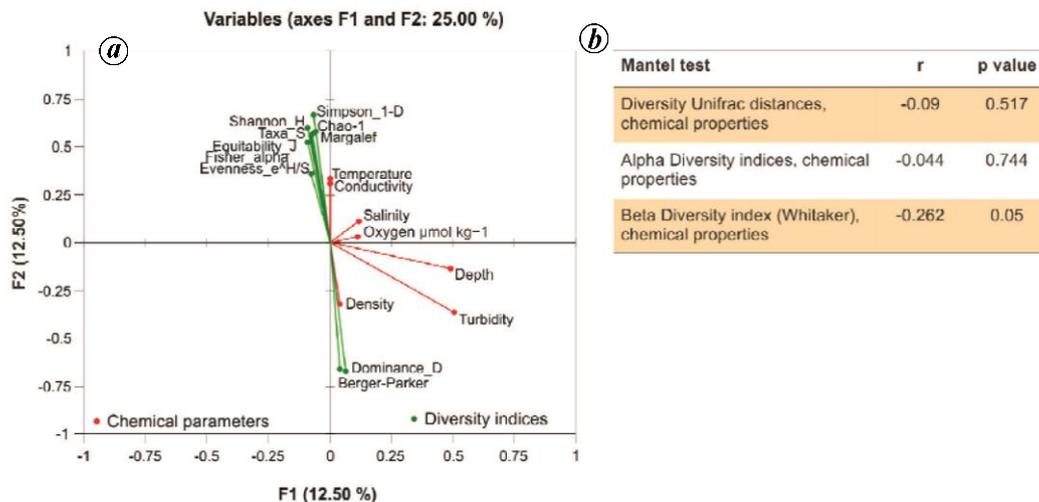

Figure 4. Association of environmental parameters and diversity indices: *a*, canonical correlation analysis; *b*, Mantel test.

analysis documented not only the rich and diverse microflora present in each sample, but also emphasized the differences in the microbial communities in the Arabian Sea.

Depth and geochemical parameters influencing the community structure

PCoA led to the identification of depth as an important determinant which influences characteristic and typical community structures of a given niche (Figure 3). Samples from similar depths clustered together, indicating that the communities in these locations are similar to each other. The correlations between environmental factors and alpha diversity indices were accessed by CCorA (Figure 4 *a*). Depth, turbidity and density were seen to influence the dominance of certain species, while temperature and conductivity correlated with the richness and evenness in samples (Figure 4 *a*). Beta diversity showed a

correlation with geochemical characters of the samples ($r = -0.262$; P value = 0.05), while alpha diversity ($r = -0.0004$; P value = 0.744) and unifrac distances among the sampling sites ($r = -0.09$; $P = 0.517$) were not affected by the geochemical characters of the samples (Figure 4 *b*). The sampling sites did not influence the community composition while depth was a major factor (PERMANOVA $F = 4.036$, $P = 0.0009$, ANOSIM, $R = 0.7222$, $P = 0.0008$) (Table 4).

OMZ versus non-OMZ samples

A comparative analysis and assessment of all samples showed the presence of 2718 species in the OMZs and 2223 species in the non-OMZs. Also, 1690 operational taxonomic units (OTUs) were common in the OMZs and non-OMZs, while 1328 and 533 OTUs were unique to the OMZs and non-OMZs respectively (Figure 5 *a*). This

clearly shows that although several common inhabitants are seen in the ocean, the depletion of oxygen changes the species pattern. Differential abundance was clearly visible when the top 50 families present at these locations were compared (Supplementary Table 1). In the case of OMZs, families SAR324, Ruthia, Arenicella, Zunongwangia, Rhodospirallacae, Nirtitreductor, Oleiobacter, Hyphomonas, Methylophaga, Clamydiales, Xanthomonadeacae, Sphingopyaix, Pararhodobacter, Anoxybacilllus, Gemella, Phenylobaterium and Legionella were present, while in the non-OMZs; Prochorobacteriacae, SAR11, Dianophyaceae, pelagomonadeacae, Delta proteobacteria, Firmicutes, Cytophagales, Flavobacteriales, Chroococcales, Dongicola, Planctomycetacia, Sphingobacteria, Draconibacterium were unique. Among the total 203 families which showed differential abundance, 86 were present in the OMZs and 117 in the non-OMZs respectively (Figure 5 b). PCoA revealed that the non-OMZ samples from similar depths clustered closely together; indicating that the communities in these locations are similar to each other and depth typically influences characteristic and typical community structures.

Functions associated with microbial communities

PICRUSt analysis is a bioinformatics software package to predict functional content from microbial community identification carried out by 16S rRNA-based metagenomic analysis. The per cent OTUs associated with different metabolic functions are xenobiotics biodegradation (2–3), glycan biosynthesis (2.3–3), energy metabolism (7–7.5) and (3–45) for lipid metabolism. Three terms under environmental information processing contain membrane transport (9–13%), signal transduction (1–2%), and signalling molecules and interaction (0.04–0.06%). DNA sequences encoding proteins such as nitrogenase FeMo-cofactor scaffold and assembly protein NifN, QscR quorum-sensing control repressor, cobalt–zinc–cadmium resistance protein CzcA; cation efflux system protein CusA, tellurite resistance protein-related protein, nitrite reductase [NAD(P)H] large subunit (EC 1.7.1.4), nitrogenase FeMo-cofactor synthesis FeS core scaffold, sulphur deprivation response regulator proteins and assembly protein NifB were found predominantly in samples from Goa compared to other samples. In seawater samples from Mangaluru, gene sequences encoding activities such as cobalt–zinc–cadmium resistance protein CzcD, sirohydrochlorin cobaltochelatase CbiK (EC 4.99.1.3)/sirohydrochlorin ferrochelatase (EC 4.99.1.4), type-IV fimbrial biogenesis protein PilY1, predicted L-rhamnose ABC transporter, methyl-accepting chemotaxis protein III, phage tail protein were enriched compared to other activities. Predominance of genes for 5-O-(4-coumaroyl)-D-quinate/shikimate 3′-hydroxylase (EC 1.14.13.36), beta-glucanase precursor (EC 3.2.1.73) (endo-beta-1,3-1,4 glucanase) (1,3-1,4-beta-D-glucan 4-glucanohydrolase), glutathione S-transferase C terminus, Shikimate/quinate 5-dehydrogenase I beta (EC 1.1.1.282), Tlr0729 protein, two-component system response regulator, putative Fe–S containing oxidoreductase, possible polygalacturonase (EC 3.2.1.15), microbial collagenase (EC 3.4.24.3), chitosanase, aspartate ammonia-lyase (EC 4.3.1.1), FtsK/SpoIIIE family protein and putative EssC component of type-VII secretion system were found in microbial communities from Goa and Calicut.
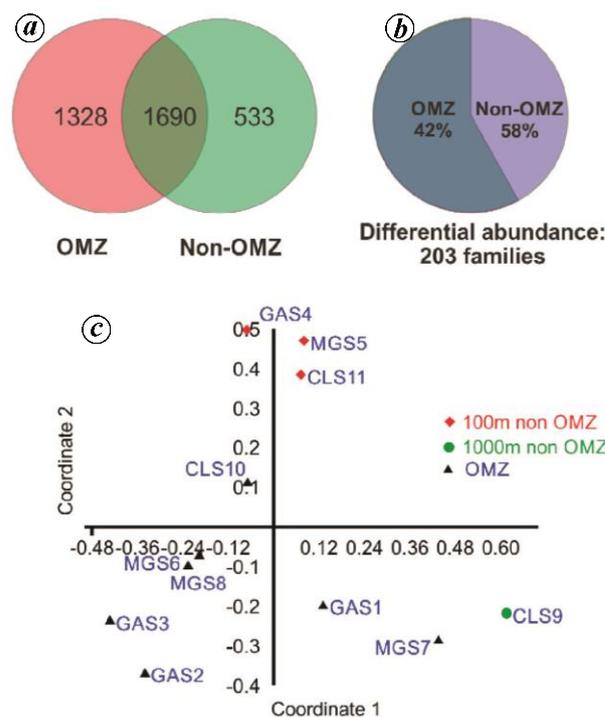

Figure 5. Comparison of microbial diversity in OMZs and non-OMZs areas. *a*, Vein diagram showing common and unique species in OMZ and non-OMZ areas. *b*, Differential abundance of families in OMZ and non-OMZ areas. *c*, Principal coordinates analysis (PCoA) representation.

Table 4. Effect of depth and sampling location on bacterial diversity of samples collected from the Arabian Sea

Test	F	P
Permanova		
Sampling site	1.041	0.416
Depth	3.503	0.0002
Pairwise comparison (*t* test)		
Goa versus Magaluru	0.9932	0.5742
Goa versus Calicut	1.294	0.2914
Calicut versus Mangaluru	0.8273	0.6308
100 m versus 200 m	4.952	0.0278
100 m versus 500 m	5.166	0.2496
100 m versus 1000 m	3.771	0.1022
200 m versus 500 m	2.472	0.1949
200 m versus 1000 m	3.668	0.0273
500 m versus 1000 m	1.75	0.2585

PICRUSt and STAMP analysis have identified OTUs associated with a few of these KO terms that differ significantly ($P < 0.05$) between samples collected from different locations. On comparison of samples from Goa and Mangaluru, dioxin degradation and translation proteins differed significantly, while processes related to polycyclic aromatic hydrocarbon degradation were enriched differentially in Calicut and Mangaluru samples. In Calicut and Goa samples, bacterial OTUs associated with phenylpropanoid biosynthesis, polycyclic aromatic hydrocarbon degradation, dioxin degradation showed differential enrichment.

Discussion

The Arabian Sea is typically characterized by the presence of vast OMZs and these are expanding further. Depletion of oxygen in the habitats changes the microbial composition and leads to alterations in the nutrient as well as elemental cycles. Analysis of the correlation between geochemical parameters and bacterial diversity is important in the understanding of the dynamics of microbial communities in OMZs. This study emphasizes that the Arabian Sea has high species richness with a complex community structure across oxygen gradients and between the depths of the Sea (Table 3 and Figure 2). Chao-1 analysis highlighted the presence of a diverse assemblage of indigenous microbial species that remain completely uncharacterized at present. The present analysis indicated that relationships between environmental variables, conductivity, temperature and oxygen concentration have a significant role in increasing the species richness and evenness in microbial communities (Figure 4). OMZ samples in the Arabian Sea displayed rich taxonomic diversity which typically showed depth-specific variation (Tables 3 and 4).

Nitrate reducing bacteria were present at all collection sites in OMZs and non-OMZs of the Arabian Sea. Reports from the suboxic zone of the Black Sea have identified single clade of nitrifying Crenarchaeota, which is closely related to *Nitrosopumilus maritimus*[39]. Global ocean sampling (GOS) database across diverse physico-chemical habitats and geographic locations has 1.2% *N. maritimus*[40]. Interestingly, *N. maritimus* is a cultured nitrifier isolated from a marine aquarium[41]. It has been shown that *N. maritimus* typically dominates low-depth samples[21]. However, in the present study, *N. maritimus* was underrepresented in low-depth samples.

Based on the metagenomic profiles of microbial assemblage, gene repertoire and predicted functions were assessed. Genes encoding nitrite/nitrate sensor proteins, nitrilase, nitrate reductase and nitrate reductase associated proteins were predominant in the datasets, emphasizing that nitrate/nitrite metabolism plays a key role in the dynamics of microbial communities and in the nitrogen

cycle in the OMZs (Figure 6). Naqvi *et al.*[42] have reported the presence of the nepheloid layer with significant amount of suspended matter caused by bacteria in the Arabian Sea, while an increase in nitrifying bacteria (both ammonium and nitrite oxidizers) has been suggested as the cause for such nepheloid layer[43]. Recent taxonomic, metagenomic and metatranscriptomic analyses of many OMZs have shown that diverse sulphur-oxidizing microbial communities are abundant and these are particularly enriched in γ-proteobacteria. Interestingly, the sulphate-reducing bacteria (SRB) were present throughout the water column at all collection sites in our analysis. The presence of SRB has been reported not only at the bottom sediments but also in aerobic surface waters and beach sediments[44]. It has been shown that SRB populations increase from the surface waters up to the oxic–anoxic boundary. Colourless sulphur-oxidizing bacteria have earlier been reported from the Arabian Sea. These bacteria are known to mediate the nitrogen cycle reductively even under autotrophic conditions[45]. SRB are also known to participate in nitrate reduction. Jayakumar *et al.*[46] and Ward *et al.*[47] reported the dominance of denitrifying bacteria in the biomass of OMZ and suggested that the denitrifying bacteria in this zone could be in a viable but non-culturable state.

The present analysis revealed that microorganisms involved in activities associated with sulphate metabolism were predominant, with sulphate permeases and reductase being predominant in the OMZs of the Arabian Sea. Additionally, together with other recent analyses[48–50], our results indicate the presence of an active sulphur oxidizing community in the Arabian Sea OMZ. It is likely that the sulphur cycle carried out by these SRB fuels nitrate reduction, thereby supplying additional substrates (nitrite and ammonia) for anammox bacteria. Comparative analysis of OMZ and non-OMZ samples revealed that species such as *Zunongwangia profunda*, *Roseovarius nubinhibens*, *hydrocarbonoclasticus*, *Prochlorococcus marinus* and *Ruegeria pomeroyi* were common in oxygen-depleted waters. These bacteria are known to be actively involved in dimethylsulfoniopropionate (DMSP) metabolism. This observation is important in the perspective of global climate change, since DMS is thought to play a key role by decreasing the absorption of solar radiation and thereby influencing temperature changes.

PICRUST analysis results obtained in this study indicate the presence of microorganisms harbouring genes such as *alkB*, *AlmA*, *CYP153A* and *AlkB* in the in Arabian Sea OMZ. Similar reports are available from the Atlantic Ocean and the Bay of Bengal[51–53]. Hydrocarbon-degrading organisms indicate the presence of alkanes and hydrocarbon. In comparison to Calicut, Goa and Mangaluru samples showed enrichment for polycyclic aromatic hydrocarbon degradation pathway, suggesting anthropogenic activities in these areas. Deep seawater samples from 100 to 500 m depth were found to be enriched in
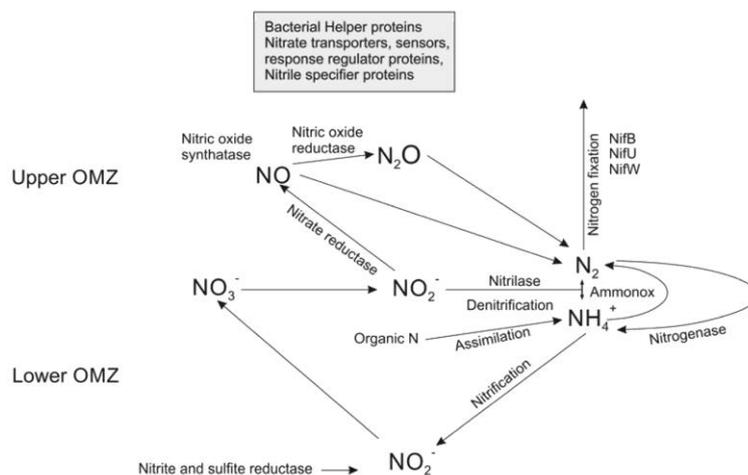
Figure 6. Probable nitrogen cycle in oxygen-depleted zones of the Arabian Sea.

(total 22) Methylobacteriaceae, Halomonadaceae, Alcanivoracaceae and Rhizobiaceae. Methylobacterium derives energy from the oxidation of thiosulphate to sulphate. The present study provides baseline data related to the diversity and potential microbial communities in oxygen-depleted waters, which could provide a basis for a understanding of the microbiological function, dynamics and distribution in the oceanic OMZs. Further, the results obtained from this study indicate location-specific functional divergence in the bacterial community. Therefore, it would be interesting to carry out detailed functional analysis for bacterial diversity from the Arabian Sea at more locations and with multiple samples during different seasons. Although our understanding of the OMZs and the interplay between geochemical processes and microbes has improved in recent years, the potential impacts of the OMZs on marine ecosystem structure and global geochemical cycling remain to be elucidated. In this context, we need to accelerate the exploration and discovery of microbes, and their interplay with geochemical processes in the OMZs.

Conflict of interest: The authors declare that they have no competing interests.

1. Naqvi, S. W. A., Some aspects of the oxygen-deficient conditions and denitrification in the Arabian Sea. *J. Mar. Res.*, 1987, **45**, 1049–1072.
2. McCreary, J. P., Yu, Z., Hood, R. R., Vinaychandran, P. N., Furue, R., Ishida, A. and Richards, K. J., Dynamics of the Indian Ocean oxygen minimum zones. *Prog. Oceanogr.*, 2013, **112**, 15–37.
3. Morrison, J. M., Codispoti, L. A., Gaurin, S., Jones, B., Manghnani, V. and Zheng, Z., Seasonal variation of hydrographic and nutrient fields during the US JGOFS Arabian Sea Process Study. *Deep Sea Res. II*, 1998, **45**, 2053–2101.
4. Naqvi, W. A., Geographical extent of denitrification in the Arabian Sea in relation to some physical processes. *Oceanol. Acta*, 1991, **143**, 281–290.
5. Stramma, L., Johnson, G. C., Sprintall, J. and Mohrholz, V., Expanding oxygen-minimum zones in the tropical oceans. *Science*, 2008, **320**, 655–658.
6. Sarmiento, J. L., Hughes, T. M., Stouffer, R. J. and Manabe, S., Simulated response of the ocean carbon cycle to anthropogenic climate warming. *Nature*, 1998, **393**, 245–249.
7. Keeling, R. F., Körtzinger, A. and Gruber, N., Ocean deoxygenation in a warming world. *Annu. Rev. Mar. Sci.*, 2010, **2**, 199–229.
8. DeSousa, S. N., Dileepkumar, M., Sardessai, S., Sarma, V. V. S. S. and Shirodkar, P. V., Seasonal variability in oxygen and nutrients in the central and eastern Arabian Sea. *Curr. Sci.*, 1996, **71**, 847–851.
9. Morrison, J. M. *et al.*, The oxygen minimum zone in the Arabian Sea during 1995. *Deep Sea Res. II*, 1999, **468**, 1903–1931.
10. Resplandy, L., Lévy, M., Bopp, L., Echevin, V., Pous, S., Sarma, V. V. S. S. and Kumar, D., Controlling factors of the oxygen balance in the Arabian Sea's OMZ. *Biogeosciences*, 2012, **9**, 5095–5109.
11. Gilly, W. F., Beman, J. M., Litvin, S. Y. and Robison, B. H., Oceanographic and biological effects of shoaling of the oxygen minimum zone. *Annu. Rev. Mar. Sci.*, 2013, **5**, 393–420.
12. Thamdrup, B., Dalsgaard, T., Jensen, M. M., Ulloa, O., Farias, L. and Escribano, R., Anaerobic ammonium oxidation in the oxygen-deficient waters off northern Chile. *Limnol. Oceanogr.*, 2006, **51**, 2145–2156.
13. Cohen, Y. and Gordon, L. I., Nitrous oxide in the oxygen minimum of the eastern tropical North Pacific: evidence for its consumption during denitrification and possible mechanisms for its production. *Deep-Sea Res.*, 1978, **25**, 509–524.
14. Diaz, R. J. and Rosenberg, R., Spreading dead zones and consequences for marine ecosystems. *Science*, 2008, **321**, 926–929.
15. Lam, P. and Kuypers, M. M., Microbial nitrogen cycling processes in oxygen minimum zones. *Annu. Rev. Mar. Sci.*, 2011, **3**, 317–345.
16. Ward, B. and Zafiriou, O., Nitrification and nitric oxide in the oxygen minimum of the eastern tropical North Pacific. *Deep-Sea Res.*, 1988, **35**, 1127–1142.
17. Ward, B., Glover, H. and Lipschultz, F., Chemoautotrophic activity and nitrification in the oxygen minimum zone off Peru. *Deep-Sea Res.*, 1989, **36**, 1031–1051.
18. Füssel, J. *et al.*, Nitrite oxidation in the Namibian oxygen minimum zone. *ISME J.*, 2011, **6**, 1200–1209.
19. Hamersley, M. R. *et al.*, Anaerobic ammonium oxidation in the Peruvian oxygen minimum zone. *Limnol. Oceanogr.*, 2007, **52**, 923–933.
20. Lam, P. *et al.*, Revising the nitrogen cycle in the Peruvian oxygen minimum zone. *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 4752–4757.
21. Stewart, F. J., Ulloa, O. and DeLong, E. F., Microbial metatranscriptomics in a permanent marine oxygen minimum zone. *Environ. Microbiol.*, 2012, **14**, 23–40.

22. Ganesh, S., Parris, D. J., Delong, E. F. and Stewart, F. J., Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. *ISME J.*, 2014, **8**, 187–211.

23. Hawley, A. K., Brewer, H. M., Norbeck, A. D., Paša-Toliae, L. and Hallam, S. J., Metaproteomics reveals differential modes of metabolic coupling among ubiquitous oxygen minimum zone microbes. *Proc. Natl. Acad. Sci. USA*, 2014, **111**, 11395–11400.

24. Ganesh, S., Bristow, L. A., Larsen, M., Sarode, N., Thamdrup, B. and Stewart, F. J., Size-fraction partitioning of community gene transcription and nitrogen metabolism in a marine oxygen minimum zone. *ISME J.*, 2015, **9**, 2682–2696.

25. Jayakumar, D. A., Francis, C. A., Naqvi, S. W. A. and Ward, B. B., Diversity of nitrite reductase genes *nirS* in the denitrifying water column of the coastal Arabian Sea. *Aquat. Microbiol. Ecol.*, 2004, **34**, 69–78.

26. Jayakumar, A., Al-Rshaidat, M. M. D., Ward, B. B. and Mulholland, M. R., Diversity, distribution, and expression of diazotroph *nifH* genes in oxygen-deficient waters of the Arabian Sea. *FEMS Microbiol. Ecol.*, 2012, **82**, 597–606.

27. Kalvelage, T. *et al.*, Aerobic microbial respiration in oceanic oxygen minimum zones. *PLoS ONE*, 2015, **107**, e0133526.

28. Stevens, H. and Ulloa, O., Bacterial diversity in the oxygen minimum zone of the eastern tropical South Pacific. *Environ. Microbiol.*, 2008, **10,** 1244–1259.

29. Glass, J. B. *et al.*, Meta-omic signatures of microbial metal and nitrogen cycling in marine oxygen minimum zones. *Front. Microbiol.*, 2015, **6**, 1–13.

30. Carlson, C. A., Morris, R., Parsons, R., Treusch, A. H., Giovannoni, S. J. and Vergin, K., Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the northwestern Sargasso Sea. *ISME J.*, 2009, **3**, 283–295.

31. Gonsalves, M. J. *et al.*, Predominance of anaerobic bacterial community over aerobic community contribute to intensify 'oxygen minimum zone' in the eastern Arabian Sea. *Continent. Shelf Res.*, 2011, **31**, 1224–1235.

32. Pitcher, A., Villanueva, L., Hopmans, E. C., Schouten, S., Reichart, G. J., Sinninghe, J. and Damsté, S., Niche segregation of ammonia-oxidizing archaea and anammox bacteria in the Arabian Sea oxygen minimum zone. *ISME J.*, 2011, **5**, 1896–1904.

33. Wyman, M., Hodgson, S. and Bird, C., Denitrifying Alphaproteo bacteria from the Arabian Sea that express *nosZ*, the gene encoding nitrous oxide reductase, in oxic and suboxic waters. *Appl. Environ. Microbiol.*, 2013, **798**, 2670–2681.

34. Lüke, C., Speth, D. R., Kox, M. A. R., Villanueva, L. and Jetten, M. S. M., Metagenomic analysis of nitrogen and methane cycling in the Arabian Sea oxygen minimum zone. *Peer J.*, 2016, **4**, e1924; doi:10.7717/peerj.1924.

35. Kim, O. S. *et al.*, Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int. J. Syst. Evol. Microbiol.*, 2012, **62**, 716–721; doi:10.1099/ijs.0.038075-0.

36. Langille, M. G. *et al.*, Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnol.*, 2013, **31**, 814–821.

37. Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M. and Tanabe, M., Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.*, 2014, **42**(D1), D199–D205.

38. Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S. and Hutten, C., Metagenomic biomarker discovery and explanation. *Genome Biol.*, 2011, **12**, R60.

39. Labrenz, M., Sintes, E., Toetzke, F., Zumsteg, A., Herndl, G. J., Seidler, M. and Jürgens, K., Relevance of a crenarchaeotal subcluster related to *Candidatus Nitrosopumilus maritimus* 830 to ammonia oxidation in the suboxic zone of the central Baltic Sea. *ISME J.*, 2010, **412**, 1496–1508.

40. Walker, C. B. *et al.*, *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine Crenarchaea. *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 8818–8823.

41. Konneke, M. E., de la Torre, J.R., Walker, C. B., Waterbury, J. B. and Stahl, D. A., Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature*, 2005, **437**, 543–546.

42. Naqvi, S. W. A., Kumar, M. D., Narvekar, P. V., de Souza, S. N., George, M. D. and D'Silva, C., An intermediate nepheloid layer associated with high microbial metabolic rates and denitrification in the Northwest Indian Ocean. *J. Geophys. Res.*, 1993, **98**, 16469–16479.

43. Koppelmann, R., Zimmermann-Timm, H. and Weikert, H., Bacterial and zooplankton distribution in deep waters of the Arabian Sea. *Deep-Sea Res. I*, 2005, **52**, 2184–2192.

44. LokaBharathi, P. A. and Chandramohan, D., Sulfate-reducing bacteria from the Arabian Sea – their distribution in relation to thiosulfate-oxidising and heterotrophic bacteria. *Bull. Mar. Sci.*, 1990, **47**, 622–630.

45. LokaBharathi, P. A., Nair, S. and Chandramohan, D., Anaerobic sulfide-oxidation in marine colorless sulfur-oxidizing bacteria. *J. Mar. Biotechnol.*, 1997, **52**(3), 172–177.

46. Jayakumar, A., O'Mullan, G. D., Naqvi, S. W. A. and Ward, B. B., Denitrifying bacterial community composition changes associated with stages of denitrification in oxygen minimum zones. *Microb. Ecol.*, 2009, **58**, 350–362.

47. Ward, B. B., Devol, A. H., Rich, J. J., Chang, B. X., Bulow, S. E., Naik, H., Pratihary, A. and Jayakumar, A., Denitrification as the dominant nitrogen loss process in the Arabian Sea. *Nature*, 2009, **461**, 78–81.

48. Lavik, G. *et al.*, Detoxification of sulphidic African shelf waters by blooming chemolithotrophs. *Nature*, 2009, **457**, 581–586.

49. Walsh, D. A. *et al.*, Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones. *Science*, 2009, **326**, 578–582.

50. Canfield, D. E. *et al.*, A cryptic sulfur cycle in oxygen-minimum-zone waters off the Chilean coast. *Science*, 2010, **330**, 1375–1378.

51. Wang, L. P., Wang, W. P., Lai, Q. L. and Shao, Z. Z., Gene diversity of CYP153A and AlkB alkane hydroxylases in oil-degrading bacteria isolated from the Atlantic Ocean. *Environ. Microbiol.*, 2010, **12**, 1230–1242.

52. Wang, W. and Shao, Z., Diversity of flavin-binding monooxygenase genes (*almA*) in marine bacteria capable of degradation long-chain alkanes. FEMS *Microbiol. Ecol.*, 2012, **80**, 523–533.

53. Rajpathak, S. N., Banerjee, R., Mishra, P. G., Khedkar, A. M., Patil, Y. M., Joshi, S. R. and Deobagkar, D. D., An exploration of microbial and associated functional diversity in the OMZ and non-OMZ areas in the Bay of Bengal. *J. Biosci.*, 2018, **43**, 635–648; https://doi.org/10.1007/s12038-018-9781-2.