Widespread microbial mercury methylation genes in the global ocean — **Source link** ⧉

Emilie Villar, Emilie Villar, Léa Cabrol, Léa Cabrol ...+1 more authors

Institutions: Aix-Marseille University, Pierre-and-Marie-Curie University, University of Chile

Published on: 01 Jun 2020 - Environmental Microbiology Reports (Wiley)

Topics: Mercury (element) and Methylmercury

Related papers:

- The Genetic Basis for Bacterial Mercury Methylation

- Microbial mercury methylation in Antarctic sea ice.

- Mercury methylation by novel microorganisms from new environments

- Global prevalence and distribution of genes and microorganisms involved in mercury methylation.

- Molecular evidence for novel mercury methylating microorganisms in sulfate-impacted lakes.

Share this paper: f 𝕏 in ✉

View more about this paper here: https://typeset.io/papers/widespread-microbial-mercury-methylation-genes-in-the-global-4fo6ormbn7

1   **Widespread microbial mercury methylation genes in the global ocean**

2   Emilie Villar[1,2], Lea Cabrol[1]*, Lars-Eric Heimbürger-Boavida[1]*

3

4   [1]Aix Marseille Université, Univ Toulon, CNRS, IRD, Mediterranean Institute of Oceanography

5   (MIO) UM 110, 13288, Marseille, France

6   [2]Sorbonne Université, Université Pierre et Marie Curie - Paris 6, CNRS, UMR 7144 (AD2M),

7   Station Biologique de Roscoff, Place Georges Teissier, CS90074, Roscoff, 29688 France

8   *Both authors contributed equally to this work.

9   Corresponding Author: Léa Cabrol, lea.cabrol@mio.osupytheas.fr

10

11   **Abstract**

12   Methylmercury is a neurotoxin that bioaccumulates from seawater to high concentrations in

13   marine fish, putting human and ecosystem health at risk. High methylmercury levels have

14   been found in the oxic subsurface waters of all oceans, yet only anaerobic microorganisms

15   have been identified so far as efficient methylmercury producers in anoxic environments.

16   The microaerophilic nitrite oxidizing bacteria *Nitrospina* has been previously suggested as a

17   possible mercury methylator in Antarctic sea ice. However, the microorganisms processing

18   inorganic mercury into methylmercury in oxic seawater remain unknown. Here we show

19   metagenomic evidence from open ocean for widespread microbial methylmercury

20   production in oxic subsurface waters. We find high abundances of the key mercury

21   methylating genes *hgcAB* across all oceans corresponding to taxonomic relatives of known

22   mercury methylators from Deltaproteobacteria, Firmicutes and Chloroflexi. Our results

23   identify *Nitrospina* as the predominant and widespread key player for methylmercury

24   production in the oxic subsurface waters of the global ocean.

25

26   **Introduction**

27   Human activities release 2500 tons of inorganic mercury (Hg) every year and have added 55

28   000 tons of Hg to the global ocean since the industrial revolution [1]. Humans are exposed to

29   Hg in the form of methylmercury (MeHg), mainly through the consumption of marine fish.

30   The Minamata Convention (www.mercuryconvention.org) aims to protect human health

31   from the adverse effects of Hg *via* the reduction of anthropogenic, inorganic Hg emissions.

32   To understand the efficacy and time-scales of lowered Hg emissions to reduce fish MeHg

33   levels, we must fully understand the origin of marine MeHg. Microorganisms play a central

34   role in Hg transformations. We must identify the Hg methylating microbes and the factors

35   controlling their distribution in order to better constrain MeHg production in the global

36   ocean.

37   As the only cultured microbes known to produce MeHg to date are anaerobic, research

38   focused for many years on a MeHg source in anoxic marine sediments [2–5]. However, several

39   lines of independent evidence speak in favour of *in situ* MeHg production in oxic seawater as

40   the main source of fish MeHg. Recent large scale oceanographic expeditions found

41   subsurface MeHg maxima in every ocean basin [4,6]. The proportion of MeHg to inorganic Hg

42   throughout the oxic seawater column is higher than those found in anoxic sediments.

43   Laboratory experiments show that Hg methylation can occur in anaerobic microniches that

44   occur within sinking particles in oxic waters [7]. Bianchi et al. [8] provide compelling evidence

45   that anaerobic microbes thrive in anoxic microenvironments of sinking particulate organic

46   matter. Independently, incubation experiments with isotopically labelled Hg spikes show

47   significant *in situ* Hg methylation in oxic seawater [9]. Additional evidence stems from Hg

48   stable isotope signatures of marine fish, that can only be explained if 60-80% of the MeHg is

49   produced in open ocean subsurface waters [10]. Lastly, a pioneering study found a compound

50   specific $\delta^{13}C$ signature of fish tissue MeHg similar to algal $\delta^{13}C$, suggesting that MeHg is

51   produced in the open ocean water column [11].

52   A major breakthrough has been made with the discovery of two key genes, *hgcA* and *hgcB*,

53   that control Hg methylation in model anaerobic Hg-methylators[5]. The presence of the *hgcAB*

54   operon predicts Hg methylation capacity in diverse microorganisms [2]. A screening of publicly

55   available microbial metagenomes found the *hgcAB* genes in nearly all anaerobic

56   environments, but the study only rarely detected the genes in pelagic marine water column

57   metagenomes in the open ocean [12]. In antarctic sea ice a marine microaerophilic nitrite

58   oxidizing bacterium belonging to the *Nitrospina* genus has been recently identified as a

59   potential Hg methylator with HgcA-like proteins [13]. We aim to resolve the paradox between

60   the wealth of geochemical evidence for *in situ* MeHg production and the absence of known

61   anaerobic Hg methylators in the open ocean. Metagenomic data from 243 *Tara* Oceans

62   samples from 68 open ocean locations covering most ocean basins was analysed to generate

63   an ocean microbial reference gene catalog[14]. We screened the *Tara* Oceans metagenomes

64   for the presence of the key *hgcA* methylating gene and provide compelling evidence on the

65   potential key players producing MeHg in the open ocean.

66   Results and Discussion

67   **Identification of HgcAB homologs in the ocean gene catalog**

68   Ten *hgcA* and 5 *hgcB* homolog genes were identified in the Ocean Microbial Reference Gene

69   Catalog [14] (OM-RGC), 6 scaftigs presenting simultaneously *hgcA* and *hgcB* (Fig. 1,

70   Supplementary Table 1). Alignment of HgcA sequences revealed 7 sequences with the

71   conserved NVWCAA motif [5] and one sequence with the modified NIWCAA motif on the 'cap

72   helix' region. Mutation experiments previously showed that the structure of the putative

73   'cap helix' region harbouring Cys93 is crucial for methylation function [15]. Two HgcA

74   sequences were truncated (OM-RGC.v1.019516181, OM-RGC.v1.015822836), preventing

75   inspection of their conserved motif, but they could be unequivocally assigned to HgcA

76   sequences based on their phylogenetic placement and high similarity with known HgcA

77   sequences (Fig. 2). The 5 HgcB sequences presented the conserved motif ECGAC [5]

78   (Supplementary Table 1).

79
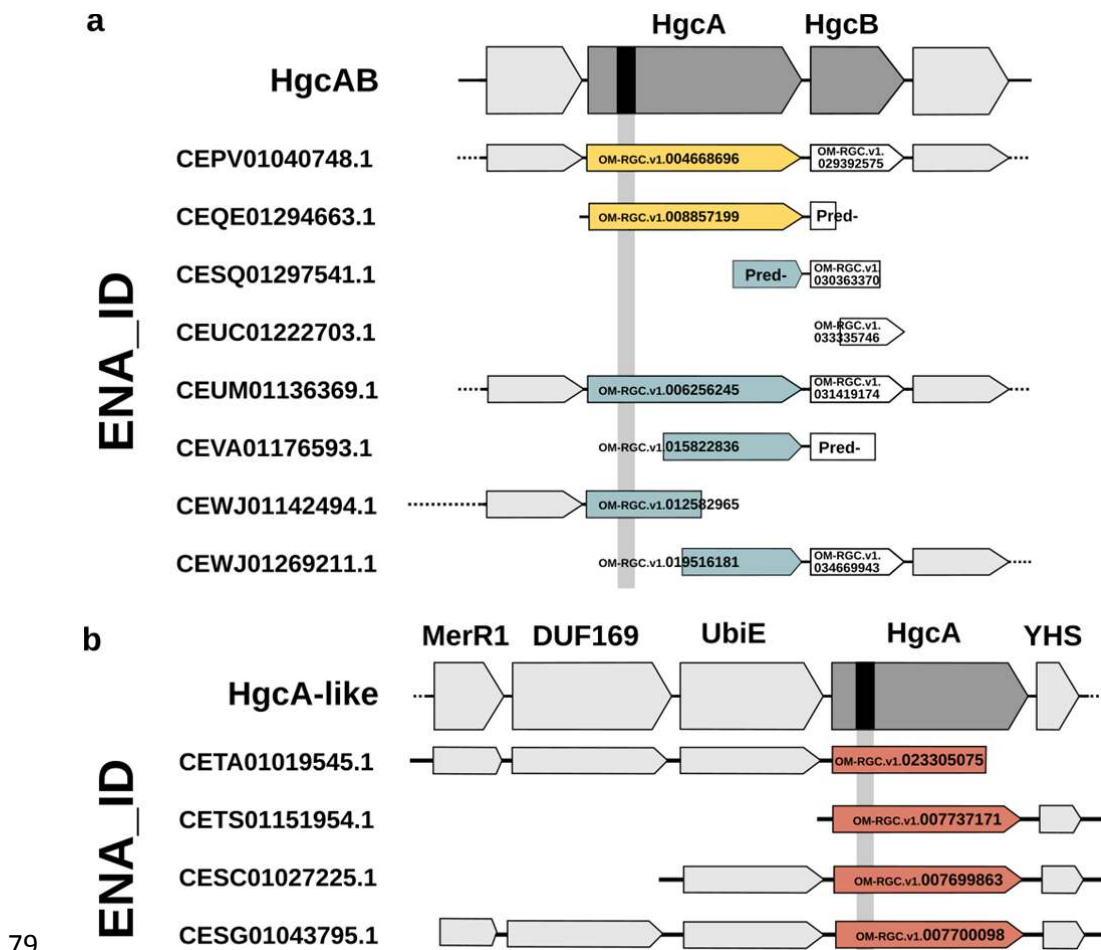
80   **Figure 1 | The genomic context of the HgcA orthologs. a,** HgcAB operon. **b,** HgcA-like proteins. The

81   12 retrieved scaftigs are identified by their ENA_ID on the left of the figure. The solid lines represent

4

82 the extent of the scaftig sequence and the dashed lines indicate that the scaftig sequence is longer

83 than the represented part. The location of the conserved motif is indicated on the HgcA box by a

84 black bar. When present in *Tara* Oceans samples, the corresponding gene identifier is indicated on

85 the bar for HgcA and HgcB, or indicated as (Pred-) if the gene was incomplete and the protein

86 sequence was partially predicted. The colour of the HgcA boxes refers to the biogeographical

87 clustering as defined in Fig.3 (Cluster 1 in blue, Cluster 2 in yellow, Cluster 3 in red). For Cluster 3

88 sequences (assigned to Nitrospina), the genomic context was enlarged to show the closest sequences

89 (MerR1: mercuric resistance operon regulatory protein, UbiE: Ubiquinone/menaquinone

90 biosynthesis C-methyltransferase, DUF169: Hypothetical protein with DUF 169 motif, YHS:

91 Hypothetical protein with YHS domain).

92

93 **HgcA sequences found in the Tara Oceans assemblies covered nearly all known Hg**

94 **methylators**

95 Phylogenetic placement of the 10 sequences found in the *Tara* Oceans assemblies covered

96 nearly all known Hg methylators (Fig. 2). Four sequences (OM-RGC.v1.007700098, OM-

97 RGC.v1.007737171, OM-RGC.v1.023305075, OM-RGC.v1.007699863) were closely related to

98 the HgcA-like proteins described by Gionfriddo et al. [13] for *Nitrospina sp.* The Nitrospinae

99 phylum has been described as a distinct phylogenetic group of lithoautotrophic nitrite

100 oxidizing bacteria exclusively found in marine environments [16], particularly abundant in

101 oxygen-deficient zones [17].

102 The remaining 6 HgcA sequences were distributed between Deltaproteobacteria, Firmicutes

103 and Chloroflexi phyla. Within Deltaproteobacteria, three orders were represented, namely

104 *Desulfovibrionales, Desulfobacterales* and *Syntrophobacterales.* OM-RGC.v1.006256245 was

105 most closely related to HgcA from *Pseudodesulfovibrio profundus*, a strictly anaerobic

106 piezophilic sulfate-reducer bacteria (SRB) previously isolated from marine sediment [18]. It

107 belongs to the Desulfovibrionales order, which contains several members with confirmed

108 Hg-methylating capacity, such as the model species *Desulfovibrio desulfuricans* with

5

109   exceptionally high Hg-methylation rates, isolated from estuarine sediments [19]. OM-

110   RGC.v1.019516181 and OM-RGC.v1.012582965 belonged to *Desulfobacterales*, a well-known

111   order of SRB containing efficient Hg-methylators such as *Desulfobulbus propionicus* and

112   *Desulfococcus multivorans*. Finally, OM-RGC.v1.004668696 belonged to

113   Syntrophobacterales. The closest relative of OM-RGC.v1.004668696 with strong confirmed

114   methylation potential was the non-SRB obligate syntroph *Syntrophus aciditrophicus* [2].

115   Syntrophic bacteria are important Hg-methylators in low-sulfate ecosystems [20,21], where

116   they degrade OM in association with H_2-consuming microorganisms such as sulfate-

117   reducers, iron-reducers and methanogens.

118   Within Firmicutes, OM-RGC.v1.015822836 was tightly related to HgcA from recently isolated

119   human gut bacteria *Khelaifiella* in the Clostridiales order [22]. Their closest relative with

120   confirmed methylation potential is the non-SRB *Dethiobacter alkaliphilus*, with low to

121   moderate Hg-methylation capacity [2].

122   OM-RGC.v1.008857199 was related to Chloroflexi, a phylum for which several *hgcAB*-

123   carriers have been identified, but for which experimental confirmation of Hg methylation

124   capacity is still needed. This sequence clusters tightly with HgcA from *Dehalococcoides*

125   *mccartyi*, which has been reported as a potential methylator, albeit in minor abundance, in

126   freshwater marshes [20]. These two sequences are separated from other Chloroflexi HgcA

127   sequences and more closely related to HgcA sequences from Syntrophobacterales, showing

128   that the taxonomy and the HgcA-phylogeny are not always congruent. The phylogenetically

129   irregular distribution of *hgcA* can be an indication of horizontal gene transfers (HGT) and/or

130   gene deletions in response to stress, suggesting the prevalent influence of environment on

131   Hg-methylation ability [23].

132

7
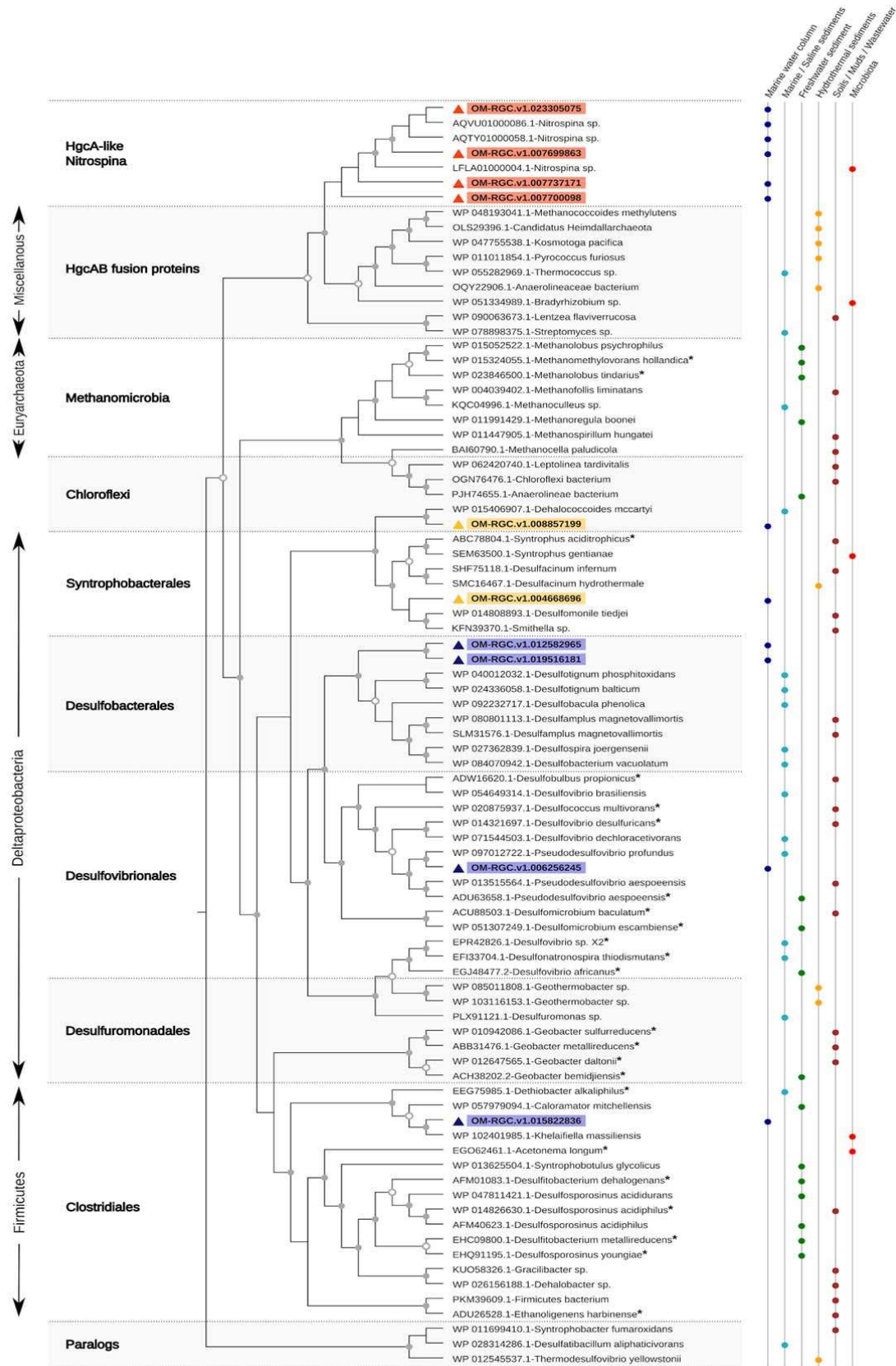
133 **Figure 2 | Phylogenetic tree of HgcA homolog sequences found in the *Tara* Oceans assemblies**

134 Maximum likelihood phylogenies were inferred using PhyML Best AIC Tree with the best model of

135 sequence evolution Blosum62+I+G+F. Branch support was calculated using the non-parametric

136 Shimodaira-Hasegawa-like approximate likelihood ratio test. The triangle colour refers to the

137 biogeographical clustering of the HgcA sequences retrieved from *Tara* Oceans assemblies, as defined

138 in Fig.3 (Cluster 1 in blue, Cluster 2 in yellow, Cluster 3 in red). The tree was rooted using 3

139 paralogues from confirmed non-Hg methylating bacteria. Sequences from experimentally confirmed

140 mercury methylators were indicated with an asterisk. Support values using 1,000 resamples are

141 shown when >50 and coloured squares indicate the isolation source.

142

143 Among the 10 HgcA sequences found in the gene ocean catalogue, none was affiliated to

144 methanogenic *Archaea*. Even if the co-existence of methanogens and sulfate-reducers has

145 been evidenced in marine sediments [24], sulfate reduction usually outcompetes

146 methanogenesis in seawater under non-limiting sulfate concentrations [25]. Our results thus

147 show that Hg-methylators in the ocean span a large taxonomic diversity, not limited to

148 sulfate-reducing bacteria.

149

150 **Biogeography distinguishes three groups of putative marine Hg methylators**

151 Once clearly identified and phylogenetically assigned, the biogeographic distribution

152 patterns of *hgcA* was evaluated. The 10 HgcA sequences were identified in 77 samples out

153 of the 243 available *Tara* Oceans metagenomes and were clearly distributed in three clusters

154 according to their abundance patterns (Fig. 3). The biogeographic clustering was consistent

155 with the HgcA-phylogeny.

156 Cluster 1 gathered Desulfobacterales, Clostridiales and Desulfovibrionales HgcA sequences,

157 exclusively present in 23 oxic surface waters (< 120 m-depth, > 10 μM_{O2}). Highest

158 abundances were found in the photic zone of the Pacific Ocean, especially in the area

159 surrounding the Marquesas Islands. This region is characterized by extensive plankton

8

160   blooms triggered by a physico-chemical phenomenon called Island Mass Effect related to

161   iron fertilization. In this Cluster, the HgcA sequence OM-RGC.v1.006256245 related to the

162   *Desulfovibrionales* order (containing most of the experimentally confirmed Hg-methylators)

163   was the most frequent and abundant in the 23 oxic samples.

164   The phylogenetic placement of the two sequences grouped in Cluster 2 is poorly supported.

165   The most abundant sequence was related to *Smithella* and *Desulfomonile tiedjei*

166   (Syntrophobacterales) while the other one was close to Chloroflexi (Fig. 2, Supplementary

167   Table 2). HgcA sequences from Cluster 2 were identified in 15 surface and subsurface

168   samples, mostly in suboxic waters : sequences found in samples with oxygen concentration

169   below 10 µM accounted for 98% of total Cluster 2 abundances (Supplementary Figure 2).

170   The highest abundances of Cluster 2 HgcA sequences were found in the subsurface waters of

171   the northern stations within the Arabian Sea Oxygen Minimum Zone (Stations TARA_036 to

172   TARA_039), under the influence of a previous major bloom event, where high particle

173   concentrations and strong anaerobic microbial respiration have been reported [26]. Cluster 2

174   sequences were also found in lower abundance in the shallow anoxic zone of the Pacific

175   North Equatorial Counter Current (Stations TARA_137 and TARA_138, see methods).

176
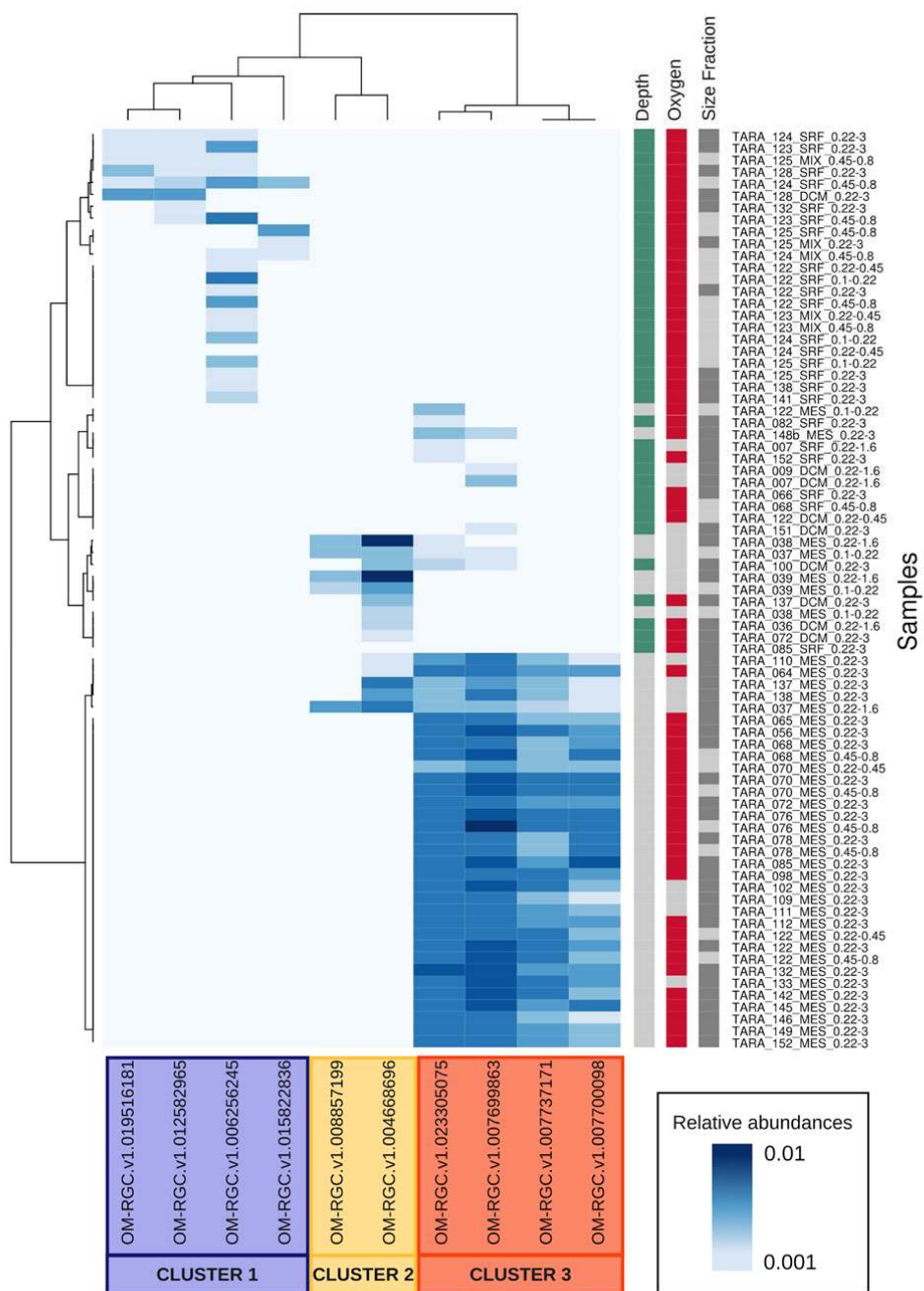
Figure 3 | Distribution of HgcA in *Tara* Oceans samples

HgcA relative abundance (from 0.01 to 0.1) is indicated by the white-blue gradient. The hierarchical

clustering highlighted three gene clusters with high abundances in specific samples with marked

environmental features, as suggested by colored squares. Surface samples were collected in the

upper layer (< 120 m-depth, in green) while subsurface were collected below 120 m-depth (in grey).

Seawater was considered as oxic when $O_2 > 10$ µM (in red) and suboxic when $O_2 < 10$ µM (in grey).

10

183   Larger size fraction samples are in dark grey (0.22-3 µm) and smaller size fractions samples (<0.8 µm)

184   are in light grey.

185

186

187   The most abundant HgcA-like proteins were grouped in Cluster 3 and were exclusively

188   assigned to *Nitrospina*. These *Nitrospina* HgcA-like proteins were found in 47 samples,

189   widespread across all sampled ocean basins. They were almost exclusively found in

190   subsurface water (> 120 m-depth) and were more frequent in the oxic waters (> 10 µM$_{O2}$).

191   Subsurface oxic waters accounted for 84% of total *Nitrospina*-HgcA abundance

192   (Supplementary Figure 2). Their highest relative abundance was found in the South Atlantic

193   and the South Pacific Oceans (Fig. 4, Supplementary Table 2). *Nitrospina* HgcA abundance

194   was positively correlated to nitrate concentration (R 0.54, P < 0.005), which is consistent

195   with *Nitrospina's* role as nitrate producer through nitrite oxidation, and with the well-known

196   nitrate enrichment with depth in the ocean.

197

198   **Nitrospina as the most predominant and widespread methylator in the open ocean**

199   The predominant and widespread HgcA-like homologs were phylogenetically extremely

200   close to the *Nitrospina*-related ones (Supplementary Figure 1) previously identified by

201   metagenomic analysis as potential Hg-methylators within Antarctic sea ice and brine, and

202   further detected by PCR in seawater samples below the ice [13]. The four *Nitrospina* HgcA-like

203   sequences from our study were distinct from HgcA in confirmed Hg-methylators, and also

204   from HgcAB fusion proteins reported in environmental metagenomes [12] (Figure 2). The few

205   cultured strains harboring a fused *hgcAB* gene (*Methanococcoides methylutens* and

206   *Pyrococcus furiosus)* were unable to produce MeHg in experimental conditions [12,27]. Through

207   sequence alignment against Protein Data Bank templates, we confirmed that the four

208   *Nitrospina* HgcA-like homologs showed high conservation of six residue positions involved in

209   cobalamin binding, which is mandatory for methyl group transfer to Hg [13] (Supplementary

210   Fig. 1). Protein structure modelling suggests that some *Nitrospina* species may be capable of

211   Hg-methylation. The observed mutations (N71 and C74) do not suppress Hg methylation

212   capacity, according to mutagenesis experiments in the model methylator *D. desulfuricans*

213   ND132 [15]. The strictly conserved cysteine facilitates the transfer of methyl groups to

214   inorganic Hg [28]. The two current *Nitrospina* isolates (*N. gracilis* and *N. watsonii*) have not

215   been experimentally tested to date for their Hg-methylation capacity. *N. gracilis* genome

216   lacks the *hgcA* gene. The complete genome of *N. watsonii* is not available. From the 12

217   *Nitrospina* genome assemblies available on NCBI at the time of writing, we found HgcA-like

218   proteins (harbouring the six mandatory residues for Hg-methylation) in three strains only:

219   SCGC AAA288-L16 (single cell whole genome from 770 m-deep ALOHA station, North Pacific

220   Ocean), AB-629-B06 (single cell whole genome from dark ocean water column) and LS_NOB

221   (from a deep-sea sponge holobiont; Supplementary Fig. 1).
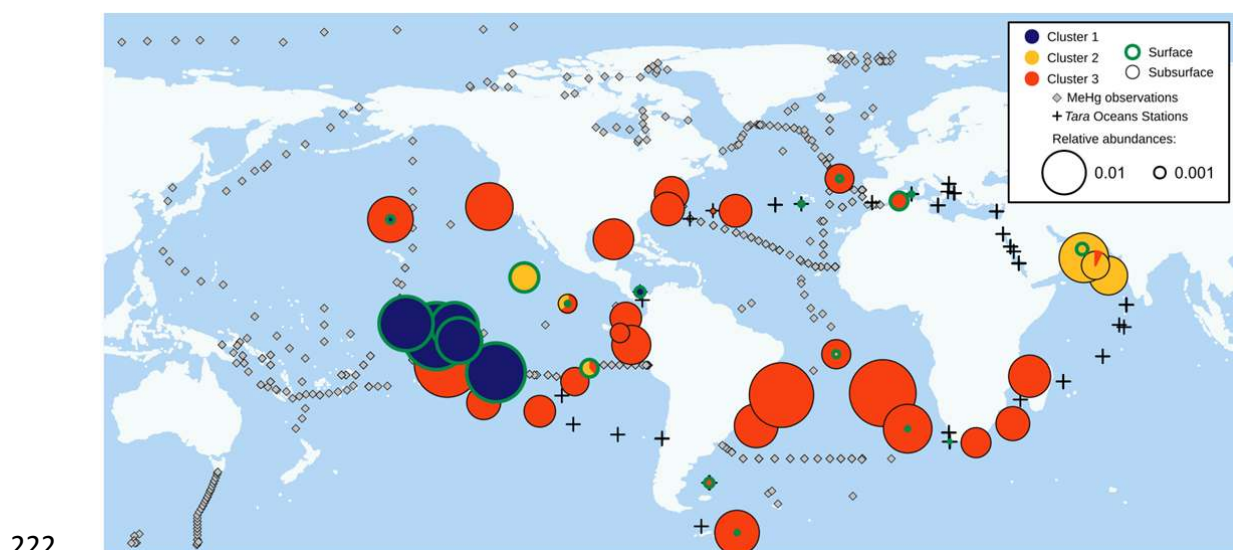
222

223   **Figure 4 | HgcA biogeography**
224   Circle sizes are proportional to the cumulated HgcA homolog genes abundances at each station. The
225   pie charts indicate the cluster attribution (legend in the chart), and their border color indicates the

226   sampling depth: surface samples (<120 m-depth) in green and subsurface samples (> 120 m-depth) in

227   grey. *Tara* Oceans stations without detected *HgcAB* genes are represented by black crosses and

228   seawater MeHg profiles from the literature (Supplementary Text 1) by black diamonds.

229

230   Mercury methylation has long been described for anaerobic environments [12] and *hgcA* genes

231   have been found exclusively in anaerobic Bacteria and Archea [5]. Yet, we find the most

232   abundant HgcA homologs are strongly dominant in oxic subsurface samples, where they

233   coincide with the subsurface MeHg concentration peaks [6], and are carried by the nitrite

234   oxidizing bacteria *Nitrospina,* usually considered as aerobic.

235   Several clues may explain this apparent contradiction. First, it is increasingly recognized that

236   anaerobic processes can occur in anoxic niches such as organic matter aggregates in the

237   middle of oxic waters [8]. *Nitrospina* sequences were predominantly present in the larger size

238   fractions (accounting for 78% of total HgcA abundances), suggesting that Hg-methylation -as

239   other anaerobic processes- might be associated with particles, where anoxic niches are likely

240   to be favourable to *Nitrospina* methylating activity. Several features suggest the adaptation

241   of *Nitrospina* to low-oxygen environments. *Nitrospina* have been detected as particularly

242   abundant (up to 10% of the bacterial community) in several upwelling and oxygen-deficient

243   zones [29].

244   Comparative genomics revealed a close evolutionary relationship between *Nitrospina* and

245   Anammox bacteria, including horizontal gene transfer events, suggesting the coexistence of

246   these organisms in hypoxic or anoxic environments [16], as confirmed in incubation

247   experiments [30,31]. Genome analysis of several *Nitrospina* strains revealed unexpected

248   adaptation features to low-oxygen environments: no ROS defence mechanism, dependence

249   on highly oxygen-sensitive enzymes for carbon fixation, and high O_2-affinity cytochromes

250   [16,32].

13

251   *Nitrospina* can play diverse ecological roles beyond the nitrogen cycle [33]. *Nitrospina* can use

252   alternative anaerobic pathways to gain energy, using other terminal electron acceptors than

253   O_2 during fermentation under hypoxic or anaerobic conditions, such as sulfur compounds or

254   metal oxides. Their capacity to cope with environmental Hg through methylation is worth

255   considering, since their genome is well equipped against other toxic compounds (arsenate-

256   and mercuric-reductase, metallic cation transporters, multidrug export system) [16,32].

257   Mercury methylation potential might have been acquired by horizontal gene transfer.

258   Within the four *Nitrospina* scaftigs harbouring *hgcA*, other neighbour genes related to

259   methyl group transfer and Hg metabolism are found, such as the *merR1* regulator of the *mer*

260   operon involved in Hg resistance, the *ubiE* methyltransferase and the putative metal-binding

261   YHS domain (Fig. 1). This genomic context can lead to hypothesize that the expression of

262   these genes, including *hgcA*, is under the same Hg-induced regulation as the *mer* operon,

263   triggered by *merR*.

264   The choice of *hgcAB* as an indicator of Hg-methylation has to be discussed. First, the

265   presence of *hgcAB* appears necessary but not sufficient for Hg methylation. Indeed,

266   unsuccessful attempts to transfer Hg-methylation capacity to a non-Hg-methylating strain

267   suggest that unidentified additional genes might be needed for effective MeHg production

268   [15]. Several critical steps are involved in the Hg-methylation process, such as Hg(II) sensing,

269   cellular uptake of Hg(II) by active transport, methyl-group providing and transfer, and MeHg

270   export from the cell. All these steps could be targeted as functional markers of Hg-

271   methylation in the environment in order to provide a more complete picture of the process.

272   Second, the exact contribution of HgcAB to Hg-methylation is not well understood. Since Hg

273   methylation does not confer Hg resistance, it cannot be considered as a protection

274   mechanism against Hg toxicity [19]. In model strains *D. dechloroacetivorans*, net Hg

14

275   methylation was not clearly induced by inorganic Hg and not significantly correlated to

276   *hgcAB* gene expression levels, but rather influenced by environmental factors, growth

277   conditions and energetic metabolism [19,34]. The variability of the methylation potential has

278   been evidenced in different strains, and the implication of *hgcAB* might also vary between

279   strains. Such functional gene approaches are powerful to track biogeochemical potentials in

280   extended environments but remain limited to well described metabolic pathways, ignoring

281   genes with unknown function [35].

282   Here, we bring metagenomic evidence for widespread presence of microbial Hg-methylators

283   in the global ocean, thus reconciling with previous geochemical hints pointing to *in situ*

284   MeHg production in the water column. The key Hg-methylating genes found across all

285   oceans corresponded to taxonomic relatives of known Hg-methylators from

286   Deltaproteobacteria, Firmicutes and Chloroflexi phyla. We further identified the

287   microaerophilic NOB *Nitrospina* as the potential dominant Hg-methylator in the global

288   ocean, ubiquitous at the DNA-level, and favoured by oxic subsurface waters (Supplementary

289   Figure 2). A critical next step would be to examine their *hgcA* expression levels and to

290   evaluate Hg-methylation capacity in *Nitrospina* cultures. Further studies should also

291   determine the physicochemical parameters controlling *Nitropina* Hg-methylation activity

292   level, in order to better understand how they will respond to expected global changes. Our

293   results open new avenues for disentangling the functional role of microorganisms in marine

294   Hg cycling. Our analysis of the *Tara* Oceans metagenomes reveals global distribution of the

295   key Hg methylating genes (*hgcA* and *hgcB*) and pinpoints *Nitrospina* as responsible for

296   widespread open ocean MeHg production in subsurface oxic seawater. Our study implicates

297   the subsurface oxic waters of all oceans as potential source of MeHg that should be

298   considered in the global Hg-cycle budgets, and identifies microbial target for further

15

299   research on marine MeHg production. We hypothesize that besides anthropogenic Hg

300   emissions, ongoing global climate change might have a previously underestimated effect on

301   *in situ* marine MeHg production by water-column microorganisms, by disturbing microbial

302   assemblages, activity, and environmental drivers governing Hg-methylation.

303   Methods

304   **Identification of HgcAB environmental sequences in oceanic metagenomes.**

305   *hgcA* and *hgcB* genes encode for a putative corrinoid protein, HgcA, and a 2[4Fe-4S]

306   ferredoxin, HgcB, serving respectively as methyl group carrier and electron donor for

307   corrinoid cofactor reduction. HgcA and HgcB homologs were retrieved by searching Hidden

308   Markov Model profiles (HMM) [36] provided by Podar et al. [12] in the Ocean Microbial

309   Reference Gene Catalog[14] (OM-RGC) using the Ocean Gene Atlas[37] (http://tara-

310   oceans.mio.osupytheas.fr/ocean-gene-atlas/). The OM-RGC is the most exhaustive catalogue

311   of marine genes to date including datasets from *Tara* Oceans metagenomic assemblies and

312   other publicly available marine genomic and metagenomic datasets. We applied an e-value

313   threshold of 1e-20. The corresponding scaftigs (i.e. the assembled sequences where the

314   homolog genes were predicted) were retrieved from the raw assemblies deposited at ENA

315   (Supp Data 1 & 4). Eight scaftigs without *Tara* Oceans mapped reads were discarded. The

316   remaining scaftigs were annotated using Prokka with default parameters [38]. The resulting

317   translated sequences were aligned separately for HgcA and HgcB using Jalview 2.10 and

318   alignments were cleaned manually[39]. For further analysis, we kept HgcA sequences if they

319   possess the conserved motif NVWCAA[5], or if the neighbouring HgcB sequence was present

320   on the scaftig.

321   **HgcA phylogenetic analysis.**

16

322   A phylogenetic tree was built from the 10 HgcA sequences kept, 55 HgcA protein sequences

323   representative of known Hg-methylator clades belonging to Archaea, Firmicutes, Chloroflexi

324   and Deltaproteobacteria, including 18 experimentally-confirmed Hg-methylators (initially

325   published by Parks et al. [5], and updated at

326   https://www.esd.ornl.gov/programs/rsfa/data.shtml), as well as 9 HgcAB fusion proteins [13]

327   and 3 HgcA-like proteins predicted from *Nitrospina* genome assemblies using Prokka [38]. The

328   tree was rooted with 3 paralogs from confirmed non-Hg-methylating strains [13]. The closest

329   sequences (i.e. best e-value match) of each environmental HgcA sequence were retrieved

330   using BLASTp against non-redundant RefSeq protein database excluding sequences from

331   uncultured organisms [40], and included in the tree.

332   The 80 sequences were aligned using MAFFT [41] and gap-containing sites were removed using

333   the mode gappyout of TrimAl [42]. Maximum likelihood phylogenies were inferred using

334   PhyML Best AIC Tree (version 1.02b) implemented in Phylemon[43] (version 2.0) with the best

335   model of sequence evolution Blosum62+I+G+F. Branch support was calculated using the

336   non-parametric Shimodaira-Hasegawa-like approximate likelihood ratio test. The final tree

337   was edited using Evolview [44], especially by annotating the isolation sites retrieved from

338   Genomes OnLine Database [45].

339   **Conserved sites in HgcA.**

340   Four sequences from OM-RGC related to *Nitrospina* were aligned with the same Protein

341   Data Bank (PDB) templates as Gionfriddo et al. [13], as well as the 3 HgcA-like proteins from

342   *Nitrospina* genome assemblies, and conserved residues were checked. The chosen PDB

343   structural templates (4djd_C, 2h9a_A, 4C1n_C, 2ycl_A) were the gamma subunit of the

344   corrinoid S-Fe acetyl-CoA decarbonylase/synthase complex, identified by Gionfriddo et al. [13]

345   as the closest and most complete relative to currently unresolved HgcA structure.

17

346 **Biogeography of HgcA.**

347 Relative HgcA abundances in *Tara* Oceans samples were obtained from the Ocean Gene

348 Atlas[37]. We screened 243 metagenomes from 68 sites covering the World Ocean except

349 Arctic, sampled at different depths from surface to 500 m-depth, covering six different size

350 fractions ranging from 0 to 3 μm. Environmental data were obtained from Pesant et al. [46]

351 (Supplementary Table 2). For the following analysis, we considered two depths classes

352 (surface samples < 120 m-depth, subsurface samples > 120 m-depth), two particle size

353 fractions (< 5 μm, < 0.8 μm), two oxic states (oxic: $O_2 > 10$ μM, suboxic: $O_2 < 10$ μM).

354 HgcA relative abundance was calculated as follows: the length-normalized count of genes

355 read was divided by the median of the length-normalized counts of a set of ten universal

356 single copy marker genes [47,48]. Thus, relative abundance represents the fraction of bacteria

357 harbouring the *hgcA* gene within the assembled genomes. A heatmap of the relative gene

358 abundances in *Tara* Oceans samples was generated in R [49] using the heatmap.2 function in

359 the ggplot CRAN library. Dendrograms were computed using hclust default parameters from

360 Ward distance index based on presence/absence of the genes ('binary' option). Genes were

361 clustered into three groups (Cluster 1, Cluster 2 and Cluster 3) according to their abundance

362 pattern on the heatmap. The geographic origin of the *hgcA* genes retrieved from the *Tara*

363 Oceans samples was plotted on a global map using the "mapplots" R package. At each

364 station, the cumulated abundance and phylogenetic affiliation of the retrieved *hgcA* genes

365 were represented on the map by the size and colour of the points. Cluster distribution was

366 also plotted against depth and oxygen concentration at each station to depict the

367 environmental conditions where each Cluster flourishes (Supplementary Figure 2). Tracks of

368 MeHg records from previous campaigns were searched in the literature (Supplementary Text

369 1) and georeferenced on the map.

370

371   References

372   1.   Outridge, P. M., Mason, R. P., Wang, F., Guerrero, S. & Heimbürger-Boavida, L. E.

373   Updated Global and Oceanic Mercury Budgets for the United Nations Global Mercury

374   Assessment 2018. *Environ. Sci. Technol.* **52**, 11466–11477 (2018).

375   2.   Gilmour, C. C. *et al.* Mercury methylation by novel microorganisms from new

376   environments. *Environ. Sci. Technol.* **47**, 11810–11820 (2013).

377   3.   Gilmour, C. C. *et al.* Sulfate-Reducing Bacterium Desulfovibrio desulfuricans ND132 as

378   a Model for Understanding Bacterial Mercury Methylation. *Appl. Environ. Microbiol.*

379   **77**, 3938–3951 (2011).

380   4.   Mason, R. P. *et al.* Mercury biogeochemical cycling in the ocean and policy

381   implications. *Environ. Res.* **119**, 101–117 (2012).

382   5.   Parks, J. M. *et al.* The genetic basis for bacterial mercury methylation. *Science (80-.).*

383   **339**, 1332–1335 (2013).

384   6.   Schlitzer, R. *et al.* The GEOTRACES Intermediate Data Product 2017. *Chem. Geol.* **493**,

385   210–223 (2018).

386   7.   Ortiz, V. L., Mason, R. P. & Evan Ward, J. An examination of the factors influencing

387   mercury and methylmercury particulate distributions, methylation and demethylation

388   rates in laboratory-generated marine snow. *Mar. Chem.* **177**, 753–762 (2015).

389   8.   Bianchi, D., Weber, T. S., Kiko, R. & Deutsch, C. Global niche of marine anaerobic

390   metabolisms expanded by particle microenvironments. *Nat. Geosci.* 1–6 (2018).

391   doi:10.1038/s41561-018-0081-0

392   9.   Lehnherr, I., St. Louis, V. L., Hintelmann, H. & Kirk, J. L. Methylation of inorganic

393   mercury in polar marine waters. *Nat. Geosci.* **4**, 298–302 (2011).

394   10.   Blum, J. D. Marine mercury breakdown. *Nat. Geosci.* **4**, 139–140 (2011).

395   11.   Masbou, J. *et al.* Carbon Stable Isotope Analysis of Methylmercury Toxin in Biological

396   Materials by Gas Chromatography Isotope Ratio Mass Spectrometry. *Anal. Chem.* **87**,

397   11732–11738 (2015).

398   12.   Podar, M. *et al.* Global prevalence and distribution of genes and microorganisms

399   involved in mercury methylation. *Sci. Adv.* **1**, (2015).

400   13.   Gionfriddo, C. M. *et al.* Microbial mercury methylation in Antarctic sea ice. *Nat.*

401   *Microbiol.* **1**, 16127 (2016).

402   14.   Sunagawa, S. *et al.* Ocean plankton. Structure and function of the global ocean

403   microbiome. *Science* **348**, 1261359 (2015).

404   15.   Smith, S. D. *et al.* Site-directed mutagenesis of HgcA and HgcB reveals amino acid

405   residues important for mercury methylation. *Appl. Environ. Microbiol.* **81**, 3205–17

406   (2015).

407   16.   Lücker, S., Nowka, B., Rattei, T., Spieck, E. & Daims, H. The Genome of Nitrospina

408   gracilis Illuminates the Metabolism and Evolution of the Major Marine Nitrite

409   Oxidizer. *Front. Microbiol.* **4**, 27 (2013).

410   17.   Spieck, E., Keuter, S., Wenzel, T., Bock, E. & Ludwig, W. Characterization of a new

411   marine nitrite oxidizing bacterium, Nitrospina watsonii sp. nov., a member of the

412   newly proposed phylum "Nitrospinae". *Syst. Appl. Microbiol.* **37**, 170–176 (2014).

413   18.   Cao, J. *et al.* Pseudodesulfovibrio indicus gen. nov., sp. nov., a piezophilic sulfate-

414   reducing bacterium from the Indian Ocean and reclassification of four species of the

415   genus Desulfovibrio. *Int. J. Syst. Evol. Microbiol.* **66**, 3904–3911 (2016).

416   19.   Gilmour, C. C. *et al.* Sulfate-reducing bacterium Desulfovibrio desulfuricans ND132 as

417   a model for understanding bacterial mercury methylation. *Appl. Environ. Microbiol.*

418   **77**, 3938–51 (2011).

419   20.   Bae, H.-S., Dierberg, F. E. & Ogram, A. Syntrophs dominate sequences associated with

420   the mercury methylation-related gene hgcA in the water conservation areas of the

421   Florida Everglades. *Appl. Environ. Microbiol.* **80**, 6517–26 (2014).

422   21.   Sorokin, D. Y., Tourova, T. P., Mußmann, M. & Muyzer, G. Dethiobacter alkaliphilus

423   gen. nov. sp. nov., and Desulfurivibrio alkaliphilus gen. nov. sp. nov.: two novel

424   representatives of reductive sulfur cycle from soda lakes. *Extremophiles* **12**, 431–439

425   (2008).

426   22.   Tidjani Alou, M. *et al.* Gut Bacteria Missing in Severe Acute Malnutrition, Can We

427   Identify Potential Probiotics by Culturomics? *Front. Microbiol.* **8**, 899 (2017).

428   23.   Regnell, O. & Watras, C. J. Microbial Mercury Methylation in Aquatic Environments: A

429   Critical Review of Published Field and Laboratory Studies. *Environ. Sci. Technol.* **53**, 4–

430   19 (2019).

431   24.   Sela-Adler, M. *et al.* Co-existence of Methanogenesis and Sulfate Reduction with

432   Common Substrates in Sulfate-Rich Estuarine Sediments. *Front. Microbiol.* **8**, 766

433   (2017).

434   25.   Pak, K. R. & Bartha, R. Mercury methylation by interspecies hydrogen and acetate

435   transfer between sulfidogens and methanogens. *Appl. Environ. Microbiol.* **64**, 1987–

436   1990 (1998).

437   26.   Roullier, F. *et al.* Particle size distribution and estimated carbon flux across the

438   Arabian Sea oxygen minimum zone. *Biogeosciences* **11**, 4541–4557 (2014).

439   27.   Gilmour, C. C., Bullock, A. L., McBurney, A., Podar, M. & Elias, D. A. Robust mercury

440   methylation across diverse methanogenic Archaea. *MBio* **9**, (2018).

441   28.   Zhou, J., Riccardi, D., Beste, A., Smith, J. C. & Parks, J. M. Mercury Methylation by

442   HgcA: Theory Supports Carbanion Transfer to Hg(II). *Inorg. Chem.* **53**, 772–777 (2014).

443   29.   Levipan, H. A., Molina, V. & Fernandez, C. Nitrospina-like bacteria are the main drivers

444   of nitrite oxidation in the seasonal upwelling area of the Eastern South Pacific (Central

445   Chile ~36°S). *Environ. Microbiol. Rep.* **6**, 565–573 (2014).

446   30.   Füssel, J. *et al.* Nitrite oxidation in the Namibian oxygen minimum zone. *ISME J.* **6**,

447   1200–1209 (2012).

448   31.   Beman, J. M., Leilei Shih, J. & Popp, B. N. Nitrite oxidation in the upper water column

449   and oxygen minimum zone of the eastern tropical North Pacific Ocean. *ISME J.* **7**,

450   2192–2205 (2013).

451 32. Ngugi, D. K., Blom, J., Stepanauskas, R. & Stingl, U. Diversification and niche

452   adaptations of Nitrospina-like bacteria in the polyextreme interfaces of Red Sea

453   brines. *ISME J.* **10**, 1383–1399 (2016).

454 33. Daims, H., Lücker, S. & Wagner, M. A New Perspective on Microbes Formerly Known

455   as Nitrite-Oxidizing Bacteria. *Trends Microbiol.* **24**, 699–712 (2016).

456 34. Goñi-Urriza, M. *et al.* Relationships between bacterial energetic metabolism, mercury

457   methylation potential, and hgcA/hgcB gene expression in Desulfovibrio

458   dechloroacetivorans BerOc1. *Environ. Sci. Pollut. Res.* **22**, 13764–13771 (2015).

459 35. Reed, D. C., Algar, C. K., Huber, J. A. & Dick, G. J. Gene-centric approach to integrating

460   environmental genomics and biogeochemical models. *Proc. Natl. Acad. Sci. U. S. A.*

461   **111**, 1879–84 (2014).

462 36. Eddy, S. R. Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**, e1002195 (2011).

463 37. Villar, E. *et al.* The Ocean Gene Atlas: exploring the biogeography of plankton genes

464   online. *Nucleic Acids Res.* **46**, W289–W295 (2018).

465 38. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–

466   2069 (2014).

467 39. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview

468   Version 2—a multiple sequence alignment editor and analysis workbench.

469   *Bioinformatics* **25**, 1189–1191 (2009).

470   40.   Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment

471   search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

472   41.   Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7:

473   improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

474   42.   Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: a tool for automated

475   alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–

476   1973 (2009).

477   43.   Sanchez, R. *et al.* Phylemon 2.0: a suite of web-tools for molecular evolution,

478   phylogenetics, phylogenomics and hypotheses testing. *Nucleic Acids Res.* **39**, W470–

479   W474 (2011).

480   44.   Gao, S. *et al.* Evolview v2: an online visualization and management tool for

481   customized and annotated phylogenetic trees. *Nucleic Acids Res.* **44**, W236–W241

482   (2016).

483   45.   Mukherjee, S. *et al.* Genomes OnLine Database (GOLD) v.6: data updates and feature

484   enhancements. *Nucleic Acids Res.* **45**, D446–D456 (2017).

485   46.   Pesant, S. *et al.* Open science resources for the discovery and analysis of Tara Oceans

486   data. *Sci. Data* **2**, 150023 (2015).

487   47.   Mende, D. R., Sunagawa, S., Zeller, G. & Bork, P. Accurate and universal delineation of

488   prokaryotic species. *Nat. Methods* **10**, 881–884 (2013).

489   48.   Sunagawa, S. *et al.* Metagenomic species profiling using universal phylogenetic

490   marker genes. *Nat. Methods* **10**, 1196–1199 (2013).

491   49.   R Development Core Team. R: A Language and Environment for Statistical Computing.

492   (2008).

493

494

Acknowledgements

Author contributions

501   E.V., L.C. and L.E.H.B. wrote the manuscript. E.V. performed the bioinformatic analyses with

502   the scientific support of L.C.

503

504