1 **Phylogenomic insights into distribution and adaptation of Bdellovibrionota in** 2 **marine waters**

3 Qing-Mei Li^{1,2}, Ying-Li Zhou^{1,2}, Zhan-Fei Wei^{1,2}, Yong Wang^{1,*}

4

5 *¹Institute of Deep Sea Science and Engineering, Chinese Academy of Sciences, Sanya,*
6 *Hai Nan, P.R. of China*

7 *² University of Chinese Academy of Sciences, Beijing, China*

8

9

10

11

12

13 ***Corresponding author:**

14 Yong Wang, PhD

15 Institute of Deep Sea Science and Engineering, Chinese Academy of Sciences

16 No. 28, Luhuitou Road, Sanya, Hai Nan, P.R. of China

17 **Phone:** 086-898-88381062

18 **E-mail:** wangy@idsse.ac.cn

19

20 **Running title:** Phylogenomics of Bdellovibrionota

21

22 **Keywords:** deep sea, bacterial predator, phylogeny, metabolism

23

24 **Abstract**

25 Bdellovibrionota is composed of obligate predators that can consume some
 26 gram-negative bacteria inhabiting various environments. However, whether genomic
 27 traits influence their distribution and marine adaptation remains to be answered. In
 28 this study, we performed phylogenomics and comparative genomics studies on 82
 29 Bdellovibrionota genomes along with five metagenome-assembled genomes (MAGs)
 30 from deep sea zones. Four phylogenetic groups, Oligoflexia, Bdello-group1,
 31 Bdello-group2 and Bacteriovoracia, were revealed by constructing a phylogenetic tree,
 32 of which 53.84% of Bdello-group2 and 48.94% of Bacteriovoracia were derived from
 33 ocean. Bacteriovoracia was more prevalent in deep sea zones, whereas Bdello-group2
 34 was largely distributed in the epipelagic zone. Metabolic reconstruction indicated that
 35 genes involved in chemotaxis, flagellar (mobility), type II secretion system, ABC
 36 transporters and penicillin-binding protein were necessary for predatory lifestyle of
 37 Bdellovibrionota. Genes involved in glycerol metabolism, hydrogen peroxide (H₂O₂)
 38 degradation, cell wall recycling and peptide utilization were ubiquitously present in
 39 Bdellovibrionota genomes. Comparative genomics between marine and non-marine
 40 Bdellovibrionota demonstrated that betaine as an osmoprotectant is probably widely
 41 used by marine Bdellovibrionota, meanwhile, all the marine genomes have a number
 42 of related genes for adapting marine environment. The chitinase and chitin-binding
 43 protein encoding genes were identified for the first time in Oligoflexia, which implied
 44 that Oligoflexia may prey a wider spectrum of microbes. This study expanded our

knowledge on adaption strategies of Bdellovibrionota inhabiting deep sea and their potential usage for biological control.

Importance

Bdellovibrionota can prey gram-negative bacteria proposed as biocontrol agent. Available Bdellovibrionota genomes showed that most are from marine environment. However, vertical distribution and adaption of Bdellovibrionota in deep sea has not been reported. Our study of Bdellovibrionota revealed four groups (Oligoflexia, Bdello-group1, Bdello-group2 and Bacteriovoracia) and their distribution pattern in oceans. We also identified the genes for different phases of predation and adaptation in deep-sea environment. Moreover, Oligoflexia genomes contain more genes for carbohydrates utilization and particularly those encoding chitin-binding protein and chitinase. Our analyses of Bdellovibrionota genomes may help understand their special lifestyle and deep-sea adaptation.

Introduction

Bdellovibrionota, previously known as Bdellovibrio and like organisms (BALOs), were gram-negative bacterial predators and were isolated from various environments. Bdellovibrionota are small vibrio-shaped bacteria with a polar flagellum (1). In 1965, the first free-living BALO, *Bdellovibrio bacteriovorus*, was reported, which lead to a discovery that predation process of Bdellovibrionota includes two steps, 1) attachment and degradation of prey cell wall with glycanases

and 2) digestion of prey cellular components with exoenzyme (2). Using *Escherichia coli* as a prey model, experimental results showed that Bdellovibrionota could penetrate the host cell and then transform into bdelloplast before being released from the host cell (3). In 1979, Bdellovibrionota was found to be an agent against pathogens (4), indicating the usage of Bdellovibrionota as a new approach of biocontrol that fights against bacterial infections, particularly when MDR bacteria are increasingly concerned globally (5). MDR bacterial infections have been successfully treated with assistance of bacterial predation by Bdellovibrionota (6). The prey spectrum of Bdellovibrionota is recently expanded to gram-positive pathogenic bacteria such as *Enterococcus* (7).

The current Bdellovibrionota phylum consists of several classes according to the Genome Taxonomy Database (GTDB, release89) (8), such as Bacteriovoracia, Bdellovibrionota, Oligoflexia and others without a taxonomic assignment. They are inhabitants that can be identified in various environments, such as soil, freshwater (river, lake and spring), ocean, wastewater treatment bioreactor and sewage (9-12). In different ecological niches, Bdellovibrionota could live with a prey-dependent or prey-independent lifestyle (11). This provides a clue of metabolic flexibility of Bdellovibrionota, as well as different predatory mechanisms when invading periplasmic space or attaching to the external surface of prey bacterial cell (1, 13). As bacterial predators distributed in a wide variety of niches with diversified predatory approaches, ecological roles of Bdellovibrionota were probably underestimated.

The ocean is probably the biggest reservoir of bacteria. So far the reported

obligatory predator bacteria mainly include *Micavibrio* (a genus of α -proteobacteria) and five families belonging to δ -proteobacteria (14). Proteobacteria is a super phylum that constitutes gram negative bacteria and dominates the microbial communities in dark deep-sea zones (15). Experiments had exhibited that *Halobacteriovorax*, affiliated with one group of Bdellovibrionota, could predate deep-sea bacterium *Vibrio parahaemolyticus* (gram-negative bacterium) (16). In nutrient-poor deep sea, gram-negative bacteria may meet difficulties in synthesis biopolymers of key cellular structural components, such as peptidoglycan. Therefore Bdellovibrionota can probably prey on more gram-negative bacteria, which likely occurs ubiquitously in deep-sea environments (17). Under oligotrophic deep-sea environment, the competition for nutrient among microorganisms was constantly intensive (18). The adaptation strategy of Bdellovibrionota in marine environment is still an unresolved issue.

In this study, available Bdellovibrionota genomes were collected for phylogenomic analysis along with five representative Bdellovibrionota genomes binned from the marine metagenomes. The environmental sources of the genomes indicated that two subclades (Bacteriovoracia and Bdello-group2) were mostly present in marine environment. Bacteriovoracia were increasingly abundant in deeper ocean, while Bdello-group2 prefers to inhabit the epipelagic zone. Comparative genomics demonstrated that marine Bdellovibrionota species evolved to have genes involved in osmolyte metabolism. Oligoflexia may be a potential biocontrol agent of pathogens beyond bacteria since we discovered their chitin-binding protein and chitinase

encoding genes.

Results and discussion

Phylogenomics of Bdellovibrionota

A total of 148 genomes of Bdellovibrionota were downloaded from GTDB database (Table S1), eight genomes were obtained from the NCBI to represent the model species that have been well studied previously (19, 20). After quality control (>70% completeness, < 10% contamination and average nucleotide identity $\geq$ 98.5%) (Tables S2 and S3), 82 genomes were retained for the subsequent study. Five MAGs that represent deep-sea Bdellovibrionota were retrieved from marine water samples of the Mariana Trench, Bashi Strait and the South China Sea with depths ranging from 1,646m to 5,992m (Table 1). The five deep-sea MAGs were in size of 2.19~5.88 Mbp with more than 70% completeness and less than 3% contamination (Table 1).

Construction of phylogenomic tree using conserved proteins displayed four phylogenetic groups (subclades), which were then named as ‘Bacteriovoracia’, ‘Oligoflexia’, ‘Bdello-group1’ and ‘Bdello-group2’ (Figs. 1A and 1B). The phylogenetic relationships between the groups were consistent with those based on 16S rRNA genes (Fig. S1). The five deep-sea MAGs were distributed into three Bdellovibrionota groups except for Bdello-group1 (Figs. 1A and 1B). The available isolation sources of the samples and genomes were collected and summarized (Table S4). These genomes were mainly obtained from marine water, subsurface sediment and bioreactor sludge, waste water and ground water. Bdello-group2 and Oligoflexia

were found in more diversified environments (Fig. 1C). Marine and ground water were the most common sources for Bdellovibrionota, which coincides with their preference for environment with low viscosity (21). About 43.40% of Bacteriovoracia and 46.81% of Bdello-group2 genomes were derived from marine waters (Fig. 1C and Table S4), implying their important role in oceans. The wide spread of Bdellovibrionota may be attributable to two-phase lifestyle and quick response to the transformation between the phases (22).

Among the four groups, the size of Oligoflexia genomes varies considerably (Fig. 2A). This is true for their genomic GC contents ranging from 32.83 % to 54.35% (Fig. 2A and Table S2), which might be stemmed from the presence of many copies of transposase genes in their genomes (Fig. S2). The average GC content of Bdello-group1 genomes was highest with a mean value more than 50% (Fig. 2A and Table S2). Note that 26.67% of the Bdello-group1 genomes were obtained from sludge bioreactors (Table S4). Whether the wide range of GC contents of Bdello-group1 genomes is ascribed to frequent contacts with the diverse microbes in sludges is still elusive.

Relative abundance of Bdellovibrionota in marine water zones

To examine vertical distribution of the four Bdellovibrionota groups (Fig. 1A) in marine water zones, the 16S rRNA sequences (16S metagenomic Illumina tags, miTags) of Tara ocean (23) and Mariana marine metagenomes (24, 25) were used to calculate relative abundance of Bdellovibrionota groups. Bacteriovoracia and

Bdello-group2 were relatively abundant in oceans, compared to the other two groups (Fig. 3A). The relative abundance of Bacteriovoracia increased with water depth, whereas a reverse trend was observed for Bdello-group2. Statistical analysis showed that epipelagic layer had significantly more abundant Bdello-group2 than the other deep-sea layers, whereas significantly more abundant Bacteriovoracia were distributed in deeper layers (t-test, $p < 0.05$) (Fig. 3B). These results strongly indicate that Bacteriovoracia was relatively more prevalent in deeper marine environment. Conversely, Bdello-group2 prefers to inhabit the epipelagic zone. Oligoflexia and Bdello-group1 are associated with very low relative abundance in marine environment (Fig. 3B). Oligoflexia showed increased tendency with significance between epipelagic and mesopelagic zones (t-test, $p < 0.05$), whereas there was no significance when comparing the epipelagic and deeper layer ($\geq 1,000\text{m}$) (Fig. 3B). Those results showed that vertical distribution pattern of Bdellovibrionota differed between the groups, suggesting considerable difference in their gene profiles.

Metabolic potentials (reconstruction) of Bdellovibrionota

Discrepancies of the Bdellovibrionota groups in metabolism and environmental adaptation were detected by comparing annotations against KEGG (Kyoto Encyclopedia of Genes and Genomes), COG (Cluster of Orthologous Groups of proteins), MEROPS and CAZY (Carbohydrate Active enZymes) databases (26-29). When group-specific genes (only present in more than 50% genomes of one group) were selected from KEGG and COG annotation results, we found that Oligoflexia

genomes contained 48 KEGG genes and 34 COGs as group-specific genes, remarkably more than other groups (Fig. 3B). The peptidase-coding genes found in MEROPS annotation were compared, which showed that 118 peptidase genes were shared by four groups and each group also contained unique peptidases (Fig. 3C). Oligoflexia differs from other groups of Bdellovibrionota with higher average number of unique peptidases (Fig. 3C). Based on CAZY annotation, Oligoflexia genomes encode the obviously more enzymes of glycosyl transferase (GT) families than other groups (Fig. 3D). GT enzymes can catalyze the transfer of sugar moieties onto aglycons, which is related to transformation of many natural products (30). This indicates that Oligoflexia may heavily depend on degradation of carbohydrate from prey bacterial cells and environment.

To disentangle the metabolic and adaptive capacities of Bdellovibrionota, metabolism reconstruction was carried out and presented (Fig. 4). If at least 50% genomes of a group harbor a gene, the group was regarded to have the gene. As bacterial predators, Bdellovibrionota genomes all harbor genes that encode an almost full set of subunits involved in mobility (FlgKGI, FliL and MotA), chemotaxis (CheYVBWAR and MCP), type II secretion system and some ABC transporters (Table S5). These chemotactic proteins may work effectively and help Bdellovibrionota to sense abundance of prey bacteria and toxic chemicals. Multicopies of *mcp* and *motA* genes were predicted in the genomes (up to 27 copies of *mcp* in an Oligoflexia genome, GCA_001907975.2, renamed as ref32 in this study) (Table S6). Artificial mutations of *motA* affected integrality of flagellar (31), which

indicated that MotA is important in flagellar assembly. MCP, a methyl accepting chemotaxis protein, was reported to act as an enhancer for predatory ability of Bdellovibrionota (32). In addition, *cheB* was predicted in all the groups (71.42% of Oligoflexia genomes; all of the Bacteriovoracia genomes; 57.14% of Bdello-group1 genomes and 89.13% of Bdello-group2 genomes) (Table S5). During bacterial rapid adaption and responding to the local gradients of attractants or repellents, protein CheB, a methylesterase, functions in covalent modification of membrane receptors (33). A gene encoding a methyltransferase CheR is present in all the Bdellovibrionota groups (Table S5) (34). Methylation enhances clockwise signals of the motors, while demethylation attenuates them by means of a feedback loop that is managed by CheB and CheR with respect to methylation rate of receptors (35). This is the nature of motor transition between clockwise and counterclockwise to response the changes of environmental factors. Bdellovibrionota may hunt other bacteria with assistance of methylated modification related chemotaxis proteins (CheB, CheR and MCP) for precise control of bacterial motility.

After invasion into periplasmatic space of prey bacteria, Bdellovibrionota bacteria begin to degrade outer cellular structures and cell membrane of the prey. N-acetyl-D-glucosamine (glcNAc), glycerol and lipid of cellular components are all nutrients for Bdellovibrionota. After glcNAc is imported into a Bdellovibrionota cell, it can be directly used for producing UDP-N-acetylmuramic acid (UDP-MurNAc) for subsequent biosynthesis of peptidoglycan that is the basic unit of cell wall (Fig. 4) (36). GlcNAc may also be degraded into glycolysis pathway. Glycerol, a simple

organic molecular, can be produced by several metabolism pathways (such as glycolysis and lipid catabolism). The coding genes of glycerol kinase (GK) and glycerol-3-phosphate-dehydrogenase (GLPD) were predicted in almost all of the *Bdellovibrionota* genomes (Table S6), which indicates that *Bdellovibrionota* can utilize glycerol for glycolysis (Fig. 4). The by-product of glycerol oxidation is H_2O_2 , which is cytotoxic (37) and can be used to destroy cell wall structure of prey bacteria after free diffusion of H_2O_2 into the prey periplasmic space. The coding gene of sn-glycerol-3-phosphate acyltransferase (PlsB), the first enzyme for membrane phospholipid biosynthesis (38), was predicted only in *Oligoflexia* genomes (Table S6). This suggests that glycerol may be used for membrane phospholipid biosynthesis in rapid growth phase of *Oligoflexia* and thus the glycerol metabolism was diversified in *Bdellovibrionota*.

The metabolism reconstruction shows that there are numerous kinds of ABC transporters for uptake of phospholipid, lipoprotein, oligopeptide and branched-chain amino acids (Fig. 4). They are also useful to import nutrients from environment in free-living phase. The strategies may help *Bdellovibrionota* to adapt to oligotrophic, low temperature and hydrostatic pressure in deep sea zones.

The degradation and assimilation of nutrients are essential processes during predatory growth phase of *Bdellovibrionota*. When *Bdellovibrionota* invaded the prey bacteria and entered into growth phase, the transcription of genes increased from 33% to 85% (22). In *Bdellovibrionota*, amino acids derived from prey bacterial cell, environmental source and oligopeptide digestion were fed into different metabolism

pathways. For example, aspartate is linked with citric acid cycle and glycine may involve into homoserine metabolism (Fig. 4). We then examined whether homoserine might be transformed to acyl-homoserine lactone as the substance for quorum sensing that can further trigger type II secretion system by searching relevant genes for downstream transformation (39). However, there are no predicted acyl-HSL synthases in *Bdellovibrionota* genomes (LuxR/LuxI, AinR/AinS and other homologs to acyl-HSL synthases) (Table S6), whereas the enzymes that catalyze homoserine to threonine, glycine, serine and L-cysteine were identified (Fig. 4). These results indicate that *Bdellovibrionota* may just use homoserine as an intermediate metabolite for amino acid metabolism. Because of the special lifestyle of *Bdellovibrionota*, there may be other protein factors serving as the quorum sensing regulators during the predatory growth phase. The latest study indicated that diffusible signaling factor (DSF) considerably delayed predation and bdelloplast lysis (40). Sec-type II system is almost complete in all *Bdellovibrionota* groups (Fig. 4 and Table S5). Type II secretion system can exude numerous bacterial toxic proteins and lytic enzymes such as proteases and lipases, which is important for bacteria to obtain nutrient from prey cell membrane and environment (41). *Bdellovibrionota* may heavily depend on their type II secretion system in the whole lifecycle. During attack phase, the secreting proteins can probably be used to catch prey bacteria and partly destruct the outer membrane structure for entry; during the predatory growth phase, they are likely useful for cellular component digestion and complete prey cell lysis.

Genes involved in survival of Bdellovibrionota in marine environment

Although many Bdellovibrionota genomes were obtained from marine environment, there is not a special group exclusively consisting of marine species (Fig. 1). Given the characteristics with high hydrostatic pressure, oligotrophy and low temperature in deep-sea, there must be some genes in Bdellovibrionota for deep-sea adaptation. To reveal these genes, we divided each group of the Bdellovibrionota genomes into two parts, non-marine and marine, according to their sample sources. The marine genomes of Bdellovibrionota mostly belonged to Bacteriovoracia and Bdello-group2 (Fig. 1C and Table S4). A comparison between the two groups showed that the marine genomes have *betT/betS* (77.78% of Bacteriovoracia marine genomes and 57.89% of Bdello-group2 marine genomes) that takes part in betaine biosynthesis via choline pathway because of its high-affinity to choline (Fig. 5). Betaine is one of known major organic molecules that protect bacteria from damage of hypersaline/high hydrostatic pressure by stabilizing the natural conformation of proteins and preventing their aggregation (42). Although it is reported that choline and glycine betaine are nearly ubiquitous in bacteria (42), the percentage of *betT/betS* in the marine group is much higher than that in non-marine group (77.78% vs. 21.05% of Bacteriovoracia and 57.89% vs. 0 of Bdello-group2) in this study (Fig. 5). This indicates that *betT/betS* genes in the marine Bdellovibrionota may be critical in protecting them from the high hydrostatic pressure by importing choline/glycine/proline betaine from the sea water or the prey bacterial cells. Apart from *betT/betS*, the gene named *mscS* (Fig. 5), contributing to normal resistance to hypoosmotic shock, was predicted in

89.47% of Bdello-group2 marine genomes. The MscS channel is integrated in the cell membrane, which prevents burst of the cell in hypotonic environments (43, 44). It may work with BetT/BetS in a feedback model for maintaining and rescuing Bdellovibrionota (Bdello-group2) under the various marine chemical gradients.

The special physicochemical parameters of ocean lead to diverse resistance mechanisms of marine microbes, such as detoxification. The H₂O₂ degrading genes were predicted in Bacteriovoracia marine genomes, and were perhaps involved in prevention against self-toxification of H₂O₂ induced by the byproduct of glycerol oxidation. The genes *osmC* (77.78%), *catA* (55.56%) and *katE* (55.56%) were present in Bacteriovoracia marine genomes (Fig. 5). These genes are probably related to protect cells from the toxic effects of H₂O₂ that acts as a bactericide (45). OsmC is a protein induced by osmotic pressure (46). Interestingly, its malfunction reduced fitness and elevated sensitivity to oxidative stress of *E. coli* (47). Therefore, OsmC may participate in defense against high salinity and H₂O₂ (a defensive molecule that may also be produced by prey bacteria) in Bacteriovoracia (Fig.4). As for *catA*, it was up-regulated in *Streptomyces coelicolor* treated with H₂O₂ during non-stationary phase and it was required for *Streptomyces coelicolor* growth under aerobic condition (48, 49). The function of *katE* has not yet been determined, but it is predicted to regulate the concentration of H₂O₂ (50, 51). H₂O₂ may be scavenged and degraded in Bacteriovoracia with co-operation of OsmC, CatA and KatE for survival in deeper ocean (Figs. 3 and 5).

In Bdello-group2 marine genomes, there are some genes involved in cell wall

degradation, recycling, organization and biogenesis (Fig. 5). Penicillin-binding proteins are important for Bdellovibrionota to break up prey cell wall (52). MltA encoded by 57.89% of Bdello-group2 marine genomes is a murein degrading enzyme that is located in outer-membrane and can interact with penicillin-binding protein 2 (53, 54), which is important for marine Bdello-group2 as the predator of other gram-negative bacteria. *MreD* gene was present in 57.89% of Bdello-group2 marine genomes (Fig. 5) and is essential for lateral peptidoglycan synthesis (55). We identified *anmK* gene in 52.63% of Bdello-group2 marine genomes (Fig. 5). The enzyme encoded by *anmK* plays a role in cell wall recycling though transforming 1,6-anhydro-N-acetylmuramic acid (anhMurNAc) to N-acetylglucosamine-phosphate (GlcNAc-P) (56). Usually, the source of anhMurNAc is derived from its own cell wall murein, but as predators, marine Bdello-group2 may reuse the component of prey bacteria for cell wall formation of next generation. ECM4 found in 52.63% of Bdello-group2 marine genomes (Fig. 5) was proposed as a protein involved in cell wall biogenesis (integrity) (57). Under deep-sea oligotrophic condition, these genes may be all associated with the whole process from murein-degrading, anhMurNAc recycling and cell wall biogenesis for proliferation of Bdello-group2 marine bacteria.

There are two genomes of Bdello-group1 (GCA_002722705.1 and GCA_002450715.1) and only one genome of Oligoflexia (R3B4) from marine. The discrepancies in genes between marine and non-marine groups were also demonstrated (Fig. S2). The genes involved in synthesis of osmoprotectants were also predicted in marine Bdello-group1 and Oligoflexia genomes, such as *betT* and *mscS* in

Bdello-group1 and *cmo* in Oligoflexia (Fig. S2). CMO can catalyze the first step of glycine betaine synthesis (58), which indicates that glycine betaine can be used as an osmoprotectant in marine Oligoflexia. In addition, there are many genes involved in transport or transformation of sugar as osmoprotectant, such as *malK*, *msmX*, *rgpF* and *algI* genes in Oligoflexia (Fig. S2) (59-62). Combining these results, betaine as an osmoprotectant is commonly used among different groups of Bdellovibrionota, all having a variety of genes for fitness in ocean.

Oligoflexia as potential defender against eukaryotic pathogens

In light of the comparison among Bdellovibrionota groups, Oligoflexia seems to be special in gene content. The chitinase gene (K01183) has many copies in Oligoflexia genomes (Fig. S3), but none in other Bdellovibrionota groups. The presence of genes encoding chitinase and chitin-binding protein (Fig. S3) indicates that some Oligoflexia may be able to degrade chitin (63, 64). The phylogenetic tree of the seven chitinases extracted from the Oligoflexia genomes and other species supports their close relationships with known homologs (Fig. 6). To further verify the predicted open reading frame (ORF) of the chitinases, multiple alignment was performed to validate the conserved motif 'DXXDXDXE' of chitinases (Fig. 6) (65). In addition, the Oligoflexia genomes contained a gene coding for chitin-binding protein (Fig. S4 and Table S8). These results imply that a part of Oligoflexia can utilize chitin in environment as nutrient and/or attack preys with chitin ultracellular structure. In this study, the chitinase and chitin-binding protein coding genes were

focused for the first time in Oligoflexia. Pathogenic fungi have chitin as a main cell wall component (66), which allows to speculate that Oligoflexia may be used to defend against eukaryotic pathogens aside from bacterial pathogens.

Conclusion

In this study, Bdellovibrionota was revealed to have four phylogenetic groups, all containing the genes in most of their genomes for two-phase predatory lifestyle. Their vertical distribution in ocean was exhibited. Bacteriovoracia was more prevalent in the deep-sea zones. Genomics traits demonstrated that marine Bdellovibrionota are able to metabolize osmoprotectant and degrade cell wall. The chitinase and chitin-binding protein encoding genes were for the first time focused in Oligoflexia. However, the function and regulation of these genes involved in pathogenic prevention need to be verified by experiments.

Materials and methods

Genome collection from public databases

We recruited 148 Bdellovibrionota genomes from GTDB database (26). Eight genomes for model Bdellovibrionota were downloaded from the NCBI. All the genomes were reclassified with GTDB-TK (v.1.0.1). Quality control of all the genomes was conducted using CheckM (v.1.1.2) (67). Mediate-quality genomes used for further analysis were selected with the following standards: 1) Completeness score was more than 70%; 2) Contamination rate was less than 10%; 3) The number of

conserved proteins derived from CheckM program was at least 22. Average Nucleotide Identity (ANI) analysis was executed by the pyani (v.0.2.09) (software available from <http://widdowquinn.github.io/pyani/>) to cluster the genomes with $\geq$ 98.5% identity (68). The sources of genomes and metagenomes were obtained from NCBI Biosample and Bioproject.

Samples collection, DNA extraction and sequencing

About 20L marine water samples were *in situ* filtered from the Mariana Trench during *R/V* TS09 by ISMIFF apparatus and sampling procedure had been described in our previous report (69). 20L water samples were also collected at different depths by Niskin bottles from the South China Sea during *R/V* TAN KAH KEE cruise and *R/V* Haigong623 equipped with remotely operated vehicle (ROV) Haixing6000 (Table 1). 0.22 μ m polycarbonate membranes (Merck Millipore, Bedford, MA) were used for microbial filtration and were then frozen at -80°C degree immediately until use. The polycarbonate membranes were cut into small pieces for DNA extraction using MO BIO Power Soil DNA Isolation Kit (MoBio, Carlsbad, CA, USA) according to the manufacturer's instruction. The quality and quantity of the DNA extraction were evaluated by 1% agarose gel electrophoresis and Qubit 3.10 Fluorometer (Invitrogen, Life, USA). The good-quality DNA was first sheared randomly to fragments of around 500bp by Focused-ultrasonicator (Covaris M220) and used for metagenomic library preparation with TruSeq® Nano DNA LT Sample Prep Kit (Illumina, San Diego, CA, USA). The high-throughput sequencing was performed on an Illumina

387 Miseq platform ($2 \times 300\text{bp}$).

388 **Raw data processing, de novo assembly and metagenomes binning**

389 The quality of metagenome raw data was assessed by FastQC (v0.11.8) (70).
 390 After Fastp (v0.20.0) (71) treatment, repeated sequences produced by sequencing
 391 platform were removed by Fastuniq (72) and then the clean reads were assembled
 392 with MEGAHIT (v1.2.5-beta) (73). The contigs $>2,000$ bp were selected for genome
 393 binning with MetaWRAP (v1.2) (models including metabat2, maxbin2 and concoct)
 394 (74). The MAGs were then taxonomically classified by gtdb-tk classify (8) and
 395 Bdellovibrionota MAGs were selected by the cut-off described above ($>70\%$
 396 completeness and $<10\%$ contamination).

397 **Calculation of Bdellovibrionota relative abundance in deep sea**

398 16S rRNA sequences of Bdellovibrionota genomes and MAGs were extracted by
 399 using rRNA_hmm_fs_wst (v.0) (75) and used to create a 16S rRNA database. The
 400 Bdellovibrionota 16S miTags in Tara Ocean (23) and Mariana (24, 69) marine water
 401 metagenomes (Table S7) were identified by blastn (v.2.9.0) (-evalue $1e-05$) against the
 402 16S rRNA sequence database of Bdellovibrionota. Only the 16S miTags longer than
 403 100bp and $>97\%$ identity to a reference in the database were selected for further
 404 calculation of their relative abundance in the marine samples. The Tara Ocean data
 405 were downloaded from <http://ocean-microbiome.embl.de/companion.html>. T-test was
 406 performed for significance analysis.

Genome annotation and metabolic reconstruction

The high-quality genomes from public databases (>90% completeness and <5% contaminant) and our five MAGs were used for gene annotation. ORFs were predicted by Prodigal (v.2.6.3) (76) and were then annotated by Kofamscan (77) (v. 1.0.0; -f mapper) against KEGG database (release 92) and by BLASTP (v.2.9.0) against COG database (COG_2019_v11.0). Carbohydrate related enzymes were searched with hmmscan (v.3.2.1) against CAZy database (dbCAN-HMMdb-V7). Peptidases were identified by BLASTP against MEROPS database (28) (pepunit.lib; -id 50; -e 1e-10). Wilcoxon test was conducted by R package to identify the genes significantly different in abundance between different groups.

Phylogenetic analysis of conserved proteins and 16S rRNA genes

The marker genes of genomes were aligned using hmmerAlign.py (called by CheckM). The alignment of 43 concatenated conserved proteins was used for reconstruction of phylogenomic tree with RAxML (78) (v.8.2.12; -# 1000; -m PROTGAMMALG) after optimization with trimAl (79) (v.1.4). All the 16S rRNA sequences extracted by rRNA_hmm (75) were treated with CD-HIT (80) (v.4.8.1; -c 0.97; -n 10; -G 1; -M 10000) to remove the same species. Two 16S rRNA sequences of *E. coli* were added as an out-group. All 16S rRNA sequences were aligned with MAFFT (81) (v.7.407) and optimized with trimAl. RAxML (v.8.2.12; GTRCAT model) was used to build a 16S rRNA phylogenetic tree with bootstrap values based on 1000 replicates. All phylogenetic trees were visualized with iTOL

(<https://itol.embl.de/>).

Acknowledgements

We are grateful to the team members aboard the *R/V* TS09 for their invaluable efforts in the sampling cruise. Great thanks are given to J. Li, J. Chen, S.X. Wang, Y.Z. Xin and D.S. Cai for their ISMIFF sampling. This study was supported by Key Research and Development Program of Hainan Province (E050010406) and the National Key Research and Development Program of China (2018YFC0310005 and 2016YFC0302501).

References

1. Stolp H, Starr M. 1963. *Bdellovibrio bacteriovorus* gen. et sp. n., a predatory, ectoparasitic, and bacteriolytic microorganism. *Antonie Van Leeuwenhoe* 29:217-248.
2. Shilo M, Bruff B. 1965. Lysis of Gram-Negative Bacteria by Host-Independent Ectoparasitic *Bdellovibrio Bacteriovorus* Isolates. *Journal of General Microbiology* 40:317-&.
3. Thomashow MF, Rittenberg SC. 1978. Intraperiplasmic Growth of *Bdellovibrio-Bacteriovorus* 109j - Solubilization of *Escherichia-Coli* Peptidoglycan. *Journal of Bacteriology* 135:998-1007.
4. Varon M. 1979. Selection of Predation-Resistant Bacteria in Continuous Culture. *Nature* 277:386-388.
5. Kadouri DE, To K, Shanks RMQ, Doi Y. 2013. Predatory Bacteria: A Potential Ally against Multidrug-Resistant Gram-Negative Pathogens. *Plos One* 8.
6. Shatzkes K, Singleton E, Tang C, Zuena M, Shukla S, Gupta S, Dharani S, Onyile O, Rinaggio J, Connell ND, Kadouri DE. 2016. Predatory Bacteria Attenuate *Klebsiella pneumoniae* Burden in Rat Lungs. *Mbio* 7.
7. Cai J, Cao Q. 2020. New *Bdellovibrio* species mutant strain BDE-1F2 used for lysing pathogenic and/or potentially pathogenic bacteria associated with aquaculture, preferably Gram-negative and/or positive pathogenic bacteria e.g. *Enterococcus*. *UNIV SOUTH CHINA TECHNOLOGY* CN111363689-A.'
8. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. 2020. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 36:1925-1927.
9. Jurkevitch E, Minz D, Ramati B, Barel G. 2000. Prey range characterization, ribotyping, and diversity of soil and rhizosphere *Bdellovibrio* spp. isolated on phytopathogenic bacteria. *Applied and Environmental Microbiology* 66:2365-2371.

- 460 10. Sockett RE. 2009. Predatory Lifestyle of *Bdellovibrio bacteriovorus*. Annual Review of
461 Microbiology 63:523-539.
- 462 11. Hobley L, Lerner TR, Williams LE, Lambert C, Till R, Milner DS, Basford SM, Capeness MJ,
463 Fenton AK, Atterbury RJ, Harris MATS, Sockett RE. 2012. Genome analysis of a
464 simultaneously predatory and prey-independent, novel *Bdellovibrio bacteriovorus* from the
465 River Tiber, supports in silico predictions of both ancient and recent lateral gene transfer from
466 diverse bacteria. BMC Genomics 13.
- 467 12. Paix B, Ezzedine JA, Jacquet S. 2019. Diversity, Dynamics, and Distribution of *Bdellovibrio*
468 and Like Organisms in Perialpine Lakes. Applied and Environmental Microbiology 85.
- 469 13. Koval SF, Hynes SH, Flannagan RS, Pasternak Z, Davidov Y, Jurkevitch E. 2013.
470 *Bdellovibrio exovorus* sp nov., a novel predator of *Caulobacter crescentus*. International
471 Journal of Systematic and Evolutionary Microbiology 63:146-151.
- 472 14. Bratanis E, Andersson T, Lood R, Bukowska-Faniband E. 2020. Biotechnological Potential of
473 *Bdellovibrio* and Like Organisms and Their Secreted Enzymes. Frontiers in Microbiology 11.
- 474 15. Zhou ZC, Tran PQ, Kieft K, Anantharaman K. 2020. Genome diversification in globally
475 distributed novel marine Proteobacteria is linked to environmental adaptation. ISME Journal
476 14:2060-2077.
- 477 16. Williams HN, Lymperopoulou DS, Athar R, Chauhan A, Dickerson TL, Chen H, Laws E,
478 Berhane TK, Flowers AR, Bradley N, Young S, Blackwood D, Murray J, Mustapha O,
479 Blackwell C, Tung Y, Noble RT. 2016. *Halobacteriovorax*, an underestimated predator on
480 bacteria: potential impact relative to viruses on bacterial mortality. ISME Journal 10:491-499.
- 481 17. Nazarenko EL, Crawford RJ, Ivanova EP. 2011. The Structural Diversity of Carbohydrate
482 Antigens of Selected Gram-Negative Marine Bacteria. Marine Drugs 9:1914-1954.
- 483 18. Hibbing ME, Fuqua C, Parsek MR, Peterson SB. 2010. Bacterial competition: surviving and
484 thriving in the microbial jungle. Nature Reviews Microbiology 8:15-25.
- 485 19. Rodionova IA, Heidari Tajabadi F, Zhang Z, Rodionov DA, Saier MH, Jr. 2019. A Riboflavin
486 Transporter in *Bdellovibrio exovorus* JSS. J Mol Microbiol Biotechnol 29:27-34.
- 487 20. Williams LE, Cullen N, DeGiorgis JA, Martinez KJ, Mellone J, Oser M, Wang J, Zhang Y.
488 2019. Variation in genome content and predatory phenotypes between *Bdellovibrio* sp. NC01
489 isolated from soil and *B. bacteriovorus* type strain HD100. Microbiology-Sgm
490 165:1315-1330.
- 491 21. Im H, Kwon H, Cho G, Kwon J, Choi SY, Mitchell RJ. 2019. Viscosity has dichotomous
492 effects on *Bdellovibrio bacteriovorus* HD100 predation. Environmental Microbiology
493 21:4675-4684.
- 494 22. Karunker I, Rotem O, Dori-Bachash M, Jurkevitch E, Sorek R. 2013. A Global Transcriptional
495 Switch between the Attack and Growth Forms of *Bdellovibrio bacteriovorus*. Plos One 8.
- 496 23. Moran MA. 2015. The global ocean microbiome. Science 350:aac8455.
- 497 24. Li WL, Huang JM, Zhang PW, Cui GJ, Wei ZF, Wu YZ, Gao ZM, Han Z, Wang Y. 2019.
498 Periodic and Spatial Spreading of Alkanes and Alcanivorax Bacteria in Deep Waters of the
499 Mariana Trench. Applied and Environmental Microbiology 85.
- 500 25. Gao ZM, Huang JM, Cui GJ, Li WL, Li J, Wei ZF, Chen J, Xin YZ, Cai DS, Zhang AQ, Wang
501 Y. 2019. In situ meta-omic insights into the community compositions and ecological roles of
502 hadal microbes in the Mariana Trench. Environmental Microbiology 21:4092-4108.
- 503 26. Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic

504 Acids Research 28:27-30.

505 27. Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a tool for
506 genome-scale analysis of protein functions and evolution. *Nucleic Acids Research* 28:33-36.

507 28. Rawlings ND, Alan J, Thomas PD, Huang XD, Bateman A, Finn RD. 2018. The MEROPS
508 database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with
509 peptidases in the PANTHER database. *Nucleic Acids Research* 46:D624-D632.

510 29. Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active
511 enzymes database (CAZy) in 2013. *Nucleic Acids Research* 42:D490-D495.

512 30. Schmid J, Heider D, Wendel NJ, Sperl N, Sieber V. 2016. Bacterial Glycosyltransferases:
513 Challenges and Opportunities of a Highly Diverse Enzyme Class Toward Tailoring Natural
514 Products. *Frontiers in Microbiology* 7.

515 31. Suzuki Y, Morimoto YV, Oono K, Hayashi F, Oosawa K, Kudo S, Nakamura S. 2019. Effect
516 of the MotA(M206I) Mutation on Torque Generation and Stator Assembly in the Salmonella
517 H⁺-Driven Flagellar Motor. *Journal of Bacteriology* 201.

518 32. Lambert C, Smith MC, Sockett RE. 2003. A novel assay to monitor predator-prey interactions
519 for *Bdellovibrio bacteriovorus* 109 J reveals a role for methyl-accepting chemotaxis proteins
520 in predation. *Environ Microbiol* 5:127-32.

521 33. Kehry MR, Bond MW, Hunkapiller MW, Dahlquist FW. 1983. Enzymatic damidation of
522 methyl-accepting chemotaxis proteins in *Escherichia coli* catalyzed by the cheB gene product.
523 *Proc Natl Acad Sci USA* 80:3599-3603.

524 34. Batra M, Sharma R, Chandra V, Aggarwal M, Agarwal U, Gupta P, Singh RP, Tomar S. 2015.
525 In silico and proteomic analysis of protein methyltransferase CheR from *Bacillus subtilis*.
526 *International Journal of Biological Macromolecules* 77:168-180.

527 35. Hamadeh A, Roberts MAJ, August E, McSharry PE, Maini PK, Armitage JP,
528 Papachristodoulou A. 2011. Feedback Control Architecture and the Bacterial Chemotaxis
529 Network. *Plos Computational Biology* 7.

530 36. Typas A, Banzhaf M, Gross CA, Vollmer W. 2012. From the regulation of peptidoglycan
531 synthesis to bacterial growth and morphology. *Nature Reviews Microbiology* 10:123-136.

532 37. Wang Y, Huang JM, Zhou YL, Almeida A, Finn RD, Danchin A, He LS. 2020. Phylogenomics
533 of expanding uncultured environmental Tenericutes provides insights into their pathogenicity
534 and evolutionary relationship with Bacilli. *Bmc Genomics* 21.

535 38. Feng YJ, Cronan JE. 2011. The *Vibrio cholerae* fatty acid regulatory protein, FadR, represses
536 transcription of *plsB*, the gene encoding the first enzyme of membrane phospholipid
537 biosynthesis. *Molecular Microbiology* 81:1020-1033.

538 39. Schuster M, Sexton DJ, Diggle SP, Greenberg EP. 2013. Acyl-Homoserine Lactone Quorum
539 Sensing: From Evolution to Application. *Annual Review of Microbiology*, Vol 67 67:43-63.

540 40. Dwidar M, Jang H, Sangwan N, Mun W, Im H, Yoon S, Choi S, Nam D, Mitchell RJ. 2020.
541 Diffusible Signaling Factor, a Quorum-Sensing Molecule, Interferes with and Is Toxic
542 Towards *Bdellovibrio bacteriovorus* 109J. *Microb Ecol* doi:10.1007/s00248-020-01585-8.

543 41. Korotkov KV, Sandkvist M, Hol WGJ. 2012. The type II secretion system: biogenesis,
544 molecular architecture and mechanism. *Nature Reviews Microbiology* 10:336-351.

545 42. Nieder JB, Hussels M, Bittl R, Brecht M. 2014. Effect of TMAO and betaine on the energy
546 landscape of photosystem I. *Biochimica Et Biophysica Acta-Bioenergetics* 1837:849-856.

547 43. Booth IR, Blount P. 2012. The MscS and MscL Families of Mechanosensitive Channels Act as

548 Microbial Emergency Release Valves. *Journal of Bacteriology* 194:4802-4809.

549 44. Bialecka-Fornal M, Lee HJ, Phillips R. 2015. The Rate of Osmotic Downshock Determines
550 the Survival Probability of Bacterial Mechanosensitive Channel Mutants. *Journal of*
551 *Bacteriology* 197:231-237.

552 45. Armoza-Zvuloni R, Schneider A, Shaked Y. 2016. Rapid Hydrogen Peroxide Release during
553 Coral-Bacteria Interactions. *Frontiers in Marine Science* 3.

554 46. Gutierrez C, Devedjian JC. 1991. Osmotic Induction of Gene Osmc Expression in
555 *Escherichia-Coli* K12. *Journal of Molecular Biology* 220:959-973.

556 47. Conter A, Gangneux C, Suzanne M, Gutierrez C. 2001. Survival of *Escherichia coli* during
557 long-term starvation: effects of aeration, NaCl, and the rpoS and osmC gene products. *Res*
558 *Microbiol* 152:17-26.

559 48. Cho YH, Roe JH. 1997. Isolation and expression of the catA gene encoding the major
560 vegetative catalase in *Streptomyces coelicolor* Muller. *J Bacteriol* 179:4049-52.

561 49. Hahn JS, Oh SY, Chater KF, Cho YH, Roe JH. 2000. H₂O₂-sensitive fur-like repressor CatR
562 regulating the major catalase gene in *Streptomyces coelicolor*. *J Biol Chem* 275:38254-60.

563 50. Sun D, Crowell SA, Harding CM, De Silva PM, Harrison A, Fernando DM, Mason KM,
564 Santana E, Loewen PC, Kumar A, Liu Y. 2016. KatG and KatE confer *Acinetobacter*
565 resistance to hydrogen peroxide but sensitize bacteria to killing by phagocytic respiratory
566 burst. *Life Sci* 148:31-40.

567 51. Rochat T, Miyoshi A, Gratadoux JJ, Duwat P, Sourice S, Azevedo V, Langella P. 2005.
568 High-level resistance to oxidative stress in *Lactococcus lactis* conferred by *Bacillus subtilis*
569 catalase KatE. *Microbiology (Reading)* 151:3011-3018.

570 52. Lerner TR, Lovering AL, Bui NK, Uchida K, Aizawa S, Vollmer W, Sockett RE. 2012.
571 Specialized peptidoglycan hydrolases sculpt the intra-bacterial niche of predatory *Bdellovibrio*
572 and increase population fitness. *PLoS Pathog* 8:e1002524.

573 53. Lommatzsch J, Templin MF, Kraft AR, Vollmer W, Holtje JV. 1997. Outer membrane
574 localization of murein hydrolases: MltA, a third lipoprotein lytic transglycosylase in
575 *Escherichia coli*. *J Bacteriol* 179:5465-70.

576 54. Jennings GT, Savino S, Marchetti E, Arico B, Kast T, Baldi L, Ursinus A, Holtje JV, Nicholas
577 RA, Rappuoli R, Grandi G. 2002. GNA33 from *Neisseria meningitidis* serogroup B encodes a
578 membrane-bound lytic transglycosylase (MltA). *Eur J Biochem* 269:3722-31.

579 55. White CL, Kitich A, Gober JW. 2010. Positioning cell wall synthetic complexes by the
580 bacterial morphogenetic proteins MreB and MreD. *Mol Microbiol* 76:616-33.

581 56. Uehara T, Suefuji K, Valbuena N, Meehan B, Donegan M, Park JT. 2005. Recycling of the
582 anhydro-N-acetylmuramic acid derived from cell wall murein involves a two-step conversion
583 to N-acetylglucosamine-phosphate. *Journal of Bacteriology* 187:3643-3649.

584 57. Teixeira MC, Fernandes AR, Mira NP, Becker JD, Sa-Correia I. 2006. Early transcriptional
585 response of *Saccharomyces cerevisiae* to stress imposed by the herbicide
586 2,4-dichlorophenoxyacetic acid. *Fems Yeast Research* 6:230-248.

587 58. Luo D, Niu X, Yu J, Yan J, Gou X, Lu BR, Liu Y. 2012. Rice choline monooxygenase
588 (OsCMO) protein functions in enhancing glycine betaine biosynthesis in transgenic tobacco
589 but does not accumulate in rice (*Oryza sativa* L. ssp. japonica). *Plant Cell Rep* 31:1625-35.

590 59. Diederichs K, Diez J, Greller G, Muller C, Breed J, Schnell C, Vornrhein C, Boos W, Welte W.
591 2000. Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter

592 of the archaeon *Thermococcus litoralis*. *Embo Journal* 19:5951-5961.

593 60. Ferreira MJ, de Sa-Nogueira I. 2010. A Multitask ATPase Serving Different ABC-Type Sugar
594 Importers in *Bacillus subtilis*. *Journal of Bacteriology* 192:5312-5318.

595 61. Kovacs CJ, Faustoferri RC, Quivey RG. 2017. RgpF Is Required for Maintenance of Stress
596 Tolerance and Virulence in *Streptococcus mutans*. *Journal of Bacteriology* 199.

597 62. Jang CH, Piao YL, Huang XQ, Yoon EJ, Park SH, Lee K, Zhan CG, Cho H. 2016. Modeling
598 and Re-Engineering of *Azotobacter vinelandii* Alginate Lyase to Enhance Its Catalytic
599 Efficiency for Accelerating Biofilm Degradation. *Plos One* 11.

600 63. Frederiksen RF, Paspaliari DK, Larsen T, Storgaard BG, Larsen MH, Ingmer H, Palcic MM,
601 Leisner JJ. 2013. Bacterial chitinases and chitin-binding proteins as virulence factors.
602 *Microbiology-Sgm* 159:833-847.

603 64. Tran HT, Barnich N, Mizoguchi E. 2011. Potential role of chitinases and chitin-binding
604 proteins in host-microbial interactions during the development of intestinal inflammation.
605 *Histology and Histopathology* 26:1453-1464.

606 65. Tyler L, Bragg JN, Wu JJ, Yang XH, Tuskan GA, Vogel JP. 2010. Annotation and comparative
607 analysis of the glycoside hydrolase genes in *Brachypodium distachyon*. *Bmc Genomics* 11.

608 66. Lenardon MD, Munro CA, Gow NA. 2010. Chitin synthesis and fungal pathogenesis. *Curr*
609 *Opin Microbiol* 13:416-23.

610 67. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing
611 the quality of microbial genomes recovered from isolates, single cells, and metagenomes.
612 *Genome Res* 25:1043-55.

613 68. Tully BJ. 2019. Metabolic diversity within the globally abundant Marine Group II
614 Euryarchaea offers insight into ecological patterns. *Nature Communications* 10.

615 69. Wang Y, Gao ZM, Li J, He LS, Cui GJ, Li WL, Chen J, Xin YZ, Cai DS, Zhang AQ. 2019.
616 Hadal water sampling by in situ microbial filtration and fixation (ISMIFF) apparatus.
617 *Deep-Sea Research Part I-Oceanographic Research Papers* 144:132-137.

618 70. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data.
619 Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.

620 71. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor.
621 *Bioinformatics* 34:i884-i890.

622 72. Xu H, Luo X, Qian J, Pang X, Song J, Qian G, Chen J, Chen S. 2012. FastUniq: a fast de novo
623 duplicates removal tool for paired short reads. *PLoS One* 7:e52249.

624 73. Li D, Liu CM, Luo R, Sadakane K, Lam TW. 2015. MEGAHIT: an ultra-fast single-node
625 solution for large and complex metagenomics assembly via succinct de Bruijn graph.
626 *Bioinformatics* 31:1674-6.

627 74. Uritskiy GV, DiRuggiero J, Taylor J. 2018. MetaWRAP-a flexible pipeline for
628 genome-resolved metagenomic data analysis. *Microbiome* 6:158.

629 75. Huang Y, Gilna P, Li W. 2009. Identification of ribosomal RNA genes in metagenomic
630 fragments. *Bioinformatics* 25:1338-40.

631 76. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal:
632 prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*
633 11:119.

634 77. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, Ogata H. 2020.
635 KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score

636 threshold. *Bioinformatics* 36:2251-2252.

637 78. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
638 large phylogenies. *Bioinformatics* 30:1312-3.

639 79. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. 2009. trimAl: a tool for automated
640 alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972-3.

641 80. Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation
642 sequencing data. *Bioinformatics* 28:3150-2.

643 81. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7:
644 improvements in performance and usability. *Mol Biol Evol* 30:772-80.

645

646

647

648 **Table 1 Information of metagenome-assembled genomes of marine**

649 **Bdellovibrionota.**

Bin_id	K1B34	M27B12	M27B96	R3B4	R4B59
Depth (m)	1,000	3,000~5,000	3,000~5,000	1,646	5,992
Sampling site (E; N)	116.47°; 18.01°	142.01°; 11.16°	142.01°; 11.16 °	120.03°; 21.87°	123.31°; 22.82°
Com. (%)	75.16	98.21	80.15	81.83	71.94
Con. (%)	0.00	0.89	2.98	2.52	0.89
N50 (bp)	6,396	295,316	5,991	7,831	3,968
No. contigs	597	24	684	923	570
Genome size (bp)	3,229,602	4,849,992	3,554,382	5,883,831	2,188,444
No. CDSs	3,741	4,431	4,172	5,667	2,566

650 * R and K in Bin_id represent the South China Sea; M denotes the Mariana Trench.

651 Com., completeness; Con., contamination;

652 N50, the shortest contig length required to cover 50% of the genome. CDS, coding

653 sequence.

Figure 1. Phylogenomic analysis and distribution of Bdellovibrionota.

(A) Unrooted maximum-likelihood (ML) phylogenomic tree of Bdellovibrionota. The ML tree was constructed based on 43 concatenated conserved proteins with the PROTGAMMALG model. The MAGs from this study were marked with a dot. (B) Rooted ML phylogenomic tree of Bdellovibrionota. The tree was built up using three *E. coli* genomes as a outgroup. The genomes with a dot were marine MAGs from this study and the references were shown as their GTDB accession numbers or species names. (C) Environmental source of Bdellovibrionota genomes. The size and color of dots represent the percentage of the sources in each group.

Figure 2. Genomic and gene content difference among Bdellovibrionota groups.

(A) GC content and genome size of Bdellovibrionota genomes. (B) Number of group-specific COGs and KEGG genes. (C) Venn diagram showing group-specific and shared peptidases of Bdellovibrionota. (D) Heatmap illustrating differences in number of carbohydrate active enzyme (CAZy) classes.

Figure 3. Vertical distribution of Bdellovibrionota in marine waters.

(A) Relative abundance of Bdellovibrionota was calculated as percentage of Bdellovibrionota miTags in metagenomes of the Tara Ocean project (depth from 5m to 1,000m) and the Mariana Trench ($\geq 1,000$ m) (Table S7) (Gao *et al.*, 2019; Li *et al.*, 2019). (B) Relative abundances of Bdellovibrionota groups in different marine layers were plotted. T-test was performed to compare the relative abundances at different layers. (****, $p < 0.0001$; *, $p < 0.05$).

Figure 6. Phylogenetic analysis and conserved motif of chitinases from Oligoflexia.

The phylogenetic tree of chitinases was built with RAxML (PROTGAMMALG model). The bootstrap values based on 1,000 replicates were denoted by the size of the black dots on the branches. The Oligoflexia genomes were marked with a green dot. The conserved motif of chitinases was highlighted with blue background in the multiple sequence alignment.

Figure S1. Rooted phylogenetic tree of Bdellovibrionota 16S rRNA genes.

Maximum-likelihood (ML) phylogenetic tree of 16S rRNA genes extracted from high-quality genomes of Bdellovibrionota were reconstructed using PROTGAMMALG model. Blue dots refer to the MAGs retrieved from marine waters of this study.

	non-marine	marine	
K15773	0.00	100.00	<i>hipB</i> ; HTH-type transcriptional regulator / antitoxin HipB
K03442	0.00	100.00	<i>mcsS</i> ; small conductance mechanosensitive channel
K02168	0.00	100.00	<i>betT</i> , <i>betS</i> ; choline/glycine/proline betaine transport protein
K06218	0.00	100.00	<i>relE</i> ; mRNA interferase RelE/StbE
K07339	16.67	100.00	<i>hicA</i> ; mRNA interferase HicA [EC:3.1.-.-]
K18843	0.00	100.00	<i>hicB</i> ; antitoxin HicB
K19092	0.00	100.00	toxin ParE 1/3/4
K08641	0.00	100.00	<i>vanX</i> ; zinc D-Ala-D-Ala dipeptidase [EC:3.4.13.22]
K01273	0.00	100.00	<i>depe</i> ; membrane dipeptidase [EC:3.4.13.19]
K05995	0.00	100.00	<i>pepE</i> ; dipeptidase E [EC:3.4.13.21]
K01177	0.00	100.00	beta-amylase [EC:3.2.1.2]
K00499	0.00	100.00	<i>cmo</i> ; choline monooxygenase [EC:1.14.15.7]
K10111	0.00	100.00	<i>malK</i> ; multiple sugar transport system ATP-binding protein [EC:3.6.3.-]
K10112	0.00	100.00	<i>msmX</i> ; multiple sugar transport system ATP-binding protein
K07272	0.00	100.00	<i>rgpF</i> ; rhamnosyltransferase [EC:2.4.1.-]
K05841	0.00	100.00	sterol 3-beta-glucosyltransferase [EC:2.4.1.173]
K03313	0.00	100.00	<i>nhaA</i> ; Na ⁺ :H ⁺ antiporter, NhaA family
K19294	0.00	100.00	<i>algI</i> ; alginate O-acetyltransferase complex protein AlgI
percent in genomes (%)			

Bdello-group1

Oligoflexia

Figure S2. Genes in genomes affiliated with Oligoflexia and Bdello-group1 from marine environment.

The percentage of the KEGG genes in the genomes was calculated for the marine and non-marine genomes of Oligoflexia and Bdello-group1. Genes identified in Bdello-group1 marine genomes were involved in betaine metabolism, type II toxin-antitoxin (TA) system and dipeptidases. Genes present in Oligoflexia marine genomes were involved in cold resistance (beta-amylase), betaine synthesis (*cmo*), multiple sugar transport and metabolism (*malK*, *msmK*, *algI* and *rgpF*).

Oligoflexia							
	desert	not available	freshwater	freshwater	freshwater	deep-sea	
K07496	7	5	7	0	0	0	putative transposase
K01183	5	2	2	0	0	0	chitinase [EC:3.2.1.14]
K02231	2	2	1	0	0	1	<i>cobP</i> ; adenosylcobinamide kinase [EC:2.7.1.156 2.7.7.62]
K03933	1	1	1	0	0	0	<i>cpbD</i> ; chitin-binding protein
K00162	1	1	1	1	1	1	<i>pdhB</i> ; pyruvate dehydrogenase E1 component beta subunit [EC:1.2.4.1]
K00631	1	1	1	1	1	1	<i>plsB</i> ; glycerol-3-phosphate O-acyltransferase [EC:2.3.1.15]
K05814	0	0	1	1	1	0	<i>ugpA</i> ; sn-glycerol 3-phosphate transport system permease protein
K05813	0	0	1	1	1	0	<i>ugpB</i> ; sn-glycerol 3-phosphate transport system substrate-binding protein
K05815	0	0	1	1	1	0	<i>ugpE</i> ; sn-glycerol 3-phosphate transport system permease protein
K01695	1	1	0	0	0	1	<i>trpA</i> ; tryptophan synthase alpha chain [EC:4.2.1.20]
K00766	1	1	1	0	0	1	<i>trpD</i> ; anthranilate phosphoribosyltransferase [EC:2.4.2.18]
K01657	1	1	0	0	0	1	<i>trpE</i> ; anthranilate synthase component I [EC:4.1.3.27]
K01658	1	1	0	0	0	1	<i>trpG</i> ; anthranilate synthase component II [EC:4.1.3.27]
K02194	1	1	1	0	0	1	<i>ccmB</i> ; heme exporter protein B
K02195	1	1	1	0	0	1	<i>ccmC</i> ; heme exporter protein C
K02197	1	1	1	0	0	1	<i>ccmE</i> ; cytochrome c-type biogenesis protein CcmE
K02198	1	1	1	0	0	1	<i>ccmF</i> ; cytochrome c-type biogenesis protein CcmF
K02199	1	1	1	0	0	1	<i>ccmG</i> ; cytochrome c biogenesis protein
K02200	1	1	1	0	0	1	<i>ccmH</i> ; cytochrome c-type biogenesis protein CcmH

Figure S3. Copy number of interested KEGG genes of Oligoflexia from different environments.

The special genes are only present in Oligoflexia genomes inhabiting different environments. Chitinase (more than one copy) and chitin-binding protein coding genes were predicted in some genomes.

Figure S4. Phylogenetic tree of chitin-binding proteins encoded by *Oligoflexia* genomes.

The red dots refer to amino acid sequences of chitin-binding proteins predicted based on KEGG annotation of *Oligoflexia* genomes.

755

756

757 Table S1. GTDB annotation of all Bdellovibrionota genomes from GTDB-TK

758 database, five MAGs of our marine water samples and eight Bdellovibrionota

759 genomes from NCBI.

760 Table S2. Information of completeness, contamination, genome size and GC content

761 of all genomes.

762 Table S3. Average nucleotide identity of all genomes.

763 Table S4. Environment source of genomes.

764 Table S5. KEGG and COG genes of Bdellovibrionota involved as predator.

765 Table S6. KEGG annotation results of genomes.

766 Table S7. The information of water samples collected from Mariana Trench with the

767 depth no less than 1,000m.

768 Table S8. Blastp results of chitinase and chitin-binding protein.

769

770

771

772