Microbial diversity of the Arabian Sea in the Oxygen minimum zones by metagenomics approach

Short Title: Microbial diversity of the Arabian Sea

Mandar S. Paingankar[#], Kedar Ahire, Pawan Mishra, Shriram Rajpathak, Deepti D. Deobagkar*

Molecular Biology Research Laboratory, Centre for Advanced Studies, Department of Zoology, Savitribai Phule Pune University, Pune - 411 007, India.

[#]Current Address - Department of Zoology, Government Science College, Chamorshi Road, Gadchiroli -442605, Maharashtra, India.

*Correspondence

Prof. Deepti D. Deobagkar,

ISRO Chair Professor, ISRO Science and technology Cell &

Department of Zoology, Centre for Advance Studies,

Savitribai Phule Pune University, Pune 411007, India.

Tel.: +91992118487

deepti.deobagkar@gmail.com

ABSTRACT

Large oxygen depleted areas known as oxygen minimum zones (OMZ) have been observed in 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 the temperature rise, acidification and deoxygenation can lead to major consequences wherein the plants, fish and other biota will struggle to survive in the ecosystem.

The current study has identified the microbial community structure using NGS based metagenomics analysis in the water samples collected at different depth from the oxygen depleted and non-OMZ areas of Arabian Sea. Environmental variables such as depth, site of collection and oxygen concentration appeared to influence the species richness and evenness among microbial communities in these locations. Our observations clearly indicate that population dynamics of microbes consisting of nitrate reducers accompanied by sulphate reducers and sulphur oxidizers participate in the interconnected geochemical cycles of the OMZ areas. In addition to providing baseline data related to the diversity and microbial community dynamics in oxygen-depleted water in the OMZ; such analysis can provide insight into processes regulating productivity and ecological community structure of the ocean.

Keywords: Arabian Sea, bacterial diversity, oxygen minimum zones, metagenomics, sulphur and nitrogen metabolism, Goa, Mangalore, Calicut.

INTRODUCTION

The Oxygen minimum zone (OMZ) in the Arabian-Sea is the second-most intense OMZ amongst the tropical oceans in the world[1,2] with a near-total depletion of oxygen at depths from 200 to 1000m [3]. In these locations, suboxic levels (≲5 μmol O2/kg) of oxygen are seen over vast areas at different depths and denitrification occurs in its upper portion[4]. Geochemical observations indicate that oxygen minimum zones have expanded over the past decades[5] and could expand further in response to the 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 the OMZ 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 the oxygen minimum zones (OMZs) in the Arabian Sea has become the major concern because of its impact on the marine ecosystems. The expansion of the OMZs due to climate change and its impacts on the ecosystems and the atmosphere is multi-dimensional and requires intense study.

OMZ is characterized by high nitrite accumulations and very low or undetectable oxygen concentrations[12]. The nitrous oxide (N_2O) concentration in the OMZ 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 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 OMZ have revealed that complex communities (such as nitrifiers) play an important role in N cycle in the OMZ[18]. Members of the Planctomycetes, Thaumarchaeota and Nitrospinae phyla have been observed to perform the majority of anammox, ammonia oxidation and nitrite oxidation and play important role in the OMZ dynamics[12,18-24]. Although some reports exist, the 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 OMZ areas[28,29]. A few studies have been carried out to understand the microbial diversity in the OMZ areas of Arabian Sea [30-34]. The special growth requirements of these microbes and abundance of uncultured organisms (over 99%) make NGS based metagenomic the method of choice in order to unravel the complexities of microbial communities, their dynamics and ecological significance.

In the current study, water samples collected from different depths of sea (100 to 1000 meters across transect) from Goa, Mangalore and Calicut (Supplementary Information Table SI1) were processed for high throughput next generation sequencing based metagenomics (based on 16S rRNA gene sequencing). The microbial diversity and predicted metabolic activities associated with these microbial communities in OMZ and non-OMZ areas in Arabian Sea of India provide valuable insight into the nature of biogeochemical processes.

MATERIAL AND METHODS

Sample Collection

Water samples at different depths were collected during the Sagar Sampada cruise (Sagar Sampada Cruse Number 340, 16 May 2015 to 08 June 2015) from Goa (GAS1, GAS2, GAS3 and GAS4; distance from coast ranging from 51 km to 90 km), Mangalore (MGS5,

MGS6, MGS7 and MGS8; distance from coast ranging from 52 km to 84 km) and Calicut (CLS9, CLS10 and CLS11; distance from coast ranging from 66 km to 109 km) (Supplementary Information Table SI1; Fig. 1) A conductivity–temperature–depth (CTD) system equipped with attached oxygen and turbidity measurement sensors was deployed to record the physical properties of the water (Supplementary Information Table SI2) and the samples were grouped into OMZ and non OMZ .

DNA extraction

1000 ml water was collected from each sampling sites and organisms collected by filtering water through 0.22 µm filter (Millex, Merck Millipore, USA) were utilised for DNA isolation using Power water DNA isolation kit (MoBio laboratories Inc. Carlsbad, CA). DNA isolation was carried out on the ship to avoid degradation of DNA. 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'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG3'; 16S Amplicon PCR Reverse Primer 5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAA TCC3') 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 thepurified PCR amplicons were quantified, normalized and an equimolar pool of all the samples was made. This multiplexed library was further subjected to QC using an Agilent Bioanalyzer DNA Chip. 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 data set. Read pairs with high sequence quality and overlapping regions were fused together to obtain a single read traversing full length of V3 and V4 region.

Bioinformatics analysis

The sequences which were less than 300 bps and sequences 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. Information related to the metagenomics reads of the samples is depicted in Table 1.

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 website link Relationship between chemical composition and (i) species diversity unifrac distances, (ii) species alpha diversity indices and (iii) species Beta diversity indices were determined by Mantel tests. P values were calculated using 9999 permutations on rows and columns of dissimilarity matrices. Principal coordinate analysis (PCO), canonical correlation analysis (CCor), permutational analysis of variance (PERMANOVA) and analysis of similarity (ANOSIM) was performed using the Past 3 software. For the 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 data sets were analysed by 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 by 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 was considered 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 (LEfSe) analysis

Linear Discriminant Analysis (LDA) Effect Size (LEfSe) analysis was utilised for identification of unique microbial communities present in different samples[38]. The LEfSe analysis with LDA score threshold of 2 using online Galaxy version 1.0 was used to identify variations in bacterial diversity at specific locations and depths.

3. RESULTS

3.1 Species diversity in Arabian Sea

Water samples (total 11) collected from Arabian Sea at specific locations and depths (Supplementary Information Table SI1, SI3) were subjected to metagenomic analysis using next generation sequencing technology of amplified rDNA libraries. A total of 498062 (45278.36 ±5369.15 per sample) high-quality sequences with 3551 (1311.72±186.45 per sample) distinct bacterial species were recorded (Supplementary Information Table 3) were identified. This data was analysed extensively to identify if there were differences in the OMZ and the non OMZ regions. In all three sampling sites in OMZ 1371 species were common. while 777 species were found common at all depths (100m, 200m, 500m and 1000m) across different sampling sites (Fig. 2A, 2B).

Community composition of the Arabian Sea

A large number of uncultured and novel microbes were abundant at these locations (Supplementary Information SI3). Proteobacteria and SAR406 were common while Firmicutes, Spirochaetes, Chloroflexi and Verrucomicrobia were present in relatively lower numbers. Alpha proteobacteria (20.43-35.51%) (Fig. 2C) 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%), Erythrobacteraceae (0.45-4.80%) were present across all samples (Fig. 2B). 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%), Nitrospinaceae (1.96-6.11%) showed high abundance in all the samples. At 1000m 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 CLS11 sample. Vaucheriales, Crenarchaeota, Pedinophyceae, Zetaproteobacteria and synergista were specific to MGS5 sample. Nitrospireae, Methanomicrobia, Bryospida were exclusive to 200m depth. SAR11-2_f (7.81-12.30%) and SAR11-1_f (6.31-12.32%) were predominant at 100m and 200m depth while Prochlorococcaceae (1.81-3.19%) was predominantly present at 100m depth. SAR406_o_uc (1.17-2.69%) was abundant at 200m 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 as compared to other samples. Correspondence analysis revealed that in the depth of 1000 meters, *Methylopila, Mycoplasma, Asticcacaulis,*

Cellulomonas, Phalanopsis were exclusive to MGS7 sample whereas *Spirochaeta, Chroococcidiopsis, Thysira, Leeuwenhoekiella* were selectively present in CLS9 sample. Water sample at 1000 meters depth from Mangalore (MGS5) revealed presence of *Terasakiella, Chlorodendrales, Vaucheria, Congregibacter, Planktotelea, Pseudoflavinifactor* whereas *Spirobacillus, Moraxella, Tiobacter, Roseburia, Marinoscillum, Thiohalophilus, Akkermansia, Caedobacter, Oceanicaulis, Epibacterium, 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 our study (Fig. 2D, Supplementary Information Table SI3).

Linear Discriminant Analysis (LDA) Effect Size (LEfSe) analysis

In order to determine the unique and predominant bacteria present at a particular location, a comparative assessment of the biodiversity LefSe was carried out. This resulted in the identification of specific marker families for different locations as well as for the various sampling depths (Supplementary information SI2; SI3). Bacterial families including Erysipelotrichi_uc_f, SAR11_uc, EU335161_o_uc, Pseudoalteromonadaceae and Alteromonadales_uc were specific to Calicut while family FJ444691_c_uc_f was seen in Mangalore. Water samples from Goa Salinisphaeraceae, EU686587_f, and Dehalococcoidales_uc showed significant enrichment. LefSe analysis with respect to the depth showed enrichment of bacterial families (total 66) such as Brumimicrobiaceae, Bacteriovoracaceae, Dinophysiaceae, Spirochaetaceae, and Chaetocerotaceae at shallower depth (100m) while bacterial families (In total 22) such as Methylobacteriaceae, Halomonadaceae and Rhizobiaceae were found to be enriched at the depth of 500-1000m.

Alpha and Beta diversity of samples

Alpha diversity analysis highlighted the rich taxonomic diversity in the sea samples (Supplementary Information Table SI4). Simpson index of all samples close to 1 for all samples indicated the presence of highly diverse microbial communities in samples. Shannon's index varied from 4.29 to 5.21 indicating the 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 emphasising the richness of bacterial species in the sea area. Chao-1 analysis predicted the number of bacterial species in each sample to be between 1106-2122 (Supplementary Information Table SI4). No significant difference was observed in alpha diversity indices when pair wise comparison was carried out between Goa, Mangalore 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 depicted in Supplementary Information Table SI5. At species level, high beta diversity was observed in all the sea samples (Supplementary Information Table SI5). This extensive analysis documented not only the rich and diverse micro flora present in each sample but also emphasised the differences in the microbial communities in the Arabian Sea.

*Depth and geochemical parameter influences the community structure*Principal component analysis (PCoA) led to the identification of depth as an important determinant which influences characteristic and typical community structures of a given niche. (Fig. 3) Samples from similar depth clustered together, indicating that the communities in these locations are very similar to each other. The correlations between environmental factors and alpha diversity indices were accessed by Canonical correlation analysis (CCorA) (Fig. 4A). 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 (Fig. 4 A). Beta diversity showed a correlation with geochemical characters of sample ($r = -0.262$; p value = 0.05) while alpha diversity ($r = -0.0004$; p value = 0.744) and unifrac

243 distances among the sampling sites (r = -0.09; p = 0.517) were not affected by the
244 geochemical characters of sample (Fig. 4B). The sampling site did not influence the
245 community composition while depth was a major factor (PERMANOVA (F= 4.036
246 *P*=0.0009ANOSIM R= 0.7222, *P*= 0.0008) (Table 2).

247 *OMZ vs Non-OMZ samples*

248 A comparative analysis and assessment of all samples showed presence of 2718
249 species in OMZ areas and 2223 species in non OMZ. It was seen that 1690 operational
250 taxonomic units (OTUs) were common in OMZ and non OMZ while 1328 and 533 are
251 unique to OMZ and non-OMZ respectively (Fig. 5A). This clearly documents that although
252 several common inhabitants are seen in the ocean the depletion of oxygen is changing the
253 species pattern. Differential abundance was clearly visible when the top 50 families present
254 at these locations were compared (Supplementary Information Table SI3). In case of OMZ,
255 SAR324, Ruthia, Arenicella, Zunongwangia, Rhodospirallacae, Nirtitreductor,
256 Oleiobacter, Hyphomonas, Methylophaga, Clamydiales, Xanthomonadeacae,
257 Sphingopyaix, Pararhodobacter, Anoxybacilllus, Gemella, Phenylobaterium and
258 Legionella families were present while in non-OMZ Prochorobacteriacae, SAR11,
259 Dianophyaceae, pelagomonadeacae, Delatproteobacteria, Firmicutes, Cytophagales,
260 Flavobacteriales, Chroococcales, Dongicola, Planctomycetacia, Sphingobacteria,
261 Draconibacterium familes were unique. Out of total 203 families which showed differential
262 abundance, 86 were present in OMZ whereas 117 in Non-OMZ respectively (Fig. 5B). The
263 principal component analysis (PCoA) revealed that non OMZ samples from similar depths
264 clustered closely together; indicating that the communities in these locations are similar to
265 each other and the depth typically influences characteristic and typical community
266 structures.

267 *Functions associated with microbial communities*

PICRUSt analysis is a bioinformatics software package designed to predict functional content from microbial community identification carried out by 16S rRNA based metagenomic analysis. The percent 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, Sulfur deprivation response regulator proteins and assembly protein NifB were found predominantly in samples from Goa as compared to other samples. In sea water samples from Mangalore, 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 as 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, putative EssC component of type VII secretion system were dominant in microbial communities from Goa and Calicut.

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

DISCUSSION

Arabian Sea is typically characterised by presence of vast areas of 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 the geochemical parameters and bacterial diversity is important in the understanding the dynamics of microbial communities in the OMZs. This study emphasises that Arabian Sea has high species richness with a complex community structure across oxygen gradients and between the depths of sea (Fig. 2). Chao-1 analysis highlighted presence of diverse assemblage of indigenous microbial species that remain completely uncharacterized at present. Our analysis indicated that relationships between environmental variables, conductivity, temperature and oxygen concentration have significant role in increasing the species richness and evenness in microbial communities (Fig. 4). OMZ samples in the Arabian sea displayed rich taxonomic diversity which typically showed a depth specific variation.

316 Nitrate reducing bacteria were present at all collection sites in OMZ and non OMZ
317 areas of Arabian Sea. Reports from the suboxic zone of the Black Sea have identified
318 single clade of nitrifying Crenarchaeota which is closely related to *Nitrosopumilus*
319 *maritimus*[39]. Global Ocean Sampling (GOS) database across diverse physiochemical
320 habitats and geographic locations has 1.2% *N. Maritimus*[40]. Interestingly, *N. Maritimus* is a
321 cultured nitrifier isolated from a marine aquarium[41]. It has been shown that *N. maritimus*
322 typically dominates low depth samples[21]. However, in the current study, *N. maritimus* was
323 underrepresented in low depth samples.

324 Based on the metagenomic profiles of microbial assemblage gene repertoire and
325 predicted functions were assessed. Genes encoding nitrite/nitrate sensor proteins, nitrilase,
326 nitrate reductase, nitrate reductase associated proteins werepredominant in the datasets,
327 emphasising that nitrate/nitrite metabolism plays a key role in the dynamics of microbial
328 communities in the OMZ areas and play important role in nitrogen cycle in OMZ (Fig. 6).
329 Naqvi *et al.* [42] have reported the presence of the nepheloid layer with significant amounts
330 of suspended matter caused by bacteria in Arabian Sea while an increase in nitrifying
331 bacteria (both ammonium and nitrite oxidizers) has been suggested to be the cause for
332 such nepheloid layer[43]. Recent taxonomic, metagenomic, and metatranscriptomic analysis
333 of many OMZs has shown that diverse sulphur-oxidizing microbial community are
334 abundant and these communities are particularly enriched in γ-proteobacteria.
335 Interestingly, the sulphate reducing bacteria (SRB) were present through-out the water
336 column at all collection sites in our analysis. The presence of SRB has been shown not
337 only at the bottom sediments but also in aerobic surface waters and beach sediments[44]. It
338 has been shown that SRB populations increase from the surface waters up to the oxic–
339 anoxic boundary. Colourless sulphur-oxidizing bacteria have earlier been reported from the
340 Arabian Sea and these bacteria are known to mediate nitrogen cycle reductively even under

341 autotrophic conditions[45]. SRBs are also known to participate in nitrate reduction.
342 Jayakumar *et al.*[46] and Ward *et al.*[47] reported the dominance of denitrifying bacteria in the
343 biomass of the OMZ and suggested that the denitrifying bacteria in this zone could be in a
344 viable but non culturable state.

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

358 PICRUST analysis results obtained in the current study indicated presence of
359 microorganisms harboring genes such as alkB, AlmA, CYP153A and AlkB are enriched in
360 the OMZ in Arabian sea similar to the reports from Atlantic Ocean and Bay of Bengal [51-53].
361 Presence of hydrocarbon degrading organisms points to presence of alkanes and
362 hydrocarbon. In comparison to Calicut, Goa and Mangalore samples showed enrichment
363 for polycyclic aromatic hydrocarbon degradation pathway suggesting anthropogenic
364 activities in these areas. Deep sea water samples from depth 100-500 m depth were found
365 to be enriched in (In total 22) Methylobacteriaceae, Halomonadaceae, Alcanivoracaceae

and Rhizobiaceae. Methylobacterium which derive energy from the oxidation of thiosulfate to sulfate. The current study provides baseline data related to the diversity and potential microbial communities in oxygen-depleted water which could provide a basis for better understanding of the microbiological function, dynamics, and distribution in the oceanic OMZ. Further the results obtained in this study indicate the location specific functional divergence in bacterial community and therefore it would be interesting to carry out detailed functional analysis for bacterial diversity form Arabian Sea at more locations and with multiple samples in different seasons. Although our understanding of the OMZ and interplay between geochemical processes and microbes has been improving in recent years, the potential impacts of OMZs on marine ecosystem structure and global geochemical cycling remains to be completely elucidated. In this context we need to accelerate the exploration and discovery of microbes and their interplay with geochemical processes in the OMZ.

Acknowledgement

We thank Prof. Dileep N. Deobagkar for valuable suggestions. The authors are grateful to Dr. M. Sudhakar, Dr. Saravannane, and Dr. R. A. Shivaji for their help during the sample collection.

Author Contributions

DDD, MSP designed the experiments; DDD, KA, MSP, SR, PM carried out the work; DDD, MSP, SR, PM interpreted the results; DDD, MSP wrote the manuscript.

Funding

The authors would like to acknowledge financial support provided by the Ministry of Earth Sciences (MoES), India under the Microbial Oceanography project. This project was coordinated through CMLRE.

Conflict of interests

The authors declare that they have no competing interests.

References

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: Topical Stud. Oceanogr.*, 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. *Ann. 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., Codispoti, L.A., Smith, S.L, Wishner, K, Flagg C, Gardner WD, *et al.,* The oxygen minimum zone in the Arabian Sea during 1995. *Deep Sea Res. II: Topical Stud. Oceanogr.,* 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. *Biogeosci.,* 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. *Ann. 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. *Ann. 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., Lam, P., Lavik, G., Jensen, M.M., Holtappels, M., Günter, M., *et al.*, Nitriteoxidation in the Namibian oxygen minimum zone. *ISME J.*, 2011, **6**, 1200–1209.

19. Hamersley, M.R., Lavik, G., Woebken, D., Rattray, J.E., Lam, P., Hopmans, E.C., *et al.,* Anaerobic ammonium oxidation in the Peruvian oxygen minimum zone. *Limnol. Oceanogr.*, 2007, **52,** 923-933.

20. Lam, P., Lavik, G., Jensen, M.M., van de Vossenberg, J., Schmid, M., Woebken, D., *et al.,* Revising the nitrogen cycle in the Peruvian oxygen minimum zone. *Proc. Nat. 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. Nat. 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. *Aqua. Microb. Ecol.*, 2004, **34**, 69-78.

462 26. Jayakumar, A., Al-Rshaidat, M.M.D., Ward, B.B. and Mulholland, M.R., Diversity,
463 distribution, and expression of diazotroph nifH genes in oxygen-deficient waters of
464 the Arabian Sea. *FEMS Microb. Ecol.*, 2012, **82,** 597-606.
465 27. Kalvelage, T., Lavik, G., Jensen, M.M., Revsbech, N.P., Löscher, C., Schunck, H, *et*
466 *al.,* Aerobic microbial respiration in oceanic oxygen minimum zones. *PLoS One.*,
467 2015, **107**, e0133526.
468 28. Stevens, H. and Ulloa, O., Bacterial diversity in the oxygen minimum zone of the
469 eastern tropical South Pacific. *Environ. Microbiol.*, 2008, **10,** 1244–1259.
470 29. Glass, J.B., Kretz, C.B., Ganesh, S., Ranjan, P., Seston, S.L., Buck, K.N., *et al.*,
471 Meta-omic signatures of microbial metal and nitrogen cycling in marine oxygen
472 minimum zones. *Front. Microbiol.*, 2015, **6**, 1-13.
473 30. Carlson, C.A., Morris, R., Parsons, R., Treusch, A.H., Giovannoni, S.J. and Vergin,
474 K., Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones
475 of the northwestern Sargasso Sea. *ISME J.* 2009, **3**, 283-295.
476 31. Gonsalves, M.J., Paropkari, A. L., Fernandes, C.E.G., LokaBharathi, P.A.,
477 Krishnakumari, L., Fernando, V., *et al.*, Predominance of anaerobic bacterial
478 community over aerobic community contribute to intensify 'oxygen minimum zone'
479 in the eastern Arabian Sea. *Continental Shelf Res.* 2011, **31**, 1224-1235.
480 32. Pitcher, A., Villanueva, L., Hopmans, E.C., Schouten, S., Reichart, G.J., Sinninghe,
481 J. and Damste´, S., Niche segregation of ammonia-oxidizing archaea and anammox
482 bacteria in the Arabian Sea oxygen minimum zone. *ISME J.*, 2011, **5,** 1896–1904.
483 33. Wyman, M., Hodgson, S. and Bird C., Denitrifying Alphaproteobacteria from the
484 Arabian Sea that express nosZ, the gene encoding nitrous oxide reductase, in oxic
485 and suboxic waters. *App. 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. *PeerJ.*, 2016, **4,** e1924. doi 10.7717/peerj.1924

35. Kim, O.S., Cho, Y.J., Lee, K., Yoon, S.H., Kim, M., Na, H., *et al.*, Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species *Internat. J. Syst. Evol. Microbiol.* 2012, **62**, 716-721, doi: 10.1099/ijs.0.038075-0

36. Langille, M.G., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., *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. *Nucl. 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., de la Torre, J.R., Klotz, M.G., Urakawa, H., Pinel, N., Arp, D.J., *et al.*, *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc. Nat. Acad. Sci. USA.*, 2010, **107**, 8818-8823.

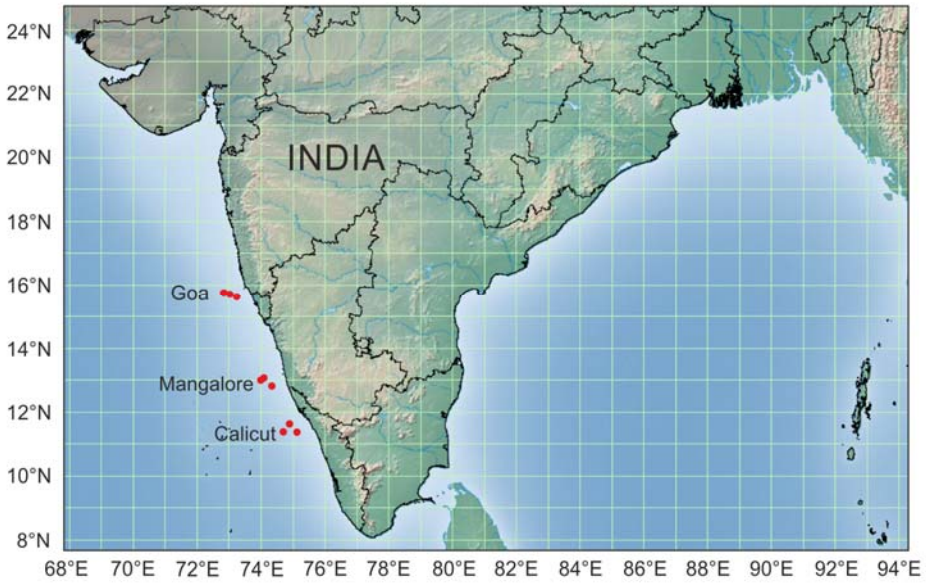
510 41. Konneke, M.E., de la Torre, J.R., Walker, C.B., Waterbury, J.B. and Stahl, D.A.
511 Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature,* 2005, **437,**
512 543-46.
513 42. Naqvi, S.W.A., Kumar, M.D., Narvekar, P.V., de Souza, S.N., George, M.D., and
514 D'Silva, C., An Intermediate nepheloid layer associated with high microbial
515 metabolic rates and denitrification in the Northwest Indian Ocean. *J. Geophys. Res.,*
516 1993, **98**, 16469-16479.
517 43. Koppelmann, R., Zimmermann-Timm, H. and Weikert, H. Bacterial and zooplankton
518 distribution in deep waters of the Arabian Sea. *Deep-Sea Res. I,* 2005, **52**, 2184–
519 2192.
520 44. LokaBharathi, P.A. and Chandramohan, D., Sulfate-reducing bacteria from the
521 Arabian Sea-their distribution in relation to thiosulfate-oxidising and hetero-trophic
522 bacteria. *Bull. Mar. Sci.*, 1990, **47**, 622–630.
523 45. LokaBharathi, P.A., Nair, S. and Chandramohan, D., Anaerobic sulfide-oxidation in
524 marine colorless sulfur-oxidizing bacteria. *J. Mar. Biotechnol.,* 1997, **52/3**, 172-177.
525 46. Jayakumar, A., O'Mullan, G.D., Naqvi, S.W.A. and Ward, B.B., Denitrifying
526 bacterial community composition changes associated with stages of denitrification in
527 oxygen minimum zones. *Microb. Ecol.*, 2009, **58,** 350–362.
528 47. Ward, B.B., Devol, A.H., Rich, J.J., Chang, B.X., Bulow, S.E., Naik, H., Pratihary,
529 A. and Jayakumar, A., Denitrification as the dominant nitrogen loss process in the
530 Arabian Sea. *Nature,* 2009, **461**, 78–81.
531 48. Lavik, G., Stuhrmann, T., Bruchert, V., Van der Plas, A., Mohrholz, V., Lam, P. *et*
532 *al.,* Detoxification of sulphidic African shelf waters by blooming chemolithotrophs.
533 *Nature,* 2009, **457**, 581-586.

534 49. Walsh, D.A., Zaikova, E., Howes, C.G., Song, Y.C., Wright, J.J., Tringe, S.G., *et al.*,
535 Metagenome of a Versatile Chemolithoautotroph from Expanding Oceanic Dead
536 Zones. *Science*, 2009, **326**, 578-582.
537 50. Canfield, D.E., Stewart, F.J., Thamdrup, B., De Brabandere, L., Dalsgaard, T.,
538 DeLong, E.F., *et al.*, A cryptic sulfur cycle in oxygen-minimum-zone waters off the
539 Chilean coast. *Science*, 2010, **330**, 1375-1378.
540 51. Wang, L.P., Wang, W.P., Lai, Q.L. and Shao, Z.Z., Gene diversity of CYP153A and
541 AlkB alkane hydroxylases in oil-degrading bacteria isolated from the Atlantic Ocean.
542 *Environ. Microbiol.*, 2010, **12**, 1230-1242.
543 52. Wang, W.P. and Shao, Z.Z., Diversity of flavin-binding monooxygenase genes
544 (almA) in marine bacteria capable of degradation long-chain alkanes. FEMS *Microb.*
545 *Ecol.*, 2012, **80,** 523–33.
546 53. Rajpathak, S.N., Banerjee, R., Mishra, P.G., Khedkar, A. M., Patil, Y.M., Joshi, S.R.
547 and Deobagkar, D.D. An exploration of microbial and associated functional diversity
548 in the OMZ and non-OMZ areas in the Bay of Bengal. *J. Biosci.*, 2018, 1-14
549 https://doi.org/10.1007/s12038-018-9781-2.
550
551
552

553 **Figure Legends**

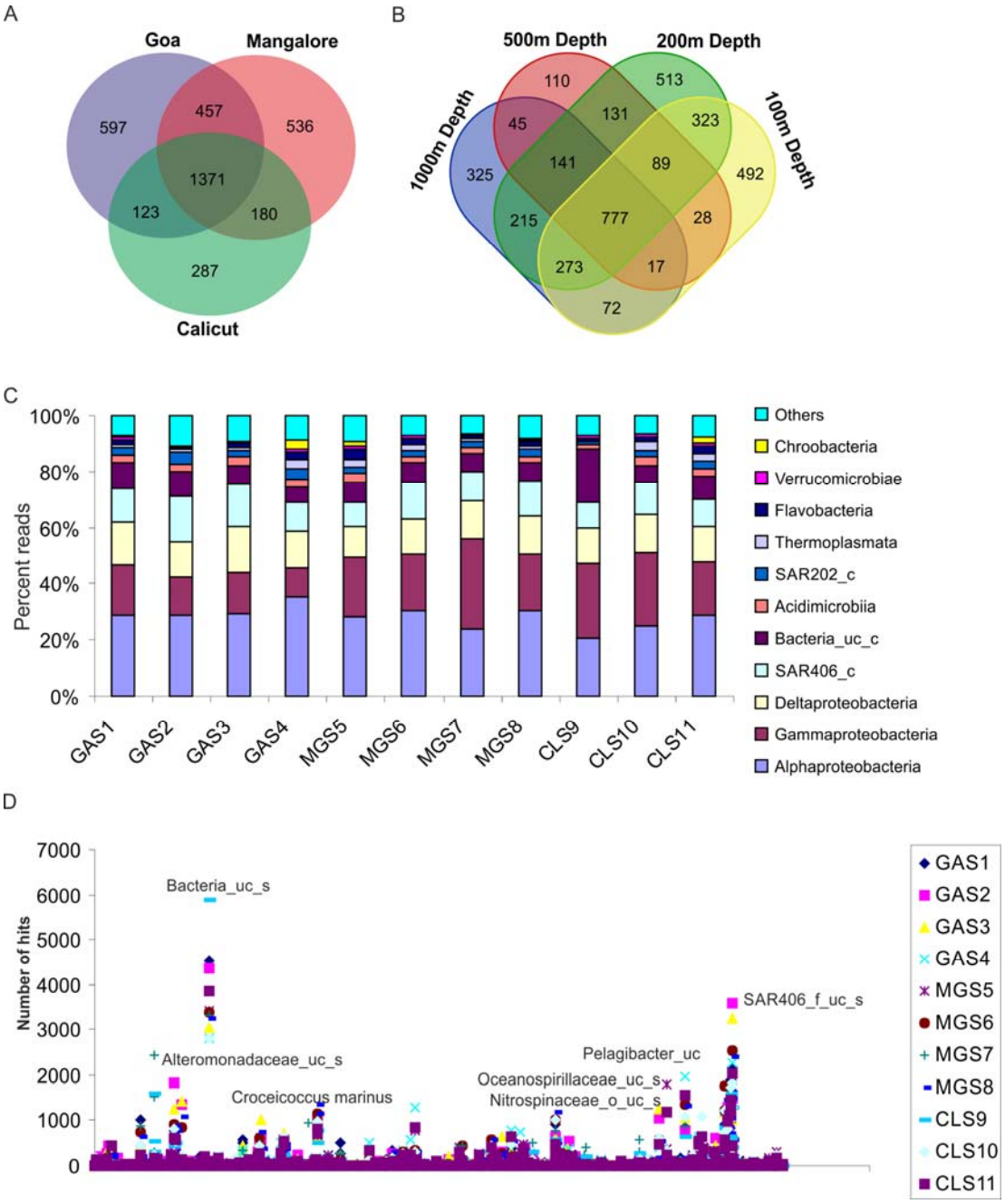
554 **Fig. 1.** Location of sampling sites in Arabian Sea. (Distance from coast is depicted in

555 Supplementary Information Table SI1)

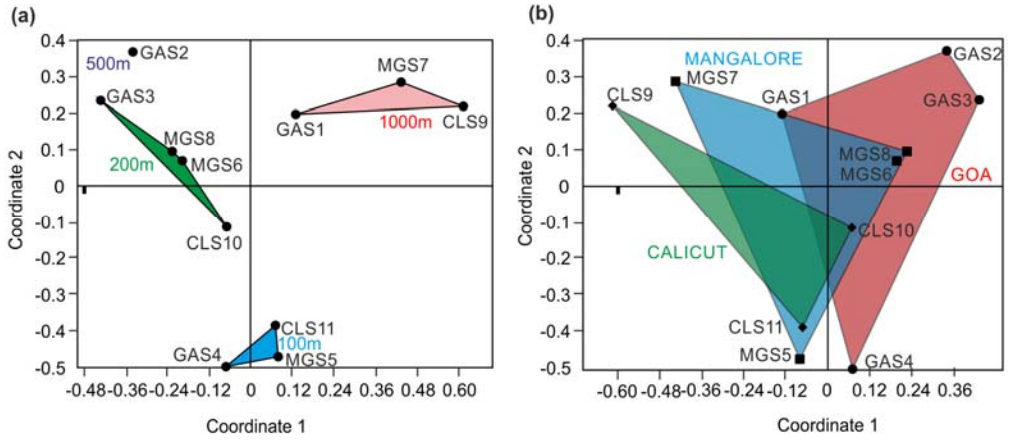


556

Fig. 2. Venn diagrams showing the number of unique and shared species between the (**A**) three sampling sites (**B**) different depth of sea. Distribution of predominant bacterial class in sample based on 16S rRNA gene sequencing. (**C**) Phyla distribution. Observations are displayed as stacked bar charts for individual mangrove sample (x-axis) against the taxa abundance (y-axis). (**D**) Abundance of species. Observations are displayed as scatter plot for individual sample (x-axis) against the species abundance (y-axis).



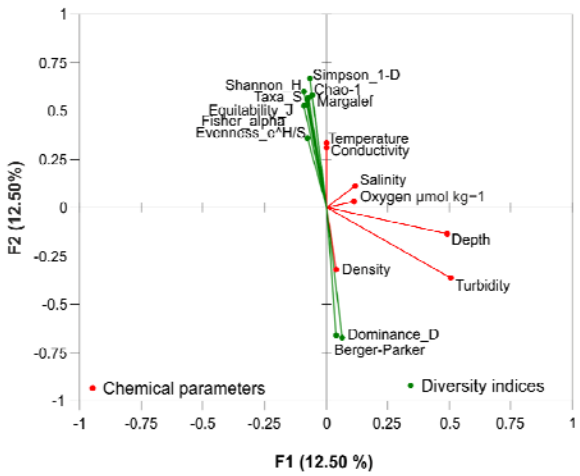
564 **Fig. 3.** Principal Coordinates Analysis (PCoA) representation of the similarity matrix
565 generated by cluster analysis. Samples from (**A**) depth and (**B**) collection site are
566 represented by a different shape, and the distance between dots represents relatedness
567 obtained from the similarity matrix.



568

569 **Fig. 4.** Association of environmental parameters and diversity indices (A) Canonical
570 correlation analysis (CCorA) (B) Mantel Test

(a)



(b)

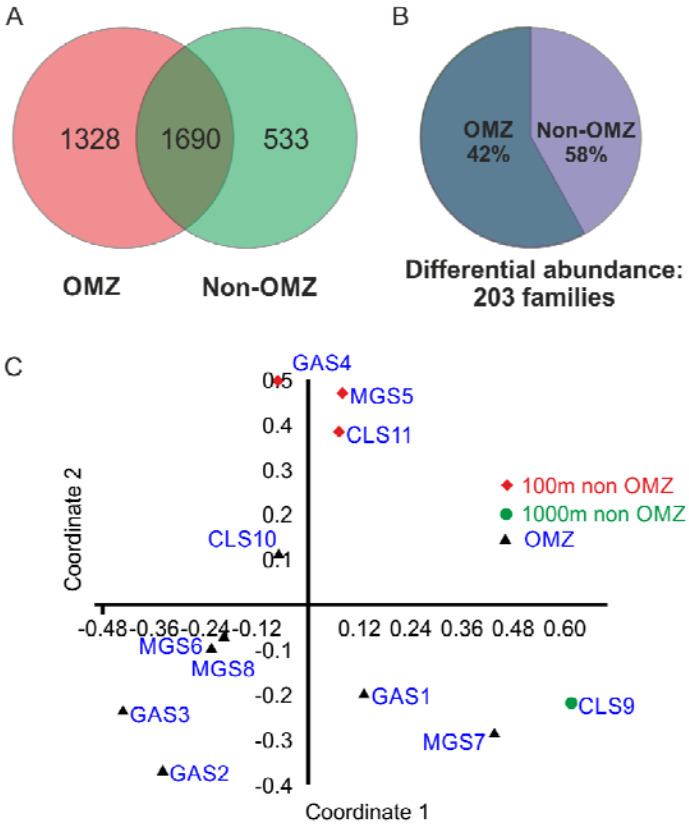
Mantel test	r	p value
Diversity Unifrac distances, chemical properties	-0.09	0.517
Alpha Diversity indices, chemical properties	-0.044	0.744
Beta Diversity index (Whitaker), chemical properties	-0.262	0.05

571

572

573 **Fig. 5.** Comparison of differentially present microbial diversity in OMZ and non-OMZ

574 areas

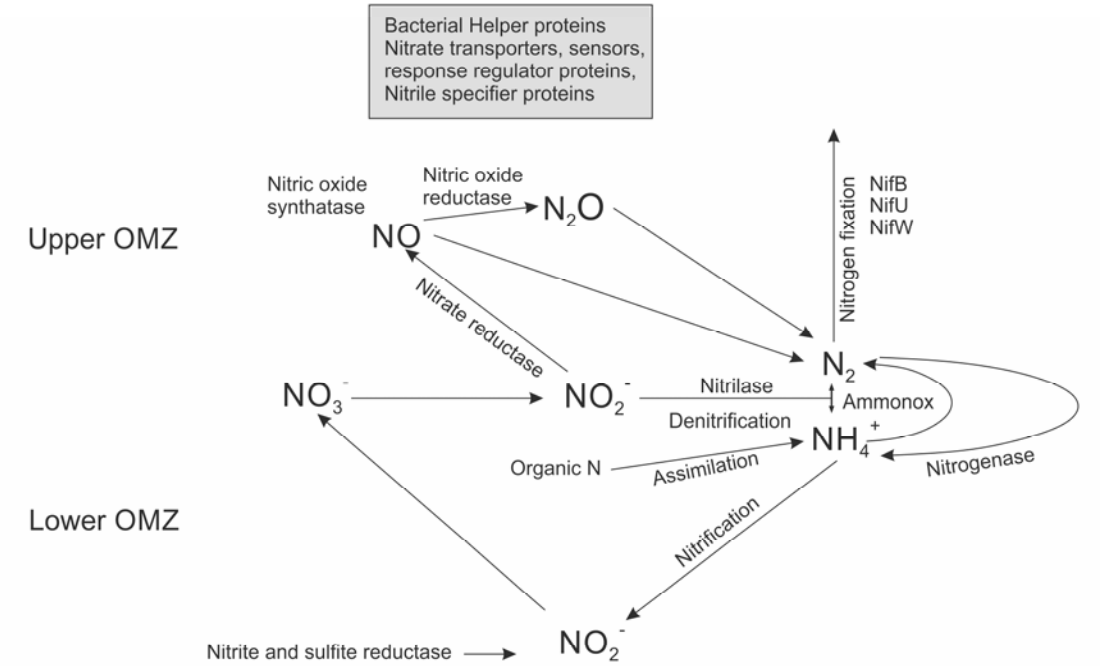


575

576

577 **Fig. 6.** Probable nitrogen cycle in oxygen depleted zones in Arabian Sea.

578



579

580 **TABLE 1. Metagenomics reads information and taxonomic affiliations of bacteria present in samples collected from Arabian Sea**

Place	Site	Valid reads	OTUs	Average Read Length	Goods Library Coverage	Bacteria			
Goa						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
Mangalore	**GAS4**	44960	34992	416.53	0.283029	285	563	897	1256
Mangalore	**MGS5**	43168	36154	421.56	0.208024	319	668	1045	1421
Mangalore	**MGS6**	47834	38732	421.25	0.242735	314	653	1052	1459
Mangalore	**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

581

582 **TABLE 2. Effect of depth and sampling location on bacterial diversity of**

583 **samples collected from Arabian Sea**

Test	**F**	***P***
Permanova		
Sampling Site	1.041	0.416
Depth	**3.503**	**0.0002**
Pair wise comparison (t test)		
Goa vs Magalore	0.9932	0.5742
Goa vs Calicut	1.294	0.2914
Calicut vs Mangalore	0.8273	0.6308
100m vs 200m	**4.952**	**0.0278**
100m vs 500m	5.166	0.2496
100m vs 1000m	3.771	0.1022
200m vs 500m	2.472	0.1949
200m vs 1000m	**3.668**	**0.0273**
500m vs 1000m	1.75	0.2585

584