Universitat Autònoma de Barcelona

CONTRIBUTION OF MARINE HETEROTROPHIC CULTURED BACTERIA TO MICROBIAL DIVERSITY AND MERCURY DETOXIFICATION

Isabel Sanz Sáez

Contribution of marine heterotrophic cultured bacteria to microbial diversity and mercury detoxification

Isabel Sanz Sáez PhD Thesis

Tesi doctoral
Programa de doctorat en Microbiologia

Universitat Autònoma de Barcelona
Departament de Genètica i Microbiologia

Novembre 2020

CONTRIBUTION OF MARINE HETEROTROPHIC CULTURED BACTERIA TO MICROBIAL DIVERSITY AND MERCURY DETOXIFICATION

Contribución de las bacterias marinas heterótrofas cultivables en la diversidad bacteriana y detoxificación del mercurio

Contribució dels bacteris marins heterotròfics cultivables en la diversitat microbiana i detoxificació del mercuri

Isabel Sanz Sáez
Departament de Biologia Marina i Oceanografia,
Institut de Ciències del Mar (ICM-CSIC)

Directores:

Dra. Silvia González Acinas
Departament de Biologia Marina i Oceanografia,
Institut de Ciències del Mar (ICM-CSIC)

Dra. Olga Sánchez Martínez
Departament de Genètica i Microbiologia,
Universitat Autònoma de Barcelona (UAB)

Tutora acadèmica:

Dra. Olga Sánchez Martínez
Departament de Genètica i Microbiologia,
Universitat Autònoma de Barcelona (UAB)

Barcelona, November 8th 2020

During the realization of this thesis, Isabel Sanz Sáez held a Ph.D Fellowship FPU (FPU14/03590) funded by the Spanish Ministry of Universities and was a student of UAB's PhD program in Microbiology.

Cover and editorial design and execution by Shook Studio (@shookstudio).

A todos los que me han acompañado durante esta etapa.

"Nada en la vida es para ser temido, es sólo para ser comprendido.
Ahora es el momento de entender más, de modo que podamos temer menos."
- Marie Curie

AGRADECIMIENTOS

Ha llegado uno de los momentos más esperados desde que empecé a escribir la tesis, aunque des del primer momento sabia que escribir los agradecimientos seria una de las partes que más me costaría. Expresar todo el agradecimiento que siento por todas aquellas personas que me han acompañado hasta aquí en unas líneas no es tarea fácil.

Quiero empezar dando las gracias a mis superjefas. Olga, gracias por abrirme las puertas al mundo de la investigación, ya que des del momento en el que entré en tu despacho a que me explicaras en que podría consistir mi máster (por allí el 2014) supe que la microbiología marina robaría mi corazón científico. Silvia, gracias por transmitirme tu pasión por la ciencia, tu energía incombustible y por hacerme ver que todo con empeño y cariño se consigue. A las dos, muchísimas gracias por toda la sabiduría que me habéis transmitido durante toda esta etapa, por todas las oportunidades que me habéis ofrecido durante estos años para ir de congresos, estancias, cursos...pero sobretodo, por vuestras risas y optimismo en todas las reuniones y conversaciones, ya que gracias a ello cogía siempre fuerzas para seguir en esta batalla llamada tesis.

Gracias también a todos los IPs, postdocs y técnic@s del ICM. En repetidas veces he escuchado que el ICM es como una gran familia, y sin ninguna duda puedo corroborar que eso es así. Me siento privilegiada por poder formar parte de esta familia, donde el talento, pero sobretodo la gran cantidad de buenas y grandísimas personas son el elemento principal.

En especial quiero dar las gracias al superteam Acinas, ya que a vuestro lado nunca faltan las risas, los buenos consejos e interesantísimas conversaciones (no solo científicas). Pablo, gracias por ayudarme siempre en mis momentos de crisis con la bioinformática y hacerme reír con tu característico sentido del humor. Carla, gracias por haber sido mi compañera incansable durante la etapa de experimentos del mercurio (sobretodo las inacabables curvas de crecimiento!!). Eli, gracias por transmitirme orden y calma en el laboratorio (cuantas colonias habremos picado!). Andrea, aunque solo hemos coincidido en esta última etapa de mi tesis, es un placer trabajar a tu lado.

No me olvido tampoco de Shini, gracias por acogerme en tu laboratorio de Zúrich y hacerme formar parte de tu equipo, ni de Guillem, no solo aprendí muchísimo contigo sino que hiciste que adentrarme en el mundo de la bioinformática fuese un reto agradable a superar.

Esta tesis no seria lo mismo sin mis compañer@s de doctorado. Anna, crec que aquesta tesi no hagués sigut la mateixa sense la teva companyia. Gràcies per tants bons moments de riures, per aguantar-me durant les meves bogeries transitòries al despatx, per les fotos al Mac utilitzant els filtres de deformació de les cares (cada vegada que en faig memòria no puc evitar riure a riallades!!), però sobretot per haver trobat una amiga tan especial durant aquest camí. Marta R., vull donar-te les gràcies no només per ser la millor companya de grup, de despatx o de congressos que hagués pogut desitjar, sinó per deixar que interrompés la teva concentració al despatx, ajudar-me sempre que ho he necessitat i donar-me sempre aquella peça clau que em faltava per seguir endavant en els meus experiments o anàlisis. Durant aquesta última etapa de la tesi, he trobat molt a faltar les meves companyes del P5-a, els cafès de les tardes, les xerrades terapèutiques i no ens oblidem del “Fashion Oracle”, gràcies per

tot. Patri, que haría yo sin ti en el ICM! Gracias por las tardes de zumba, los marujeos que tanto nos gustan, las risas por cualquier cosa, las visitas al despacho, por las canciones compartidas de spotify que nos recuerdan la una a la otra, pero por encima de todo, por escucharme siempre que lo he necesitado y darme el consejo justo y necesario para poder coger energías y seguir adelante. Maria, crec que no hagués superat amb èxit aquesta etapa si no t'hagués conegut. Gràcies per compartir les penúries a l'hora d'omplir "papeleos" que mai s'acaben per la beca FPU, per enviar-me àudios sempre en el moment idoni i per ser l'amiga que et diu que tu pots amb tot. Yaiza, mi "twin" rubia, gracias por acogerme nada más llegar al ICM, por esos abrazos que tanto nos gustan y por hacer que los largos días de campana se hagan amenos. Y gracias al resto del supergrupo "Malagente": Janire, Marina P., Adrià, Aleix, Manu, Andreu, Carlota, Claudia, Gui, Dani, Marina B., Guillem, Stefano, Charlie, Joan, Ainara, Miguel y tantos otros que han ido y venido. Gracias por las magníficas horas de la comida donde siempre había hueco en la mesa para uno más, por esas conversaciones que te hacían desconectar de la tesis y tus experimentos, y por las magníficas tardes de Friday Beer!

Fuera del ICM, también he contado con el apoyo indispensable de mis amigos de toda la vida. Laia, Gemma y Carlos, gracias por sacarme a respirar aire fresco siempre que lo he necesitado, por escuchar mis "batallitas" de la tesis, por echarme de menos cuando he estado fuera, por estar ahí en los buenos y, sobretodo, en los malos momentos, por todas las risas, y por mil y una cosas que hemos compartido durante estos 26 años que nos conocemos (madre mía!) y muchas más que faltan por llegar.

A mis compis de carrera, Sonia y Mireia. Quien nos iba a decir cuando empezamos en esto de la Microbiología que acabaríamos haciendo un doctorado! La carrera no hubiese sido lo mismo sin vuestra compañía, con esos apuntes compartidos, esas prácticas en las que a veces íbamos un poco perdidas, y durante el doctorado, aunque no tan frecuente como nos gustaría, esas quedadas para ponernos al día y sentirnos tan unidas como antes.

A mi familia, por supuesto, por interesarse y preocuparse siempre por mi, y en especial a mis padres, Antonia y Teo, que siempre han luchado lo máximo posible para darnos a mi y a mi hermana el mejor de los futuros. Mama, gracias por animarme siempre, por tu positivismo, y por darme ese empujón que necesitaba para elegir una carrera, en este caso la Microbiología, una decisión de la que no me arrepiento. Papa, gracias por apoyarme en cada una de mis decisiones y pasos tanto en la vida profesional como en la personal por muy locas que te pudieran parecer a veces. Y como no, tengo que dar las gracias a mi hermanita Laura (siempre serás mi hermanita por muy mayores que nos hagamos!). Gracias por tener siempre una sonrisa para mi, por estar ahí siempre, y por esa conexión tan especial que tenemos con la que con solo una mirada sabemos como estamos y lo que pensamos. Estas líneas no son suficientes para describir lo afortunada y feliz que me siento de tener unos padres y una hermana como vosotros, os quiero muchísimo y esta tesis es tanto mía como vuestra.

También quiero darle las gracias a mi familia adoptiva, Ana, Jose, Kike, Iván y el más pequeño de todos, a Óscar, gracias por hacerme sentir siempre en casa y acogerme como una más de la familia.

No me puedo olvidar de mi compañero peludo, Kovu. Solo tu compañía (sobretodo durante estos últimos meses caóticos) ya me transmite tranquilidad y calma. Gracias por tus visitas pidiendo mimos, por tus locuras perrunas y porque eres tu el que me saca a mi de paseo y no al revés.

Y a ti, mi "Chuloplayas" favorito. Esta tesis va especialmente dedicada a ti. Gracias por aguantarme en mis días buenos y en mis días malos (que no son pocos), por creer siempre en mi, por ser participe de mis locuras, por hacerme ver que para todo problema hay una solución, por ser mi psicólogo personal, por hacerme reír (hasta cuando parece imposible sacarme una sonrisa), por ser mi mejor amigo, por hacerme feliz, por todo eso y mucho más que me es imposible describir con palabras...Gracias! Conocerte ha sido una de las mejores cosas que me ha pasado en la vida. Te quiero muchísimo y estoy impaciente por ver que aventuras nos esperan en el futuro.

CONTENTS

Summary
Resumen
Resum

SUMMARY

The world's oceans sustain the life for an estimated total of 10^{29} microbial cells. Marine bacteria are responsible for most part of the ocean respiration and are key in most biogeochemical cycles of the Earth. Accordingly, the study of the bacterial diversity present in different marine ecosystems is essential, and having access to their genomes through isolation or genomic centric studies is important to decipher their metabolic potential.

Isolation of marine microorganisms is fundamental to gather information about their physiology, ecology and genomic content. To date, most of the bacterial isolation efforts have focused on the photic ocean leaving the deep ocean less explored. In this thesis, standard plating techniques allowed to create a marine culture collection of heterotrophic bacteria (MARINHET). More than 2000 isolates were retrieved from samples collected from a variety of oceanographic regions, from different depths including surface, mesopelagic and bathypelagic waters, and also covering different seasons and years. **Chapter 1** describes the taxonomy, the phylogenetic diversity and the biogeography of culturable heterotrophic marine bacteria, and reveals that an important percentage of the strains (37 %) are 100 % identical in their partial 16S rRNA gene between photic and aphotic layers. In addition, we identified *Alteromonas* and *Erythrobacter* genera as the most frequently retrieved heterotrophic bacteria from the ocean in standard marine agar medium.

It is a long-standing observation that traditional culture techniques only retrieve a small fraction of the microbial diversity found in natural environments including marine ecosystems, what is known as "the great plate count anomaly". In addition, most of the retrieved isolates belong to the so-called rare biosphere. However, we do not know if these patterns, usually described for bacteria living in the surface/photic ocean, also apply for the deep ocean bacteria. In **Chapter 2** of this thesis, I combined results from culture-dependent and -independent techniques by comparing the 16S rRNA partial sequences of the MARINHET isolates with 16S rRNA amplicon Illumina TAGs (16S iTAGs) and metagenomic TAGs (miTAGs) from surface, mesopelagic and bathypelagic samples globally distributed. A high proportion of bacteria inhabiting the deep ocean could be retrieved by pure culture techniques and a significant fraction of the isolates preferred a lifestyle attached to particles. Additionally, I revised the axiom that "less than 1% of bacteria can be cultured", finding variability between mesopelagic and bathypelagic samples, where up to 3% of the cells could be cultured.

Bacterial isolates also represent a valuable genetic reservoir for biotechnology applications, such as bioremediation strategies of marine polluted environments. Mercury is one of the most toxic heavy metals in the planet and its most dangerous form, methylmercury (MeHg), is being bioaccumulated in the marine food web. However, little is known about the tolerance capacity and phenotypic characterization of marine bacteria codifying the mercury resistance operon (*mer* operon). **Chapter 3** describes the functional screening of *merA* and *merB* genes, which are key in the mercury detoxification process, in well know marine genera with described genetic potential for mercury detoxification, such as *Alteromonas* and *Marinobacter*. I reported that the *merAB* genes from these two genera are widely distributed in different oceanographic regions and depths. In addition, I selected a promising candidate, phylogenetically affiliated to *Alteromonas mediterranea*, for future bioremediation studies due to its high tolerance and degradation ability of different mercury forms.

Overall, this thesis provides a significant advance in the knowledge of the diversity of culturable heterotrophic bacteria from the less explored deep ocean, pointing out that heterotrophic cultured bacteria represents a significant fraction of the total bacterial diversity in deeper waters, while identifying some key marine bacteria with interesting potential for mercury bioremediation.

RESUMEN

Los océanos contienen aproximadamente un total de 10^{29} células microbianas. Las bacterias marinas son responsables de la mayor parte de la respiración que se produce en el océano y son esenciales en los ciclos biogeoquímicos de la Tierra. Estudiar la diversidad bacteriana de los ecosistemas marinos y tener acceso a los genomas mediante estudios dependientes e independientes de cultivo es importante para descifrar el potencial metabólico de las bacterias marinas.

Los cultivos nos aportan información sobre la fisiología bacteriana, ecología y contenido genómico, pero la mayoría de los esfuerzos en aislar bacteria marinas provienen de la zona fótica del océano, dejando las profundidades marinas menos exploradas. En esta tesis, técnicas estándar de cultivo han permitido crear una colección marina de bacterias heterótrofas (MARINHET), compuesta por más de 2000 aislados, recuperados de varias regiones oceanográficas, de varias profundidades (superficie, mesopelágico y batipelágico), y cubriendo varias estaciones y años. El **Capítulo 1** describe su taxonomía, diversidad filogenética y biogeografía y revela que un 37% de las cepas son 100% idénticas en la secuencia parcial del gen ribosomal 16S (16S rRNA) entre la zona fótica (superficie) y afótica (mesopelágico y batipelágico). Además, hemos identificado *Alteromonas* y *Erythrobacter* entre los géneros marinos heterótrofos más comunes que recuperamos en cultivo usando un medio marino estándar.

Las técnicas tradicionales de cultivo generalmente solo recuperan una fracción pequeña de las comunidades bacterianas naturales, fenómeno conocido como "la gran anomalía de recuento en placa" y muchas de las cepas que se aíslan pertenecen a la biosfera rara. Sin embargo, no conocemos si estos patrones, normalmente descritos para las bacterias de superficie, también se aplican en las profundidades. En el **Capítulo 2** he combinado resultados obtenidos mediante técnicas dependientes e independientes de cultivo comparando las secuencias del 16S rRNA de la colección MARINHET contra los fragmentos de secuenciación masiva del 16S rRNA (de amplicones y metagenomas), obtenidos de muestras globalmente distribuidas y de diferentes profundidades. Una mayor proporción de las bacterias del océano profundo son cultivables y una fracción importante de los aislados tiene preferencia a un estilo de vida adherido a partículas. Además, confirmamos que el dogma "menos del 1% de las bacterias son cultivables" deber ser revisado ya que encontramos variabilidad en las muestras de profundidad, donde hasta un 3% de las células se han podido aislar.

Los aislados bacterianos son un excelente material para aplicaciones biotecnológicas, como la biorremediación de zonas marinas contaminadas. El mercurio es un metal pesado tóxico y su forma más peligrosa, el metilmercurio (MeHg), se bioacumula en la cadena trófica marina. No obstante, se conoce muy poco la tolerancia de bacterias marinas frente al mercurio o la fisiológia de aquellas cepas que codifican los genes de resistencia (operón *mer*). El **Capítulo 3** describe los resultados del mapeo funcional de los genes *merA* y *merB*, clave en la detoxificación, en una fracción de la colección MARINHET. Nos centramos en dos géneros marinos, con un potencial genético para la degradación del mercurio previamente descrito en la literatura, como son *Alteromonas* y *Marinobacter*. Desvelamos que los genes *merAB* están ampliamente distribuidos en diferentes regiones oceanográficas y en varias profundidades. Adicionalmente, hemos seleccionado una cepa de *Alteromonas mediterranea* para futuros estudios de biorremediación debido a su alta tolerancia y capacidad de degradación de diferentes formas de mercurio.

En conjunto, esta tesis proporciona avances significativos en el área de conocimiento de la diversidad bacteriana heterótrofa cultivable en zonas menos exploradas como es el océano profundo. Describimos ciertos patrones ecológicos, señalamos que las bacterias heterótrofas marinas son un componente importante en las comunidades bacterianas de las profundidades e identificamos cepas para futuros estudios de biorremediación del mercurio.

RESUM

Els oceans conten aproximadament un total de 10^{29} cèl·lules microbianes. Els bacteris marins són responsables de la major part de la respiració que es produeix en l'oceà i són essencials en els cicles biogeoquímics de la Terra. Estudiar la diversitat bacteriana dels ecosistemes marins i tenir accés als genomes mitjançant estudis dependents o independents de cultiu és important per desxifrar el potencial metabòlic dels bacteris marins.

El cultius ens aporten informació sobre la fisiologia bacteriana, ecologia i contingut genòmic, però la majoria dels esforços en aïllar bacteris marins provenen de la zona fòtica de l'oceà, deixant les profunditats marines menys explorades. En aquesta tesi, tècniques estàndard de cultiu han permès crear una col·lecció marina de cultius de bacteris heterotròfics (MARINHET), composada per més de 2000 aïllats, recuperats de diverses regions oceanogràfiques, de diverses profunditats (superfície, mesopelàgic i batipelàgic) i cobrint diverses estacions i anys. El **Capítol 1** descriu la seva taxonomia, diversitat filogenètica i biogeografia i revela que un 37% de les soques aïllades són 100% idèntiques en el gen ribosomal 16S (16S rRNA) entre la zona fòtica (superfície) i afòtica (mesopelàgic i batipelàgic). A més, hem identificat *Alteromonas* i *Erythrobacter* entre els gèneres marins heterotròfics més comuns que recuperem en cultiu usant un medi marí estàndard.

Les tècniques tradicionals de cultiu generalment només recuperen una fracció petita de les comunitats bacterianes naturals, fenomen conegut com "la gran anomalia de recompte en placa" i moltes de les soques que s'aïllen pertanyen a la biosfera rara. Tanmateix, no coneixem si aquests patrons, normalment descrits per als bacteris de la zona fòtica, també s'apliquen als bacteris de les profunditats. En el **Capítol 2** he combinat els resultats obtinguts mitjançant tècniques dependents i independents de cultiu comparant les seqüències del 16S rRNA de la col·lecció MARINHET contra fragments de seqüenciació massiva del 16S rRNA, tant d'amplicons com de metagenomes, obtinguts de mostres globalment distribuïdes i de diferents profunditats. Trobem que una major proporció dels bacteris de l'oceà profund són cultivables i una fracció important dels aïllats té preferència a un estil de vida adherit a partícules. A més, confirmem que el dogma "menys de l'1% dels bacteris són cultivables" ha de ser revisat ja que trobem variabilitat en les mostres de l'oceà profund, on fins a un 3% de les cèl·lules han sigut cultivades.

Els aïllats bacterians són un excel·lent material per a aplicacions biotecnològiques com la bioremediació de zones marines contaminades. El mercuri és un metall pesat tòxic i la seva forma més perillosa, el metilmercuri (MeHg), es bioacumula a la cadena tròfica marina. No obstant això, es coneix molt poc la tolerància de bacteris marins enfront del mercuri o la fisiologia d'aquelles soques que codifiquen l'operó dels gens de resistència (operó *mer*). El **Capítol 3** descriu els resultats del mapeig funcional dels gens *merA* i *merB*, clau en la detoxificació, en una fracció de la col·lecció MARINHET. Ens centrem en dos generes marins, amb un potencial genètic per a la degradació del mercuri conegut, com són *Alteromonas* i *Marinobacter*. Revelem que els gens *merAB* estan àmpliament distribuïts en diferents regions oceanogràfiques i en diverses profunditats. Addicionalment, hem seleccionat una soca d' *Alteromonas mediterranea* per a futurs estudis de bioremediació degut a la seva alta tolerància i capacitat de degradació de diferents formes de mercuri.

En conjunt, aquesta tesi proporciona avanços significatius en l'àrea de coneixement de la diversitat bacteriana heterotròfica cultivable en zones menys explorades com l'oceà profund. Descrivim certs pa-

trons ecològics, assenyalem que els bacteris heterotròfics marins són un component important en les comunitats bacterianes de les profunditats i identifiquem soques per a estudis futurs de bioremediació del mercuri.

Introduction

INTRODUCTION

The marine environment

The oceans, covering 3.6 x 10^8 km^2 (71% of the Earth's surface), with an average depth of 4,000 m and reaching up to 11,000 m in the Marianas Trench, represent the largest habitable ecosystems on Earth. The seas and oceans around the world are interconnected environments where microorganisms live in a network with other organisms. Microbial life fills almost an unlimited number of ecological niches across vertical, horizontal and latitudinal gradients, which are linked to different biotic and abiotic factors determining the interactions between organisms. Vertically, we found the pelagic environment (ocean water column) and the benthic environment (seafloor). The pelagic realm, where this thesis is focused, can be divided on the basis of the water column depth into the epipelagic (0-200 m), mesopelagic (200-1000 m), bathypelagic (1000-4000 m), abyssopelagic (4000-6000 m) and hadalpelagic zones (6000-11000 m) (Tait and Dipper, 1998a; Webb, 2020). The epipelagic zone is the region where enough light penetrates to support photosynthesis, so it is also called the photic layer. This layer presents diurnal and seasonal changes of light intensity and temperature. Nutrients values in this layer are highly variable depending on season stratification and primary production levels (Bolhuis and Cretoiu, 2016). On the other hand, the mesopelagic zone, also called the "twilight" zone, is characterized by an increase of hydrostatic pressure, diminished light that prevents photosynthesis activity, high inorganic nutrients concentrations and episodic food supply. It is not a homogeneous layer, but encompasses strong gradients in environmental parameters, particularly at the interface with the epipelagic layer, and in the regions with oxygen minimum zone (OMZ) areas (Robinson et al., 2010), defined by low concentrations of oxygen and high concentrations of nitrate and phosphate (Bristow et al., 2017). Finally, the bathypelagic, abyssopelagic and hadalpelagic layers constitute the aphotic ocean and are characterized by darkness, low temperatures, high pressures and nutrient enriched waters (Tait and Dipper, 1998a; Bolhuis and Cretoiu, 2016). In this thesis, however, we will use the terms photic ocean to refer to the epipelagic layer, and aphotic or deep ocean to refer to the mesopelagic and bathypelagic layers, which were the main regions sampled. Horizontally, the oceans are not uniform neither. The most comprehensive and widely accepted classification of the epipelagic or photic zone is the Longhurst's partitions of the ocean (Longhurst, 2007) into four major biomes (polar, westerlies, trade winds, coastal) and into approximately 50 ocean biogeochemical provinces. These partitions are mainly based on the spatial variability of physical properties, such as temperature, salinity, mixing state, and observations in chlorophyll concentration and estimates of primary production. The boundaries between provinces are generally persistent, but are also spatially and temporally variable seasonally and interannually (Vichi et al., 2011). On the other hand, these horizontal classifications into different biogeographic areas are currently not so well defined in the deep ocean (Sutton et al., 2017), which is also composed by different deep water-masses that hardly mix and maintain distinctive physical features and organic and inorganic nutrients (Rahmstorf, 2003; Teira et al., 2006; Agogué et al., 2011).

The connectivity between the different oceanic environments found in the vertical and horizontal axis is mainly influenced by the combined effect of temperature, pressure and salinity, which determines water density. Higher temperatures and lower salinity of the surface layer of the ocean, compared to the waters underneath, lead to a vertical stratification of the water column (Talley et al., 2011). These temperature and density stratification of the water column generally prevents the nutrient-rich deep water

from mixing with the surface water. However, under certain conditions this nutrient-rich deep water may be brought to the surface through the process of convectional mixing or upwelling (Tait and Dipper, 1998b; Wells, 2015). In addition, connectivity along the vertical axis is also promoted by particles that originate in the surface layers and sink into the deep ocean mobilizing organic and inorganic nutrients between layers. An example could be the sedimentation of particulate organic matter (POM) exported from the photic layers, in various forms like detrital matter, fecal pellets or dead cells, which is called the biological carbon pump (Raven and Falkowski, 1999; Ducklow et al., 2001; Herndl and Reinthaler, 2013; Bopp et al., 2015). Moreover, connectivity between surface waters is assured due to winds originated by the differential heating of air masses and the rotation of the Earth (Wells, 2015). In the deep ocean, differences in salinity and temperature between water masses create density differences that fuel a relatively slow worldwide circulation of the deep oceanic water, the thermohaline circulation system or the *global conveyor belt* (Rahmstorf, 2003). The deep ocean is also influenced by specific geographical features, such as submarine mountains, that also compartmentalize it into different basins. This may influence water circulation and connectivity as well as organism dispersal (Schauer et al., 2010; Moalic et al., 2012; Salazar et al., 2016).

Therefore, even though the sea may seem rather homogeneous from a superficial point of view, the physical and chemical variations found across oceans determine different biotic areas where diverse organisms dwell, highly adapted to their environmental conditions defined by light, nutrients, and water physical properties.

Marine microbial diversity

Microorganisms, referred here as single-celled organisms, are invisible to the naked eye. They are the closest living descendant of the original forms of life on Earth and all organisms present on the planet may have evolved from these (Hunter-Cevera et al., 2005). Current marine microbes comprise organisms that belong to the three domains of life: Archaea, Bacteria and Eukarya (Woese et al., 1990). However, this thesis is focused in Bacteria, also known, together with Archaea, as prokaryotes due to the lack of nuclear membrane. In the ocean, it is estimated that one milliliter of seawater contains 10^6 prokaryotic cells, while the total number of cells in the global ocean is approximated to 10^{29} cells (Whitman et al., 1998). In addition, marine microbes may account for 70-90% of the total ocean's biomass (Fuhrman and Azam, 1980; Whitman et al., 1998; Pomeroy et al., 2007; Bar-On et al., 2018).

The number of microbial species on Earth is believed to be on the order of 10^{12} (Locey and Lennon, 2016) but there is only around 20,214 species validly classified (Accessed in October 2020, List of Prokaryotic Names with Standing in Nomenclature) (Parte et al., 2020) and from these only 653 different species have been catalogued from marine origin (Mora et al., 2011), a number that may have increased in the last years since new prokaryotic species are continuously described in the International Journal of Systematic and Evolutionary Microbiology. The most abundant bacteria in the oceans belong to phyla *Cyanobacteria*, *Proteobacteria* and *Bacteroidetes* (Amaral-Zettler et al., 2010; Overmann and Lepleux, 2016). Photoautotrophic *Cyanobacteria* as *Prochlorococcus* and *Synechococcus* predominate within photosynthetic bacteria with cell numbers up to 10^5-10^6 per milliliter (Heywood et al., 2006; Scanlan et al., 2009). Among *Proteobacteria*, *Alphaproteobacteria* and specifically the SAR11 clade constitute the dominant heterotrophic bacterial group. The SAR11 clade plays a critical role in carbon cycling, repre-

senting around 25% of the plankton biomass, and it is found throughout the oceans, although it reaches largest numbers in stratified oligotrophic gyres (with low levels of nutrients and biotic activity) (Amaral-Zettler et al., 2010; Giovannoni, 2017). *Alteromonas*, belonging to the *Gammaproteobacteria* class also within the *Proteobacteria* phylum, is perhaps one the most abundant genus in the global bathypelagic ocean (Salazar et al., 2016). Furthermore, *Alteromonas* is among the most common culturable heterotrophic bacteria living in open marine waters all around the world (García-Martínez et al., 2002), and this genus is also thought to be important in the upper ocean as one of the most significant contributors of the dissolved organic carbon (DOC) consumption and nutrient mineralization (Pedler et al., 2014). Finally, the phylum *Bacteroidetes* (formerly known as Cytophaga-Flavobacterium-Bacteroides, or CFB group) is mainly composed of three large classes of bacteria found in a wide variety of environments, but the *Flavobacteriaceae* family is the most abundant in the oceans (Kirchman, 2002; Bowman, 2006). The phylum *Bacteroidetes* may contribute up to 22% of total bacteria in different marine ecosystems (Cottrell and Kirchman, 2000; Alonso-Sáez and Gasol, 2007; Acinas et al., 2015), having a pivotal role in the processing of macromolecules, such as polysaccharides and proteins (Kirchman, 2002; McBride, 2014), and displaying a wide array of different ecotypes both in photic and aphotic layers (Díez-Vives et al., 2019). Additionally, some members of this phylum, together with other phylogenetically distant bacteria such as SAR11, contain proteorhodopsin, which is a light dependent proton pump that enables a photoheterotrophic lifestyle by using light to obtain energy (Giovannoni et al., 2005; González et al., 2008).

Role of marine microbes in biogeochemical cycles

Bacteria are responsible for 95% of the respiration in the oceans (del Giorgio and Duarte, 2002) and are crucial in most key transformations in the cycles of carbon, nitrogen, phosphorous, sulphur, iron and other metals (Whitman et al., 1998; Gasol and Kirchman, 2018). Photosynthetic bacteria (e.g. *Prochlorococcus* and *Synechococcus*) together with eukaryotic microalgae contribute up to 45% of Earth's primary production (Field et al., 1998). A considerable proportion of organic carbon from primary producers can enter the trophic food-webs directly by being consumed by heterotrophic microbes or animals such as copepods, or indirectly through the microbial-loop (Azam, 1998; Fenchel, 2008). The microbial loop can be defined as the transformation of dissolved organic matter (DOM) into living particulate organic matter (POM) by heterotrophic bacteria and, therefore, a way to supply organic matter to higher trophic levels (Azam, 1998; Fenchel, 2008). Between 1 and 40% of primary production can also reach aphotic layers in the form of sinking POM (Ducklow et al., 2001), which can be gradually degraded by heterotrophic microbes (Smith et al., 1992). Nevertheless, sinking POM is not apparently the only source of carbon in the aphotic ocean due to the evidences of a high respiratory activity, which is difficult to reconcile with the rates of supply of organic carbon produced in the photic layer (Arístegui et al., 2002, 2009; del Giorgio and Duarte, 2002; Baltar et al., 2010a; Herndl and Reinthaler, 2013). Thus, alternative sources of carbon has been described including non-sinking particulate organic carbon (non-POC), or *in situ* production by autochthonous microbial chemolithoautotrophs (Baltar et al., 2010b). Hence, microbes are at the base of the oceanic food web due to the transformation of inorganic carbon into organic matter through primary production and the flux of this organic matter through heterotrophic metabolism.

Marine bacteria are also responsible for mediating some biochemical cycles that other organisms are

unable to achieve (Hunter-Cevera et al., 2005). As an example, the nitrogen cycle in the ocean is driven by complex microbial transformations and the biochemical process of fixing nitrogen is confined to a diverse but limited number of bacterial and archaeal lineages (Zehr and Ward, 2002; Voss et al., 2013; Gruber, 2016). Further, some bacteria are involved in the sulphur cycle and in the formation of clouds by cycling compounds such as dimethylsulfoxide into the atmosphere (Voss et al., 2013) influencing planet's climate.

Lastly, marine bacteria are eventually exposed to unfavorable conditions when toxic substances are released into marine ecosystems. Some marine bacteria have developed a genetic machinery to cope with these situations and are involved in the decomposition of different pollutants, such as hydrocarbons (Hassanshahian et al., 2014), plastics (Andrady, 2011), pharmaceutical compounds (Benotti and Brownawell, 2009) or heavy metals (Hanan et al., 2011). Therefore, marine microorganisms not only have a vital role in fueling biogeochemical cycles taking place in the oceans, but also represent a unique genetic reservoir for medical and biotechnological research areas.

How do we study marine bacteria?

The study of Earth's biodiversity, including marine bacterial diversity, is today more important than ever in order to better predict the ecosystems responses to the rapid loss of biodiversity that we are witnessing due to human activities and their effects on climate change. Moreover, microbial diversity studies are the basis to detect and describe those bacteria that could be used as a tool for bioremediation strategies in human impacted ecosystems. Most marine microbial ecologists that study biodiversity aim to get answers to crucial questions such as: (i) how many different bacterial species can be found in a specific environment?, (ii) who are they and how are they distributed?, (iii) which are their functions in the ecosystem? To answer these questions, two main type of methodologies can be applied: culture-dependent and culture-independent techniques.

Culture-dependent techniques

Marine microbial diversity has been traditionally studied by culture-dependent techniques, i.e. those methods that rely on growing bacteria under laboratory conditions. One of the firsts scientists to isolate bacteria in axenic cultures from marine samples was Claude E. Zobell (Zobell, 1941), who is considered one of the founders of modern marine microbiology. Culture-dependent techniques include both more traditional methods as enrichment cultures (Zobell, 1941; Schut et al., 1993; Pinhassi et al., 1997; Kai et al., 2017) and more recent innovative cultivation techniques using advanced technologies, such as microencapsulation or cultivation chips (Ingham et al., 2007; Joint et al., 2010; Gutleben et al., 2017) (Figure 1). Enrichment cultures rely on establishing a certain group of conditions (nutrients, incubation temperatures or light) that selects for specific microorganisms. The most efficient enrichment cultures are those that are able to mimic the resources and conditions for a specific habitat. Even though these enrichment cultures can be a powerful tool, most of the times, we could only retrieve those bacteria that grow fast, usually copiotrophic bacteria (i.e. bacteria adapted to rapid growth upon encountering high substrate levels) that are more prone to grow and utilize the high concentrations of nutrients generally provided in the culture media (Overmann et al., 2017). The complexity to mimic the environmental con-

ditions at the laboratory arise by different factors, such as: (i) the composition of the growth medium and incubation conditions, specially the concentration of nutrients available that generally exceed the ones inherent in the ocean (Giovannoni and Stingl, 2007; Joint et al., 2010), (ii) because other organisms, or the metabolites produced by them, are needed for the growth of certain species (co-culture) (Wang et al., 2014; Overmann et al., 2017), (iii) due to the lifestyle of certain prokaryotes living in association on organic matter aggregates (Azam, 1998) or in symbiosis (Zehr et al., 2008), or (iv) because the physiological status of the cell, as dormancy can limit its availability to be active and grow again (Buerger et al., 2012).

For these reasons, new culture-dependent techniques have been developed in the last years to expand the range of bacteria that can be cultured, like microfluidics (Ma et al., 2014; Boitard et al., 2015) cultivation chips (Ingham et al., 2007; Hesselman et al., 2012; Gao et al., 2013), manipulation of single cells (Ben-Dov et al., 2009; Park et al., 2011), and high-throughput cultivation techniques named "culturomics" (Giovannoni and Stingl, 2007; Lagier et al., 2012). A common feature of all these new techniques is that they are all based on the same principles of re-creating the nutrient conditions of natural environments and overcoming the tendency of rapidly growing cells to overwhelm species that divide less often. The use of these new methodologies allowed the isolation of previously uncultivable groups such as the *Alphaproteobacteria Candidatus* Pelagibacter ubique in the SAR11 clade, probably the most abundant bacteria in the ocean (Morris et al., 2002; Rappé et al., 2002) or *Sphingopyxis alaskensis*, which is one of the best studied oligotrophic bacterium isolated from coastal samples off Alaska (Cavicchioli et al., 2003). Altogether, these culturing depending approaches are still fundamental for microbial ecologists to fully understand the ecology, function and biotechnological potential of microorganisms in marine ecosystems.

Culture-independent techniques

Culture-independent techniques include basic microscopy methodologies as 4′,6-diamidino-2-phenylindole staining (DAPI) or acridine orange which allow the quantification of microorganisms in natural samples (Francisco et al., 1973; Porter and Feig, 1980). However, these techniques usually do not infer much about the physiology or phylogeny of bacteria, and it was not until the emergence of DNA-based techniques, about 30 year ago, that the study of bacterial diversity in marine communities experienced a revolution. Carl Woese and George Fox used the sequencing of the 16S rRNA gene, a marker gene reflective of the evolutionary history of the organisms, to organize the life's diversity within a phylogenetic framework (Woese and Fox, 1977). They were pioneers in the transformation of our understanding of marine microbial diversity, previously based mainly on the knowledge acquired from bacteria that grew in the laboratory. With this new approach the possibility to describe uncultivable bacteria was opened and marine microbial ecologists started to study the taxonomic composition of whole bacterial communities. First, clone libraries containing few tenths of 16S rDNA sequences from environmental samples were developed (Pace et al., 1986) using universal primers and the polymerase chain reaction (PCR). These were later followed by samples with increasing numbers of sequences that once they were analyzed by sequencing the different 16S rRNA genes, revealed the magnitude of the marine prokaryotic diversity for Bacteria (Giovannoni et al., 1990; Ward et al., 1990; Mullins et al., 1995; Acinas et al., 1999) and Archaea (DeLong, 1992; Fuhrman et al., 1992; Massana et al., 2000). Contemporary to clone libraries, fingerprinting molecular techniques existed, including TRFLP (Liu et al., 1997), DGGE

(Muyzer et al., 1993), or ARISA (Fisher and Triplett, 1999), which allowed the characterization of microbial communities based on sequence dissimilarities in the targeted gene without the need of sequencing. Since 2006, high-throughput sequencing (HTS) techniques like 454 or Illumina emerged, and a second revolution in the description of microbial biodiversity started. These sequencing techniques allow the analysis of massive amounts of short DNA sequences from a sample or a combination of samples, characterizing the taxonomical diversity of microbial communities with a resolution at an unprecedented scale (Sogin et al., 2006). With these methods, it also became possible to directly sequence the whole DNA content of environmental samples (i.e. metagenomics) or RNA (metatranscriptomics) without the need for targeting (i.e. amplifying by PCR) a specific gene (Venter et al., 2004) and, then, to obtain not only taxonomic information but also the metabolic potential of microbial communities as well as, for example, the interaction between organisms (Lima-Mendez et al., 2015). Lately, different strategies have been developed in order to link phylogenetic information with metabolism, an essential goal in microbial ecology. These strategies include the use of metagenomic data for building genomes from the DNA content of the whole community (Metagenome-Assembled Genomes, MAGs) (Parks et al., 2017) and the direct amplification of DNA from previously sorted individual cells without the need of cultivation (Single Amplified Genomes, SAGs) (Stepanauskas and Sieracki, 2007). A diagram of the culture-independent techniques revised here is included in Figure 1.

Insights from culture-dependent and culture-independent studies

The great plate count anomaly

The firsts evidences that not all bacteria from a given environment would grow on laboratory media came from microscopy counts, where the number of cells observed far outweighed the number of cells or colonies growing on petri dishes (Hobbie et al., 1977). This phenomenon has been traditionally known as the "great plate count anomaly" and was coined by Staley and Konopka (1985). The recovered proportion of cells using selective media and standard plating techniques, when compared to both microscopy analyses by direct staining and to 16S rRNA sequencing, only represented among 0.001-1% of the community (Kogure et al., 1979; Staley and Konopka, 1985; Amann et al., 1995). This phenomenon lead to the known paradigm that "less than 1% of the bacterial community can be cultured". Different interpretations of this paradigm can be found through the literature (Martiny, 2019) but the closest one seems to refer to the original studies of Torsvik and Øvreås (2002) or to Amann et al. (1995), where comparisons between agar plate and direct-staining counts were presented. Recently, different studies had revised this paradigm and the "great plate count anomaly" phenomenon. Some claimed that given the coordinated effort to cultivate abundant microorganisms across different environments, including marine ecosystems, and the actual methodological improvements, the proportion of microorganisms similar to culturable relatives should be higher than 1%. In fact, Martiny (2019) compared Sanger 16S rRNA sequences to known cultured relatives classified in the Ribosomal Data Project (RDP) and concluded, for example, that in ocean communities 50% of the sequences, that corresponded to 35% of the taxa, had known relatives to cultured organisms at 97% sequence similarity. As similar results were found across biomes, Martiny (2019) claimed that a high proportion of bacteria is already culturable. However, Martiny (2019)'s statement has been argued in other studies where metagenomic and metatranscriptomic data (less skewed to cultured organisms compared to Sanger sequences that depend

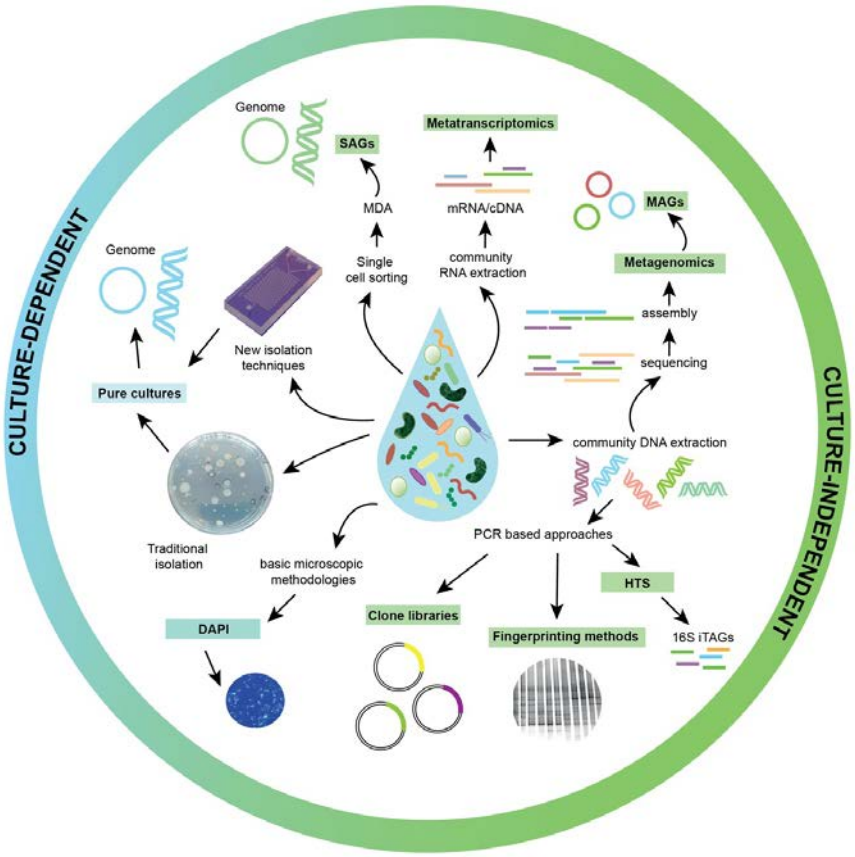

Figure 1: Range of culture-dependent and culture-independent techniques that can be used for microbial ecology studies. Blue side of the image and blue rectangles cover culture-dependent techniques, while the green part covers culture-independent techniques. Note that not all techniques that could be used are included in the diagram, but only those mentioned in the text. See Ferrera and Sánchez (2016) and Pedrós-Alió et al. (2018) for a more complete version. SAGs: Single amplified genomes; MDA: multiple displacement amplification; MAGs: metagenome-assembled genomes; HTS: high-throughput analysis; DAPI: 4',6-diamidino-2-phenylindole staining.

on PCR amplifications) had been analyzed and compared to known culturable organisms concluding that still higher proportions of bacteria remain uncultured across biomes (Lloyd et al., 2018; Steen et al., 2019). Nevertheless, these studies agree that the repeated axiom that only "1% of microbial cells can be cultured" should be retired or, at least, revised, because the meaning is often nebulous. In addition, even though the median of culturable cells from environmental samples is estimated around 0.5%, individual studies where culture-dependent and culture-independent techniques had been applied to diverse biomes show variability in the proportion of recovery rates (Staley and Konopka, 1985; Kaeberlein et al., 2002; Browne et al., 2016; Lloyd et al., 2018). On top of that, this type of studies combining cultures with HTS data are scarce (Selje et al., 2005; Zeng et al., 2012; Stefani et al., 2015) and they are mainly done at local or regional scales. In marine ecosystems, there is a lack of global culture depending studies focusing not only in the photic ocean but also in less explored deeper layers. Hence, part of this thesis is focused on the combination of culture-dependent and culture-independent techniques to elucidate the proportions of culturable bacteria at different ocean depths and the possible vertical connectivity between taxa along layers at a broader scale.

General patterns of microbial communities extracted from global circumnavigation expeditions

General microbial diversity patterns emerged from 16S rRNA amplicon tags (16S TAGs) (Zinger et al., 2011) and metagenomic 16S fragments (16S miTAGs) (Logares et al., 2014) thanks to global ocean surveys that stablished coherent sampling, sequencing and data analyses protocols. The pioneering survey was the Global Ocean Sampling expedition (GOS) launched in 2003 where HTS techniques were applied to describe the global ocean planktonic diversity (Venter et al., 2004). Afterwards, the International Census of Marine Microbes (ICoMM, 2005-2010) (Amaral-Zettler et al., 2010) was the first international effort to study marine microbial diversity through consistent standardized HTS, collecting samples from a wide variety of marine environments. Nowadays, the two main worldwide surveys conducted after the ICoMM initiative have been the *Tara* Oceans Expedition (2009-2013) (Karsenti et al., 2011) and the Malaspina 2010 Circumnavigation Expedition (2010-2011) (Duarte, 2015). *Tara* oceans sampled seven different size fractions of plankton diversity (from marine virus to zooplankton) in the global ocean including samples from surface waters to the mesopelagic layer (approximately 1000 m depth) of the temperate and polar latitudes. On the other hand, the Malaspina Expedition targeted the diversity of life in the ocean focusing in the deeper layers, mainly the bathypelagic (down to 4000 m depth).

Some of the microbial patterns that raised collectively from these ocean surveys include the key concept that microbial communities are formed by a few abundant taxa and a long tail of very low abundant species (rare biosphere) (Sogin et al., 2006). It has also been detected that temperature is the main driver for microbial community composition (Sunagawa et al., 2015). Moreover, biogeography studies describing how microbial communities distribute along space, vertically and horizontally, were also possible thanks to the above mentioned global samplings. It was found that there is a vertical segregation between photic and aphotic samples (Amaral-Zettler et al., 2010; Sunagawa et al., 2015) but at the same time there exists a vertical connectivity between surface and deep taxa through sinking particles (Mestre et al., 2018; Ruiz-González et al., 2020). Patterns of latitudinal gradient of diversity (LGD) was recently revealed for most marine planktonic organisms showing a decline towards the poles and maximum diversity at mid latitudes for prokaryotes (Ibarbalz et al., 2019). In addition, with the avail-

able data it seems that alpha diversity (richness in a sample) is higher in the deeper layers (Pommier et al., 2010; Sunagawa et al., 2015), and gamma diversity (total diversity in a landscape) is lower in deep waters compared to surface waters (Zinger et al., 2011; Sul et al., 2013; Sunagawa et al., 2015; Salazar et al., 2016). Then, marine bacterial communities are more similar among themselves in the bottom than at the surface of the oceans, despite finding higher richness in deeper layers (Pedrós-Alió et al., 2018). Horizontally, it was observed that there is a limited dispersion of bacterial communities across space (Galand et al., 2010; Pommier et al., 2010; Zinger et al., 2014; Salazar et al., 2015; Sunagawa et al., 2015). Nevertheless, these microbial diversity patterns are mainly based on very short fragments of the 16S rRNA gene and a limited sequencing effort to reach to the total microbial diversity (Duarte et al., 2020). Our current HTS techniques and sampling strategies are still far from retrieving the full microbial diversity of the oceans (Curtis et al., 2002; Locey and Lennon, 2016).

The importance of cultures in the "omics age"

Despite the fact that molecular techniques and more recently HTS techniques have allowed the description of marine microbial diversity from diverse habitats at an unprecedented scale, as we have already overviewed, isolates are still a valuable resource of knowledge. Usually there is a few overlap between taxa retrieved by molecular techniques and those retrieved by isolation (Lekunberri et al., 2014; Crespo et al., 2016). This is mainly caused because molecular techniques usually recover the abundant bacteria present in a given environment, while cultures retrieve those taxa that belong to the rare biosphere and are present in very low abundances (Sogin et al., 2006; Pedrós-Alió, 2012). Thus, isolation is still a reliable technique to decipher the full spectra of diversity of an ecosystem. Furthermore, isolates help us to test ecological hypotheses raised from metagenomic and metatranscriptomic data, to interpret multi-species interactions, evolutionary principles or population dynamics, and at the same time, improve the available databases needed for the correct annotations of HTS data (Giovannoni and Stingl, 2007; Gutleben et al., 2017) (Figure 2). Moreover, isolation is still key for the official procedures for classification and characterization of novel bacterial species (Parker et al., 2019), which, at the same time, allow us to have access to their complete genomes, and the functional characterization of novel genes (Muller et al., 2013) (Figure 2). Finally, the short generation times and the nearly 4 billion years of evolution of marine microorganisms has resulted in an enormous biodiversity and a plethora of metabolic pathways which make isolates an excellent material for biotechnology research (Luna, 2015) including bioremediation of polluted ecosystems.

Marine bacteria are thought to be good candidates for bioremediation studies as they are naturally exposed to unfavorable conditions and live in a very dynamic environment with fluctuations in light, temperature, salinity, pH, pressure or nutrients. Moreover, most of the marine bacteria that can be retrieved currently by culture-dependent techniques are copiotrophs (Overmann et al., 2017), which usually contain larger genomes (>4 Mb) and diverse metabolisms, so they respond to changes in environmental conditions by controlling the transcription of many genes (Cottrell and Kirchman, 2016) and can be potential detoxifiers in bioremediation. Nowadays, the literature shows different examples of marine bacteria used in bioremediation studies confirming their clear potential (Deppe et al., 2005; Sekiguchi et al., 2010; Panwichian et al., 2011). Nevertheless, the potential metabolisms of bacteria inhabiting less explored habitats, such as the mesopelagic or bathypelagic layers of the oceans, is far for being totally uncovered and the isolation and physiological characterization of such bacteria could be

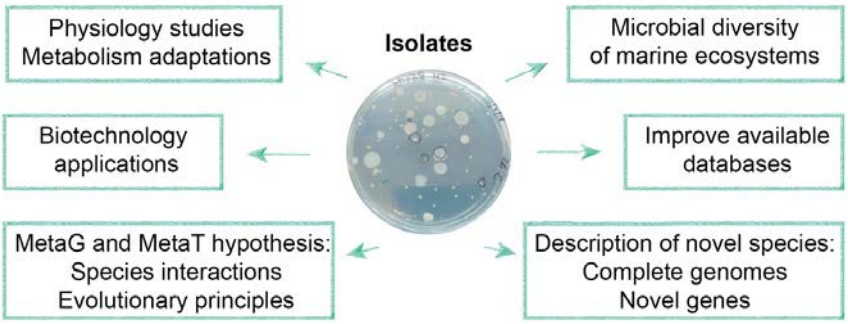

Figure 2: Outline of the importance and usefulness of bacterial isolates in microbiology studies. MetaG: metagenomics; MetaT: metatrasnscriptomics.

very promising for future clean-up strategies of marine contaminated areas. Consequently, part of this thesis focused on the recovery of marine bacteria from different ocean layers, will also try to address this challenge.

Marine pollution and bioremediation by marine bacteria

Oceans not only harbor an incredible diversity of life but they also provide food, energy and water, and are helpful in sustaining the livelihoods of hundreds of millions of people. Besides, as we mentioned earlier, they produce half of the oxygen that we breath and are the main stabilizers of world's climate. Marine pollution, defined as "*the introduction by man of substances into the marine environment resulting in such deleterious effects as harm to living resources, hazards to human health, hindrance to marine activities including fishing, impairment of quality for use of seawater and reduction of amenities*" (GESAMP, 1969), have been impacting and reshaping the chemistry of the global ocean. This has occurred at different scales since late 1700s or early 1800s, that is, at the beginning of the Anthropocene (Crutzen, 2002; Steffen et al., 2007) when societies became a geological and ecological driving force. However, it has been in the last century when oceans have suffered an increment of contamination due to the increase of human populations and the uncontrolled development of multiple human activities in, for example, industry, transport, agriculture, and urbanization. The activities derived from industrial discharges and the runoff from agricultural activities and coastal cities are among those of major concern (Wilhelmsson et al., 2013). Some of the main and common contaminants derived from human activities include persistent organic pollutants (POPs) (Wenning and Martello, 2014; Lohmann and Dachs, 2019), pesticides (Köck-Schulmeyer et al., 2019; Tsygankov, 2019), plastics (Auta et al., 2017; Sharma and Chatterjee, 2017), personal care products and pharmaceuticals (López-Pacheco et al., 2019), hydrocarbons (Mojiri et al., 2019), radionuclides (Uddin et al., 2020), or toxic heavy metals (Yılmaz et al., 2017; UN Environment, 2019), which are more or less ubiquitous in the global ocean. Moreover, the effects of climate change in the oceans worsen the accumulation effects of certain substances, affecting their biogeochemical cycles and their presence in marine food webs (Wilhelmsson et al., 2013; Alava et al., 2017, 2018).

The conservation and sustainable management of the biodiversity and the multiple marine environments are the foremost priorities in the global environmental agenda of the United Nations Sustainable

Development Goals (SDGs) as part of the UN 2030—Agenda for Sustainable Development[1]. In addition, specific programs as the Minamata Convention has been created to reduce, in this case, levels of mercury releases into the environment (Eriksen and Perrez, 2014). In this framework, it is necessary to pursue new research oriented in finding solutions for the correct management and protection of marine ecosystems where ecotoxicologists, marine biologists or environmental scientists are involved. Blue biotechnology approaches like bioremediation using marine bacteria extracted from the natural ecosystems could be a powerful solution, as this approximation is an attempt to accelerate naturally occurring degradation by using the metabolic potential of bacteria and optimizing the limiting conditions for growth (Dash et al., 2013).

Mercury pollution in the ocean

Mercury is one of the most toxic heavy metals with little or no biochemical function and, while tolerable in minute quantities, exhibit toxic effects above critical concentrations (Clarkson, 1993; Clarkson and Magos, 2006). Mercury is emitted to the atmosphere and released to waters and lands as a result of anthropogenic activities (discharge of sewage waters from the chlor-alkali plants, incineration of residues, the combustion of carbon or mining) and natural sources (geological activity, biomass burning or rock weathering) (Li et al., 2009; Gworek et al., 2016). The mercury volatile form (elemental mercury, Hg^0) can be transported in the atmosphere around the world and eventually be deposited in soils, plants or water, including marine ecosystems (Mason and Sheu, 2002; Soerensen et al., 2010; Kim and Zoh, 2012; Enrico et al., 2016; Saiz-Lopez et al., 2018). Besides the mercury elemental form, other chemical species can be found in the oceans including divalent mercury (Hg^{2+}) or the highly toxic organic form, methylmercury (MeHg) (Figure 3) (Mason et al., 2012). Methylmercury is produced by both aerobic and anaerobic bacteria (Monperrus et al., 2007; Malcolm et al., 2010; Podar et al., 2015; Gionfriddo et al., 2016; Munson et al., 2018), and is bioaccumulated in the aquatic food webs (Wiener et al., 2002; Mason et al., 2012; Harding et al., 2018; UN Environment, 2019) reaching humans through fish consumption. This last form has received the most attention largely due to notorious methylmercury poisoning events in Japan and Iraq following high exposures (Bakir et al., 1973; Harada, 1995). Methylmercury poisoning is associated with adverse effects on brain development, especially in fetuses, but it can also affect the cardiovascular and immune systems after low-level but chronical exposures (UN Environment, 2019).

Mercury levels are different in each particular ocean presenting an average concentration of 1.5 pM (Lamborg et al., 2002). The Mediterranean sea and the North Atlantic Ocean have recorded the higher concentrations, around 2.5 and 2.0 pM, respectively (Cossa et al., 1997; Mason et al., 1998). The Arctic and Antarctic Oceans present the lower concentrations but are neither free of mercury, with levels of 0.8 pM detected (Laurier et al., 2004; Sunderland and Mason, 2007). These measures, when compared to a pre-industrial era, show an increase of mercury concentrations of 230%, 25% and 12% in surface, intermediate and deep waters, respectively (Sunderland and Mason, 2007; Lamborg et al., 2014a; UN Environment, 2019). They usually refer to total mercury levels including the different chemical species that can be found in the water column. The average levels of MeHg in oceanic waters fell in the range of 2–35% of this total mercury (Sunderland et al., 2009) and it is estimated that its concentration has also increased since pre-industrial era by a 200-500% (Lamborg et al., 2014a). Also, mercury and its various forms are variably distributed along the water column. It is assumed that the point of entry of mercury

[1] https://sustainabledevelopment.un.org/post2015/transformingourworld

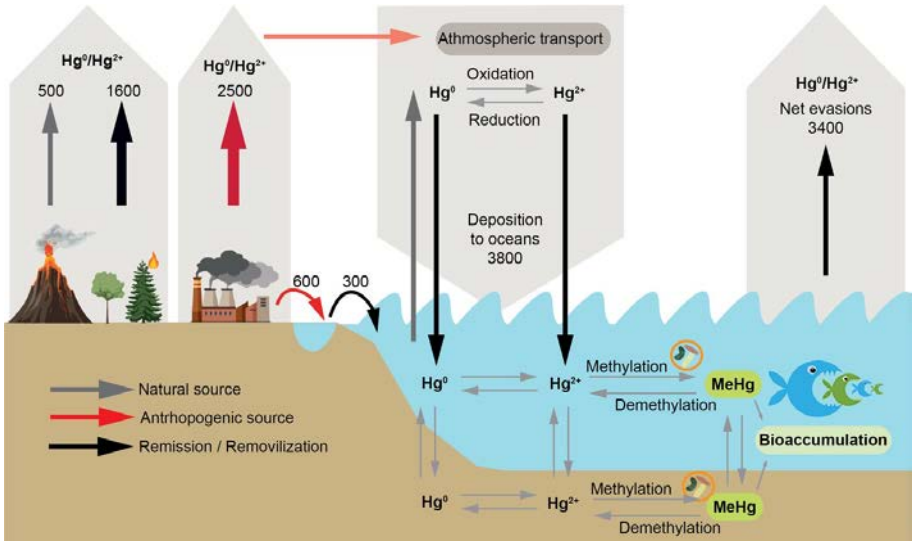

Figure 3: Mercury cycle and global mercury budgets released into the atmosphere and cycled through the oceans and atmosphere. Mercury transformations indicated are focused in those mainly occurring in marine ecosystems including the water column (blue area) and the sediments (brown area). Values above arrows are registered in tons/year and were extracted from the UNEP Global mercury assessment (UN Environment, 2019). Hg^0: elemental mercury; Hg^{2+}: divalent mercury ion; MeHg: methylmercury.

into marine ecosystems is the surface mixed layer (up to 100 m depth) (Gworek et al., 2016). From this layer the reactive mercury is transported mainly with suspended particles to regions of the ocean where methylation occurs, and MeHg reaches its maximum normally below the thermocline where low oxygen levels are present (up to 1000 m depth) (Cossa et al., 1997; Sunderland et al., 2009; Mason et al., 2012). This methylated form can be transported back to surface layers where it can be demethylated and converted to Hg^0 which is eventually released into the atmosphere (Mason and Sheu, 2002; Gworek et al., 2016).

The mercury resistance operon

Due to the elevated distribution of mercury species in the ocean water column, it is not surprising to find not only the bacteria that methylate mercury, but also those with resistance to its different forms. These resistant microorganisms harbour the *mer* operon and about 1–10% of cultured heterotrophic aerobic microbes from various environments possess this set of genes (Barkay, 1987). The operon encodes different genes including: the homodimeric flavin-dependent disulfide oxidoreductase enzyme mercuric reductase (MerA), an organomercury lyase (MerB) that is not always present, a periplasmic Hg^{2+} scavenging protein (MerP), one or more inner membrane spanning proteins (MerT, MerC, MerE, MerF, MerG) that transport Hg^{2+} to the cytoplasm where it is reduced by MerA, and one or two regulatory proteins (MerR, MerD) (Boyd and Barkay, 2012) (Figure 4).

This operon seems to have evolved from a geothermal environment where exposures to geological sources of mercury would have driven the origin of MerA (Boyd and Barkay, 2012). After that, other genes have gradually increased the size, complexion and functionality of the operon and it has become a tightly regulated and efficient mercury detoxification system in diverse environments, including marine ecosystems (Nakamura et al., 2000; De and Ramaiah, 2007; Møller et al., 2014; Oyetibo et al., 2015; Ramadan et al., 2019). However, few studies targeting the physiological characteristics of mer-

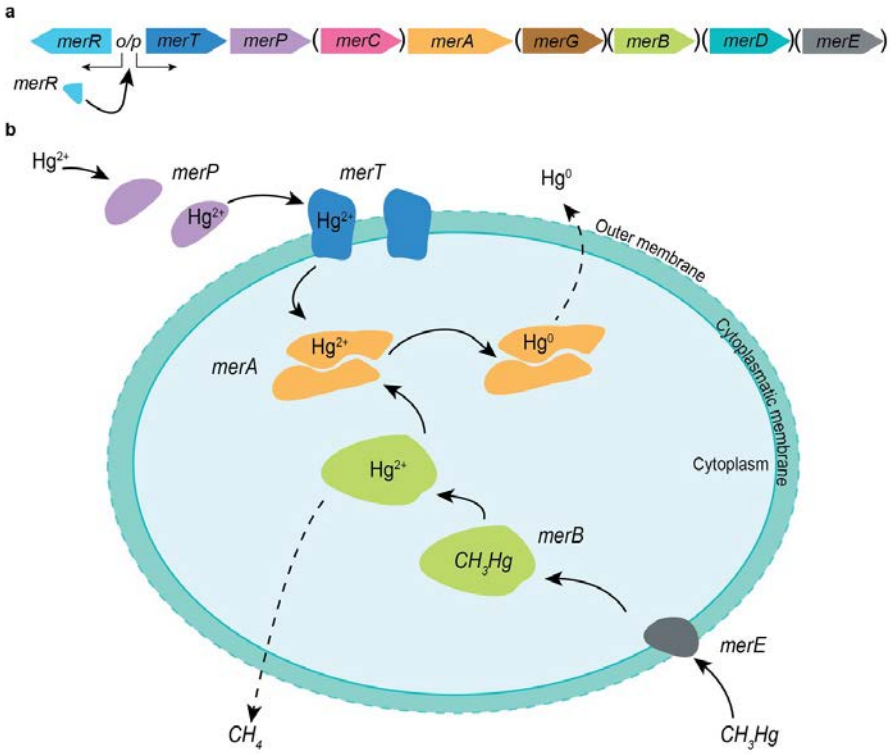

Figure 4: The *mer* system. **(A)** Representation of the genes included in the *mer* operon. Genes between brackets are not always present in the operon. **(B)** The cellular *mer*-encoded mercury detoxification mechanisms. The outer cell wall is depicted by a broken line illustrating that not all microbes have an outer membrane. Broken line arrows depict diffusion while solid line arrows indicate transport or mercury transformations. Illustration modified from Boyd and Barkay (2012).

cury resistance marine bacteria distributed across different oceans and layers can be found through the literature. Given the beforehand need for finding biological alternatives for the removal of mercury from contaminated sites, the detection and characterization of marine mercury resistant bacteria, possibly adapted to different environmental conditions and with diverse metabolisms, seems relevant for future bioremediation processes.

Aims, outline and objectives of the Thesis

AIMS, OUTLINE AND OBJECTIVES OF THE THESIS

The general purpose of this thesis is to **explore the distribution of marine heterotrophic culturable bacteria retrieved from different ocean layers and oceanographic regions, decipher its contribution in the bacterial diversity thanks to global marine microbial diversity surveys, and gain insights of their genetic capabilities concerning mercury detoxification** (Figure 5).

The achievement of this major goal is structured in **three chapters**. In the **first chapter**, a marine culture collection of heterotrophic bacteria (MARINHET) was created with isolates retrieved through traditional culture-dependent techniques from a wide variety of oceanographic regions and depths, and its taxonomic diversity and distribution along different layers was studied. This culture collection led to the work presented in the second and third chapters. In the **second chapter**, "the great plate count anomaly" phenomenon was revised in the mesopelagic and bathypelagic waters by comparing flow cytometry counts with plate colony counts. Additionally, the known axiom that "less than 1% of microbial cells can be cultured" was tested by comparing the 16S rRNA sequences of the isolates from the culture collection to the HTS 16S rRNA sequences (Illumina 16S amplicons TAGs (16S iTAGs) and metagenomic 16S iTAGs) obtained from the same samples and others worldwide distributed from *Tara* Oceans and Malaspina expeditions. Finally, in the **third chapter** the genetic potential of a fraction of the MARINHET collection was assessed in order to find biological alternatives for future bioremediation studies in mercury impacted sites. Specifically, the ability of *Alteromonas* and *Marinobacter* isolates for the degradation of mercury toxic species was assessed.

The outline of the different chapters can be further detailed under three general objectives and several specific tasks as follows:

Objective 1: to create a marine culture collection of heterotrophic marine bacteria from a variety of oceanographic regions and depths using standard culture-dependent techniques to extent the knowledge of the diversity of marine microbial communities.

This extensive culture-dependent study, which kind is scarce in the literature, targeted a fraction of the culturable bacterial community at a large scale, including samples from different oceans and layers, as well as mesopelagic and bathypelagic samples. In **Chapter 1**, a great isolation effort has been done to:

- Identify the taxonomy of the heterotrophic isolates retrieved and explore their phylogenetic diversity and the potential differences between depths.
- Reveal the most common distributed heterotrophic culturable bacteria across oceans and depths.
- Describe the biogeography of the most abundant recovered isolates.
- Unveil some novel isolated bacterial strains.

Objective 2: to explore the contribution of heterotrophic marine cultured bacteria in global deep ocean microbial diversity surveys.

Studies comparing the bacterial diversity using both traditional culture-dependent techniques and culture-independent techniques in the same samples at wide scale are missing. In **Chapter 2**, the

diversity of photic and aphotic zones that could be retrieved by culture-dependent techniques was examined and compared with the HTS data of different bacterial size fractions of the *Tara* Oceans and Malaspina Expeditions in order to:

- Enumerate which proportion of the photic and aphotic bacteria from the HTS data could be isolated and classify those among the abundant or the rare biosphere.
- Test if the "great plate count anomaly" applies in the deep ocean by comparing colony counts to flow cytometry counts from the same samples.
- Analyze if the isolated bacteria are more prone to be associated to free-living or particle-attached bacterial communities at the different studied depths in order to reveal a vertical connectivity pattern.

Objective 3: to characterize the genetic and physiological potential of *Alteromonas* and *Marinobacter* isolates to detoxify mercury toxic compounds.

There are few culture-dependent studies aiming to detect marine mercury resistant bacteria from different depths and at a large scale. In **Chapter 3**, 247 *Alteromonas* and 46 *Marinobacter* isolates were examined to:

- Detect marine bacteria from the MARINHET collection codifying *mer* genes, including mercuric reductase (*merA*) and organomercury lyase (*merB*).
- Assess the tolerance of the strains harboring those genes facing different concentrations of inorganic and organic mercury.
- Describe the potential of the most resistant strains in the removal of mercuric forms, as well as their biogeographic distribution in the marine environment thanks to HTS data.

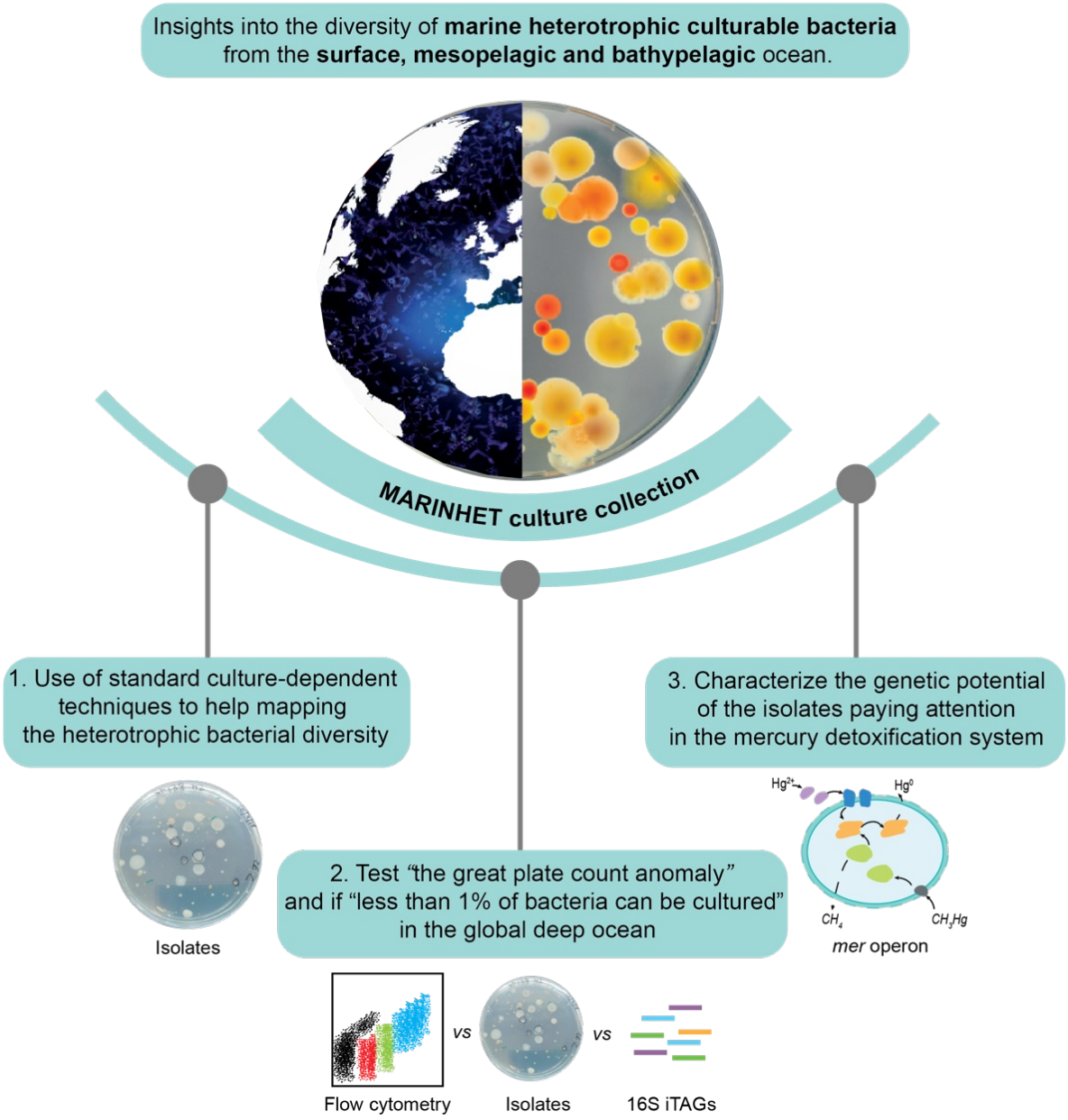

Figure 5: Graphic representation of the overall purpose of this thesis and its three main objectives. World map modified from "Investigación y Ciencia".

Chapter 1

Diversity and distribution of marine heterotrophic bacteria from a large culture collection

Abstract

Isolation of marine microorganisms is fundamental to gather information about their physiology, ecology and genomic content. To date, most of the bacterial isolation efforts have focused on the photic ocean leaving the deep ocean less explored. We have created a marine culture collection of heterotrophic bacteria (MARINHET) using a standard marine medium comprising a total of 1561 bacterial strains, and covering a variety of oceanographic regions from different seasons and years, from 2009 to 2015. Specifically, our marine collection contains isolates from both photic (817) and aphotic layers (744), including the mesopelagic (362) and the bathypelagic (382), from the North Western Mediterranean Sea, the North and South Atlantic Ocean, the Indian, the Pacific, and the Arctic Oceans. We described the taxonomy, the phylogenetic diversity and the biogeography of a fraction of the marine culturable microorganisms to enhance our knowledge about which heterotrophic marine isolates are recurrently retrieved across oceans and along different depths.

The partial sequencing of the 16S rRNA gene of all isolates revealed that they mainly affiliate with the classes *Alphaproteobacteria* (35.9%), *Gammaproteobacteria* (38.6%), and phylum *Bacteroidetes* (16.5%). In addition, *Alteromonas* and *Erythrobacter* genera were found the most common heterotrophic bacteria in the ocean growing in solid agar medium. When comparing all photic, mesopelagic and bathypelagic isolates sequences retrieved from different stations, 37% of them were 100% identical. This percentage increased up to 59% when mesopelagic and bathypelagic strains were grouped as the aphotic dataset and compared to the photic dataset of isolates, indicating the ubiquity of some bacterial isolates along different ocean depths. Finally, we isolated three strains that represent a new species, and the genome comparison and phenotypic characterization of two of these strains (ISS653 and ISS1889) concluded that they belong to a new species within the genus *Mesonia*.

Overall, this study highlights the relevance of culture-dependent studies, with focus on marine isolated bacteria from different oceanographic regions and depths, to provide a more comprehensive view of the culturable marine bacteria as part of the total marine microbial diversity.

Introduction

Traditional culturing methods allow the isolation of microorganisms from natural samples with the possibility to sequence their genome, perform physiological/experimental assays and, thus, infer their functional and ecological role in detail. Moreover, microbial cultures can retrieve diversity usually not recovered by molecular methodologies, particularly bacteria belonging to the "rare biosphere", i.e. bacterial species that are present in very low abundances in the environment (Pedrós-Alió, 2012; Shade et al., 2012). The overlap between isolated microorganisms and those belonging to the uncultured majority is relatively low in molecular surveys, and efforts to culture bacteria from the ocean often yield isolates that do not have their corresponding 16S rRNA gene sequences deposited in sequence databases (Suzuki et al., 1997; Lekunberri et al., 2014). As a consequence, isolation of microorganisms by culture-dependent techniques, and their comparison to data obtained from high-throughput sequencing techniques (HTS), remains a fundamental tool to fully understand the whole range of bacterioplankton diversity found in the ocean. In addition, isolation is so far a requisite for the description of new microbial species.

Most of the studies targeting the marine heterotrophic culturable bacteria have focused on the upper ocean (0-200 m depth) or on specific oceanographic regions (Pinhassi et al., 1997; Eilers et al., 2000; Lekunberri et al., 2014; Yang et al., 2016) while studies covering different depths are less frequent (Castro Da Silva et al., 2013; Kai et al., 2017; Liu et al., 2018) . Efforts to culture bacteria from the deep ocean (>200 m) have focused mostly on isolates from hydrothermal vents (Nakagawa and Takai, 2008; Ferrera et al., 2014; Grosche et al., 2015), whale carcasses (Tringe et al., 2005), trenches (Eloe et al., 2011b), and deep-sea sediments (Zobell and Morita, 1957; Sahm et al., 1999; Castro Da Silva et al., 2013; Hwang et al., 2015; Chen et al., 2016). Thus, very few studies have analyzed the diversity of isolates from mesopelagic (in particular from regions with oxygen minimum zone areas) (Finster and Kjeldsen, 2010; Liu et al., 2018; Menezes et al., 2018; Mulla et al., 2018), the bathypelagic and abyssopelagic waters (Tabor et al., 1981; Kaye and Baross, 2000; Yuan et al., 2015; Kai et al., 2017), and those available were mainly done at a local or regional scale. Therefore, a study of the culturable microorganisms covering different layers including underexplored areas such as the mesopelagic and the bathypelagic areas is missing.

Here we present an extensive marine heterotrophic bacterial culture collection (MARINHET) with 1561 marine bacteria retrieved from different ocean depths from the Mediterranean Sea, the North and South Atlantic Oceans, the Indian, the Pacific, and the Arctic Oceans, covering diverse latitudes, from different seasons and years from 2009 to 2015. We used well established marine solid media (Zobell agar and Marine Agar 2216) in order to describe the fraction of the bacterioplankton community than can be commonly isolated under laboratory conditions (nutrient rich medium, standard oxygen concentrations and atmospheric pressure). Therefore, to the best of our knowledge, we have created the first extensive marine heterotrophic bacterial culture collection, including isolates from different depths and oceanographic regions, that were retrieved through a standard methodology. Analyses of the partial 16S rRNA gene sequences (average 526 bp, covering V3 to V5 regions), allowed us: (i) to identify the taxonomy of those isolates distributed along the water column, (ii) to explore the phylogenetic diversity and the potential differences between depths, (iii) to reveal the most common distributed heterotrophic culturable bacteria across oceans and depths, (iv) to describe the biogeography of the most abundant isolates recovered, (v) to compare the isolates 16S rRNA sequences with available HTS 16S rRNA se-

quences derived from samples of the same oceanographic expeditions, and (vi) to unveil some novel isolated bacterial strains.

Materials and methods

Study areas and sampling

A total of eight photic-layer, four mesopelagic, and seven bathypelagic samples were taken during different oceanographic cruises in several sampling stations distributed along a wide range of latitudes (Figure 6). Photic-layer samples (Table 1) were collected in the Atlantic and Indian Oceans during the *Tara* Oceans expedition in 2009-2013 (Karsenti et al., 2011), and from the Arctic Ocean during the ATP cruise in 2009 (Lara et al., 2013). Additionally, surface seawater samples from the Blanes Bay Microbial Observatory, (BBMO)[2] in the NW Mediterranean Sea were collected in May 2015. Mesopelagic samples (Table 1) were taken from the Indian and the Pacific Oceans also during the *Tara* Oceans expedition in 2009-2013 (Karsenti et al., 2011). All the mesopelagic samples were collected in regions with oxygen minimum zone (OMZ) areas, ST39 from the Arabian Sea, ST102 and ST111 from the Eastern Tropical South Pacific, and ST138 from the Eastern Tropical North Pacific. Bathypelagic samples (Table 1) from the Atlantic Ocean at approximately 4000 m depth were taken from six different stations during the Malaspina 2010 Circumnavigation Expedition (Duarte, 2015). One of the stations sampled was located in the North Atlantic, whereas the other five stations were located in the South Atlantic. ST43 could specially differ from the rest South Atlantic samples because it was particularly placed in the Agulhas Ring, where deep waters from the South Atlantic converge and mix with Indian Ocean deep-water masses (Villar et al., 2015). In addition, one bathypelagic sample was collected at 2000 m depth in the NW Mediterranean during the MIFASOL cruise in September 2014.

In each of these stations, seawater was collected using Niskin bottles attached to a rosette sampling system, except at BBMO, where samples were collected with a bucket. Seawater was sequentially filtered through 200 µm and 20 µm meshes to remove large plankton cells and to keep the free-living bacterial community together with the one attached to particles (< 20 µm). Duplicate 2 ml seawater of each station were kept in Eppendorf tubes with dimethyl sulfoxide (DMSO) 7% final concentration and stored at -80ºC until further processing in the laboratory.

Geographical coordinates of stations, sampling date, sampled depth, *in situ* temperature and total number of sequenced isolates are listed in Table 1.

Culturing and isolation

Isolates were obtained by plating 100 µl of undiluted and 10x diluted seawater from the photic, mesopelagic and bathypelagic samples, in triplicates, onto Zobell agar plates (i.e. 5 g peptone, 1 g yeast extract and 15 g agar in 750 ml of 30 kDa filtered seawater and 250 ml of Milli-Q water) or Marine Agar 2216 (Difco™) plates, which is based also on the Zobell medium formulation (Zobell, 1941). Our medium culturing strategy was only focused to retrieve heterotrophic marine bacteria that could

[2] http://www.icm.csic.es/bio/projects/icmicrobis/bbmo

Table 1: Characteristics of the different samples used for isolation of marine heterotrophic bacteria. Non-redundant isolates stand for the number of different isolates remaining after removing those that were 100% identical in their partial 16S rRNA gene.

Station	Sampling date	Oceanic location	Latitude	Longitude	Depth (m)	*In situ* temperature (ºC)	N° of sequenced isolates	N° of non-redundant isolates
Tara								
ST 39	March 2010	Indian Ocean	18º 35.2' N	66º 28.22' E	5.5	26.2	104	25
ST 39	March 2010	Indian Ocean	18º 35.2' N	66º 28.22' E	25	26.8	243	53
ST 39	March 2010	Indian Ocean	18º 43.12' N	66º 21.3' E	268.2	15.6	88	18
ST 67	September 2010	South Atlantic	32º 17.31' S	17º 12.22' E	5	12.8	115	49
ST 72	October 2010	South Atlantic	8º 46.44' S	17º 54.36' W	5	25	71	33
ST 76	October 2010	South Atlantic	20º 56.7' S	35º 10.49' W	5	23.3	89	27
ST 151	March 2012	North Atlantic	36º 10.17' N	29º 1.23' W	5	17.3	76	33
ST 102	April 2011	Pacific Ocean	5º 16.12' S	85º 13.12' O	475.6	9.2	97	15
ST 111	June 2011	Pacific Ocean	16º 57.36' S	100º 39.36' O	347.1	10.9	98	35
ST 138	December 2011	Pacific Ocean	6º 22.12' N	103º 4.12' O	444.9	8.2	79	34
ATP								
AR_1	June 2009	Arctic Ocean	78º 20.00' N	15º 00.00' E	2	6.2	13	9
AR_2	June 2009	Arctic Ocean	76º 28.65' N	28º 00.62' E	25	-1.2	20	9
Malaspina								
ST 10	December 2010	North Atlantic	21º 33.36' N	23º26' W	4002	2	20	9
ST 17	February 2011	South Atlantic	3º 1.48' S	27º 19.48' W	4002	1.7	93	24
ST 23	August 2011	South Atlantic	15º 49.48' S	33º 24.36' W	4003	1.5	94	39
ST 32	January 2011	South Atlantic	26º 56.8' S	21º 24' W	3200	2.5	39	16
ST 33	January 2011	South Atlantic	27º 33.2' S	18º 5.4' W	3904	1.7	5	5
ST 43	April 2011	South Atlantic	32º 48.8' S	12º 46.2' E	4000	1.2	4	4
MIFASOL								
ST 8	September 2014	NW Mediterranean	40º 38.41' N	2º 50' E	2000	13.2	127	36
BBMO								
IBSURF	May 2015	NW Mediterranean	41º 40' N	2º 48' E	5	17.7	86	43

grow easily under laboratory conditions (nutrient rich medium, standard oxygen concentrations and atmospheric pressure) using two similar culturing media. The only difference between Zobell agar and Marine Agar 2216 plates is the use of natural seawater (Zobell agar), or the addition of the minerals and salts contained in natural seawater to distilled water (Marine Agar 2216). Indeed, we did not observe significant differences (Fisher test analyses, data not shown) in the bacterial isolation between both media.

Photic-layer and mesopelagic samples were incubated at room temperature (RT, 20ºC approx.) while bathypelagic samples were incubated at their *in situ* temperature, which ranged from approximately 4ºC (in the Atlantic Ocean at 4000 m depth) to 12ºC (NW Mediterranean at 2000 m depth) (Table 1 and Table S1), but also at RT in order to assure bacterial recovery from all stations. In all cases, triplicates of each temperature condition and dilution were incubated in the dark until no more colonies appeared (10-30 days).

A total of 1561 bacterial isolates were randomly selected for DNA amplification and partial sequencing of their 16S rRNA gene (Table 1 and details below). Similar number of isolates were sequenced from photic layers (817; average: 102 isolates per station) and from deep oceans (744; average: 67 isolates per station) with 362 isolates from the mesopelagic and 382 from the bathypelagic. In most of the bathypelagic samples we collected all colonies appearing in the plates, which ranged from 6 to 129 including all replicates. Colonies were streaked on agar plates in duplicate to ensure their purity and avoid contamination. The isolates were stored in the broth medium used with glycerol 25% in cryovials at -80ºC.

PCR amplification and sequencing of the 16S rRNA gene

Available DNA used for template in Polymerase Chain Reaction (PCR) was extracted from 200 µL of isolates liquid cultures placed in 96 well plates, diluted 1:4 and heated (95ºC, 15 min) to cause cell lysis. The partial 16S rRNA gene sequences were PCR amplified using bacterial primers 358F (5'-CCT ACG GGA GGC AGC AG-3') (Muyzer et al., 1993) and 907Rmod (5'-CCG TCA ATT CMT TTG AGT TT-3') (Sánchez et al., 2007). The complete 16S rRNA gene was amplified for *Mesonia* strain ISS653 after DNA extraction using the DNeasy Blood |& Tissue kit (Qiagen), following the manufacturer's recommendations, and using the modified primers from Page et al. (2004) 27F (5'- AGR GTT TGA TCM TGG CTC AG -3') and 1492R (5'- TAC GGY TAC CTT GTT AYG ACT T -3'). Detailed PCR conditions are described in Supplementary Information Chapter 1. Purification and OneShot Sanger sequencing of 16S rRNA gene products was performed by Genoscreen (Lille, France) with primer 358F for partial sequences, and with both 27F and 1492R for complete sequences. ChromasPro 2.1.8 software (Technelysium) was used for manual cleaning and quality control of the sequences.

Data processing and taxonomic classification

The 16S rRNA sequences of our cultured strains were clustered at 99% sequence similarity (Acinas et al., 2003) in order to define different operational taxonomic units (iOTUs or isolated OTUs) and construct iOTU-abundance tables for the different stations and layers studied (Table S2) using UCLUST algorithm from the USEARCH software (Edgar, 2010). The different iOTUs were taxonomically classified using the lowest common ancestor (LCA) method in SINA classifier (Pruesse et al., 2012), using both SILVA (release 132 in 2017) and RDP (Ribosomal Database Project, release 11) databases. Parallelly, isolates sequences were submitted to BLASTn (Altschul et al., 1990) with two subsets of the RDP database, one including only the uncultured bacteria (Closest Environmental Match, CEM), and another including only the cultured bacteria (Closest Cultured Match, CCM) in order to extract the percentages of similarity with both datasets (Tables S3 and S4), and to assess whether our isolates were similar to effectively published cultured organisms.

Additionally, a more restrictive clustering at 100% sequence similarity (USEARCH software) was also used to define iOTUs and to detect how many bathypelagic, mesopelagic and photic-layer bacterial isolates were identical, and thus, to identify bacterial taxa or strains that could distribute along different water depths. Such comparisons were done with: (i) photic and mesopelagic isolates sequences retrieved from the ST39 vertical profile and (ii) the whole isolates dataset.

Phylogenetic analyses

The phylogeny was inferred for the representative isolates of each iOTU defined at 99% and 100% sequence similarity. The closest sequence to each isolated iOTU in SILVA v.132 database was found and collected using BLASTn (Altschul et al., 1990). Alignment of the isolates and reference sequences was performed with MUSCLE from the Geneious software v.11.0.5 (Kearse et al., 2012). The alignment was trimmed to the common 16S rRNA gene fragment covered by both sets of sequences. Phylogeny was constructed using maximum-likelihood inference with RAXML-NG 0.9.0 (Kozlov et al., 2019) and the GTR

evolutionary model with optimization in the among-site rate heterogeneity model and the proportion of invariant sites (GTR+G+I), and 100 bootstrap replicates.

Eventually, some isolates among our culture collection presented partial 16S rRNA sequences with a percentage of similarity below the 97% with public databases. In this case, the complete 16S rRNA gene was sequenced for ISS653, with which two more strains (ISS1889 and ISS2026) clustered at 100% similarity, and a phylogenetic tree was constructed to support its putative novelty. The tree included their complete and partial sequences of the 16S rRNA gene, their best hits from uncultured and cultured microorganisms, extracted from local alignments against RDP 11, SILVA LTP (Living Tree project), and National Center for Biotechnology Information (NCBI) databases, and the reference 16S rRNA genes from their related genera. Details on the phylogenetic tree constructions are explained in Supplementary Information Chapter 1.

Comparisons between layers and statistical analyses

All data treatment and statistical analyses were conducted in the R statistical software version 3.4.3 (R core team, 2017) and packages *stats*, *vegan* version 2.5-3 (Oksanen et al., 2018), *ape* version 5.1 (Paradis et al., 2004), *picante* version 1.6-2 (Kembel et al., 2010) and *EcolUtils* (Salazar, 2018a). In general, analyses were performed using the non-rarefied iOTU-abundance tables, but for specific analyses, such as the detection of iOTUs present along different depths, the iOTU-abundance table constructed with the sequences clustering at 99% was sampled down to the lowest sampling effort (362 isolates in the mesopelagic). In this manner, the rarefied or subsampled iOTU table was obtained using the function *rrarefy.perm* with 1000 permutations from the R package *EcolUtils* (Salazar, 2018a).

Rarefaction curves were performed with the package *vegan* to estimate the sampling effort in each studied layer. We also calculated bacterial richness/diversity metrics from each depth using two approaches: an OTU-based approach (i.e. considering the iOTUs as unrelated biological entities), and a phylogenetic approach (i.e. considering the evolutionary relationships among iOTUs with the complete computed phylogeny). The number of iOTUs, the Chao extrapolative richness estimator (Colwell and Coddington, 1994) and the Shannon entropy index (Shannon, 1948) were computed as OTU-based metric using the non-rarefied iOTU abundance table, while the Faith's phylogenetic diversity (PD) (Faith, 1992), the PD divided by the number of iOTUs (PD/iOTUs), and the mean nearest taxon distance (MNTD) (Webb et al., 2002) were used as phylogenetic measures for diversity. Differences between photic, mesopelagic and bathypelagic for richness/diversity measures were tested using an ANOVA test followed by the Tukey's post hoc test, as data normality was assured. To assess significance, the statistical analyses were set to a conservative alpha value of 0.01.

The Good's coverage (C) for each of the depths was also calculated by the equation $C=[1-(n_1/N)]*100\%$, where N is the number of iOTUs being examined and n_1 represents the number of iOTUs occurring only once or singletons (Giovannoni et al., 1995).

Comparison to environmental 16S rRNA Illumina sequences

Isolates were compared to denoised zOTUs (zero-radius OTUs, i.e. OTUs defined at 100% sequence similarity) (Edgar, 2010) from high-throughput sequencing (HTS) of the 16S rRNA sequences (16S iTAGs) obtained from *Tara* Oceans and Malaspina Expedition datasets which covered surface, mesopelagic and bathypelagic layers. Further description of those datasets, sample collection, DNA extraction, sequence processing and data treatment are described in Supplementary Information Chapter 1. All isolates sequences were compared to zOTUs sequences at 100% similarity respectively, by running global alignment using the *-usearch_global* option from the USEARCH v10.0.240 (Edgar, 2010). The results were filtered by coverage of the alignment at 100% and in those cases where isolates had more than one hit, only the ones with the higher percentage of identity were kept. Primers used to obtain the 16S rRNA genes of the isolates were different from the ones used to obtain the 16S rRNA iTAGs, but both amplified the V4 and V5 hypervariable region of the gene, so comparisons could be done by this method. For each dataset compared we calculated the mean percentage of reads or iTAGs, and zOTUs of the bacterial community that matched at 100% similarity with the 16S rRNA sequences of the strains isolated by traditional culture techniques. These percentages were calculated from the rarefied zOTU-abundance tables.

Genomes of ISS653 and ISS1889 and fragment recruitment analysis in marine metagenomes

Genomes of ISS653 and ISS1889 were sequenced and analysed by the Spanish Culture Collection of Type Strains (CECT). The accession number of ISS653 16S rRNA gene sequence and draft genome are MH732189 and CABVMM01, respectively, while the accession number of ISS1889 16S rRNA gene is MN836382. Detailed description of genome sequencing and analyses can be found in Lucena et al. (2020b). Metagenomic reads from some selected *Tara* Oceans stations (ST38: SUR (surface), DCM (deep chlorophyll maximum), MES (mesopelagic); ST39: DCM, MES; ST76: SUR, DCM, MES; ST102: SUR, DCM, MES; and ST151: SUR, DCM) were recruited competitively against the pool of the assembled contigs of the two isolates genomes. All metagenomes were subsampled to the shallower sequencing depth (129,995,612 fragments; mesopelagic from ST38) with *bbtools reformat.sh* (v38.08[3]). BLASTn v2.7.1+ (Altschul et al., 1990) was used to map the reads with the following alignment parameters: *-perc_identity* 70, *-evalue* 0.0001. Only those reads with more than 90% coverage and mapping at identities equal to or higher than 95% were considered to be true positives. In order to remove possible false hits mapping to the conserved regions of rRNA genes, reads aligning to the regions annotated as ribosomal genes were not considered for the analysis. Reads mapping with the same probability to any of the genomes were assigned at random.

Nucleotide sequences accession number

The 16S rRNA gene sequences of the bacterial isolates retrieved in this study were deposited in GenBank under the accession number MH731309-MH732621 and MK658870-MK659428.

[3]https://sourceforge.net/projects/bbmap/

Table 2: Summary of isolates, iOTUs, singletons and coverage per depth. Results derived from isolates 16S rRNA sequences clustering at 99% similarity to construct the non-rarefied and rarefied iOTU-abundance table (sampled down to the layer with the lowest number of isolates, i.e. mesopelagic with 362 isolates). Singletons: iOTUs appearing only once.

	99% (non-rarefied)			99% (rarefied)		
	Photic	Mesopelagic	Bathypelagic	Photic	Mesopelagic	Bathypelagic
Number of isolates	817	362	382	346	362	368
Number of iOTUs	100	57	59	61	57	59
Number of Singletons	39	25	20	18	25	20
Good's coverage	61%	56.1%	66.1%	70.5%	56.1 %	66.1%

Results

Taxonomic and phylogenetic diversity of the MARINHET culture collection

A total of 1561 bacterial strains were isolated from 19 marine stations, eight photic-layer, four mesopelagic, and seven bathypelagic samples (Figure 6A). The partial 16S rRNA sequences of the cultured strains were grouped into operational taxonomic units (isolated OTUs, referred hereafter as iOTUs) using 99% similarity thresholds. *Alphaproteobacteria* and *Gammaproteobacteria* iOTUs dominated in all stations (Figure 6B). *Bacteroidetes* isolates were present in all photic stations, but were not retrieved in the Indian mesopelagic sample ST39 or in the Atlantic bathypelagic samples ST10, ST33 and ST43. *Actinobacteria* isolates were retrieved only from six stations including photic and mesopelagic but not from bathypelagic samples. Finally, *Firmicutes* could be only isolated from photic samples of the Arctic and Indian Ocean during the time of sampling (Figure 6B).

If we group the different stations per depth, Good´s coverage analyses per layer, which is an estimator of the percentage of total species represented in a sample, ranged from 56.1 to 70.5% (Table 2). These results indicated that the isolates dataset, even if not saturated, represents a reasonable inventory of the culturable heterotrophic marine bacteria. The number of iOTUs detected was slightly higher in the photic layer for the non-rarefied iOTU table, but similar in all depths for the rarefied iOTU table, being the mesopelagic the layer with the lowest observed values (Table 2). Rarefaction curves showed also slightly higher richness for the photic samples compared to the mesopelagic and the bathypelagic, but they did not reach an asymptote (Figure 7A and S1A). On the other hand, rank abundance plots of the non-rarefied (Figure 7B) and rarefied iOTU tables (Figure S1B) presented, for the three depths studied, a steep curve, which is indicative of low evenness. Thus, there were a few abundant iOTUs with a large number of representatives and a large proportion of iOTUs that had few representatives (rare iOTUs). Therefore, we also calculated the richness and diversity metrics of each depth using OTU-based and phylogenetic approaches. All three metrics of OTU-based alpha diversity used (Species observed (S.obs) or nº of iOTUs, Chao1 and Shannon indexes, 7C) decreased with depth but not significative differences were found between layers (ANOVA test: S.obs: P-value = 0.152; Chao richness estimator: P-value = 0.191; Shannon diversity index: P-value = 0.183). The three measures of phylogenetic diversity, Faith's phylogenetic diversity (PD) (Faith, 1992), the PD divided by the number of iOTUs (PD/iOTUs), and the mean nearest taxon distance (MNTD) (Webb et al., 2002), were not significantly different between depths (ANOVA test: PD: P-value = 0.093; PD/iOTUs: P-value = 0.159; MNTD: P- value = 0.107), although

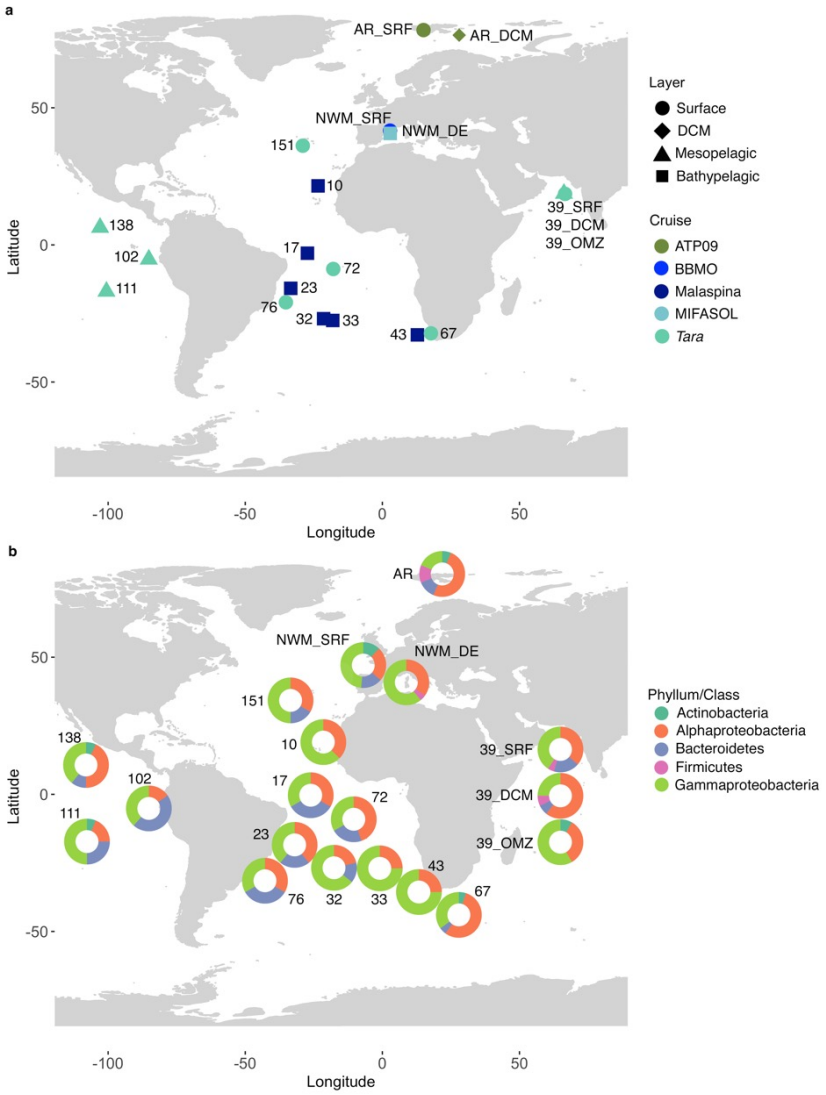

Figure 6: Map showing the sampling stations of the present study. **(A)** Position of the samples used for isolation. DCM: deep chlorophyll maximum. **(B)** Pie charts indicating the proportion of isolates retrieved affiliating with the different phyla, or classes in the case of *Proteobacteria*.

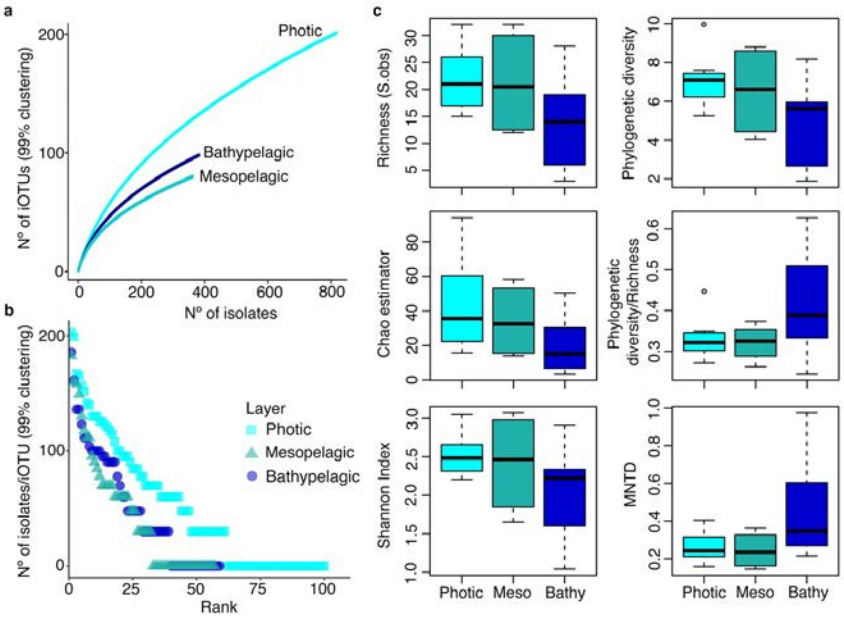

Figure 7: Diversity measures per layer studied. **(A)** Rarefaction curves extracted from the non-rarefied iOTU table (99% clustering). **(B)** Rank abundance plots showing the number of isolates per iOTU (at 99% clustering) obtained in the three layers studied also for the non-rarefied iOTU table. Y axis are in log10 scale. Photic: surface and deep chlorophyll maximum (DCM); Meso: mesopelagic; and Bathy: bathypelagic ocean. **(C)** Alpha-diversity measures using OTU-based (left panels) and phylogenetic (right panels) approaches. MNTD: mean nearest taxon distance.

a higher mean in phylogenetic diversity was observed in the photic layer than in the mesopelagic and the bathypelagic samples, while the phylogenetic diversity per iOTU and the MNTD was slightly higher in the bathypelagic layer (Figure 7C).

Shared diversity between photic, mesopelagic, and bathypelagic samples across oceans

We explored the similarity between iOTUs from different layers (photic, mesopelagic and bathypelagic). First, we started with samples from Indian Ocean ST39 because it was the only station with a vertical profile covering samples of the photic (surface and deep chlorophyll maximum, DCM) and the aphotic layer (mesopelagic). A total of 34 iOTUs were obtained from the independent clustering at 99% sequence similarity of all isolates from ST39. This clustering revealed that only 5 iOTUs were shared between photic and mesopelagic, while the rest could only be recovered from one depth, being the photic layer with the highest number of different iOTUs (Figure 8A), results that could be biased due to the higher presence of both surface and DCM isolates in the ST39 photic layer in comparison with the mesopelagic isolates. However, it was surprising that the shared iOTUs comprised the 63.6% of the total isolates (Figure 8A). At this point, we also examined the connectivity between different layers and across distant oceans covering large spatial and latitudinal scales. The non-rarefied iOTU table, including all the photic (817), mesopelagic (362), and bathypelagic (382) isolates, as well as the rarefied iOTU table, sampled down to the layer with the lowest number of isolates (mesopelagic), were used for the analyses, and because minor differences were observed among them (Tables S5 and S6), the results mentioned here refer only to those obtained after rarefaction. Fifteen out of 122 iOTUs (Figure 8B) included isolates from all layers, accounting for 52.7% of the total isolates sequences (Figure 8B), with an average number of 37.6 isolates per iOTU. Further, eight iOTUs (12.7% of the isolates) were common to photic and

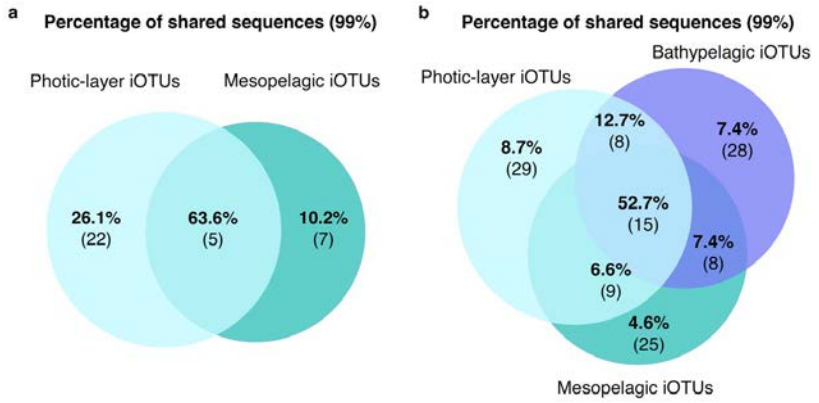

Figure 8: iOTUs retrieved from photic-layer and deep-sea waters. **(A)** Venn diagram representing the percentages of the sequences shared between photic and mesopelagic layer only from the vertical profile samples of the station 39. Numbers inside brackets indicate the number of shared iOTUs corresponding to that percentage of sequences. **(B)** Venn diagram showing the percentages of the sequences shared between photic, mesopelagic and bathypelagic layers. Numbers inside brackets indicate the number of shared iOTUs corresponding to that percentage of sequences. Numbers displayed in all Venn diagrams are extracted from the rarefied iOTU-abundance tables.

bathypelagic isolates, nine (6.6%) to photic and mesopelagic isolates, and eight (7.4%) to mesopelagic and bathypelagic isolates (Figure 8B). Nevertheless, as observed in ST39, a substantial proportion of isolates were only retrieved from one of the layers: 29 iOTUs were only found in the photic samples, 25 in the mesopelagic, and 28 in the bathypelagic samples (Table S7), with an average number of 3.2, 1.4, and 2.9 isolates per iOTU, respectively (Figure 8B).

The taxonomic classification of all these iOTUs, using the lowest common ancestor (LCA) method, designated a total of 59 different genera and 10 iOTUs that could not be classified at the genus level. From these 59 genera, 13 were widely distributed along the different depths studied representing 75% of the total isolates. On the other hand, the photic ocean was again the layer with the highest number of retrieved genera that were not observed in the other two depths, even though they accounted for only 5.6% of the isolates (Table S8).

If the comparative analysis is repeated with a more restrictive clustering, (instead of 99% at 100% similarity) we found that 37% of the isolates (578 out of 1561) were 100% identical at their partial 16S rRNA genes regardless the origin or layer. We found *Alteromonas, Cobetia, Erythrobacter, Leeuwenhoekiella, Halomonas, Idiomarina, Marinobacter*, and *Mesonia* between the shared genera, indicating taxa widely distributed along different depth layers. This shared percentage was even higher, up to 58.9%, when considering all mesopelagic and bathypelagic samples as aphotic and comparing them to all the photic isolates.

Biogeography of the commonly isolated heterotrophic bacteria

The most abundant and common culturable genera, i.e. those that occurred in all or most (around 80%) of the 19 stations studied, and the ones only retrieved locally (around 25% of the samples) with a restricted distribution were identified. *Erythrobacter* and *Alteromonas* were the most abundant and recurrent genera retrieved, representing 41.3% of the isolates (338 and 333 isolates respectively), and appearing in 94% of the samples studied regardless their origin, season and year of sampling (Figure 9A).

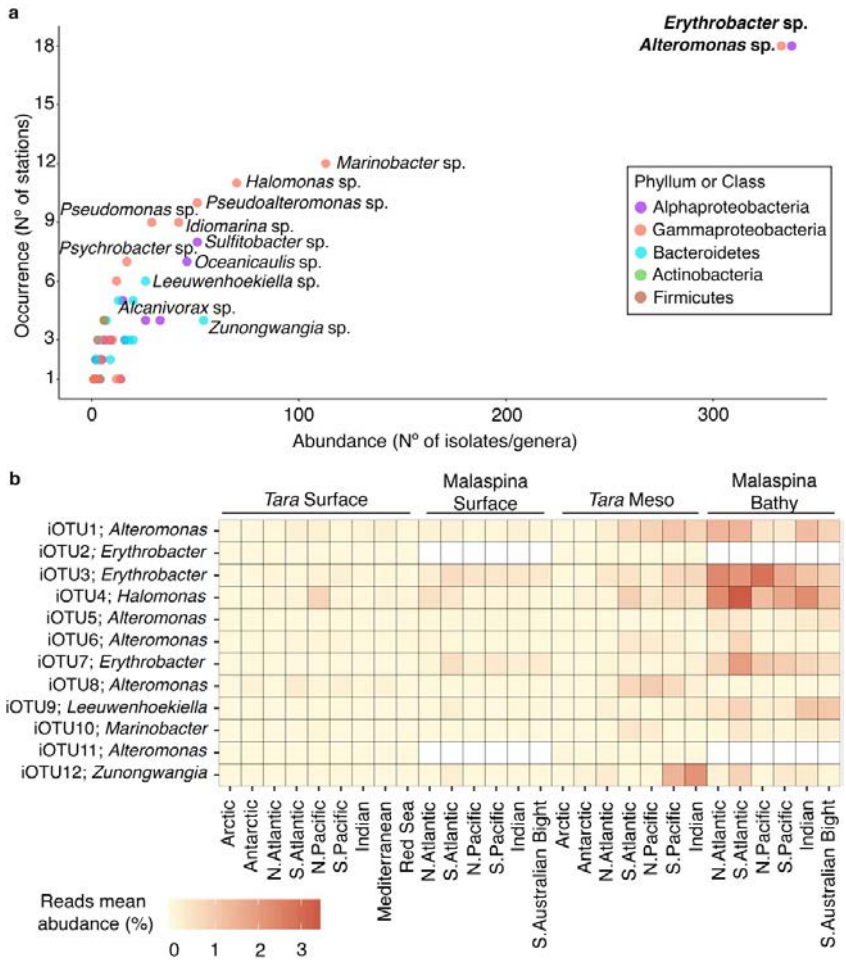

Figure 9: Abundance and biogeography of the isolates retrieved. **(A)** Abundance versus occurrence of the genera retrieved in the total culture collection. The most abundant and common genera are indicated in bold, and in regular type those with a more regional distribution. The color of the dots indicates the taxonomic (phylum or class) affiliation of the iOTUs. **(B)** Heatmap representing the mean abundance of reads (%) from zOTUs (zero-radius OTUs) of the top 12 isolated OTUs (rows) along the different oceanographic regions studied in the *Tara* and Malaspina expeditions samples (columns). Subsampled zOTU-abundance tables from the different datasets have been used.

Less abundant genera such as *Marinobacter* (113 isolates), *Halomonas* (70 isolates), *Pseudoalteromonas* (51 isolates), *Idiomarina* (42 isolates), *Pseudomonas* (29 isolates), *Sulfitobacter* (51 isolates), or *Oceanicaulis* (46 isolates) were present in more than 25% of the samples (Figure 9A) and covered almost all the oceanographic regions (Table S9). These could be considered, thus, regionally distributed. Some genera such as *Psychrobacter*, *Leeuwenhoekiella* or *Alcanivorax* had lower numbers of isolates, but were recovered in more than 25% of the samples (Figure 9A). Other genera, in turn, such as *Zunongwangia*, were retrieved in less than 25% of the samples but presented 54 isolates (3.5% of the strains). All the mentioned genera were found in the photic, mesopelagic, and bathypelagic layers, except *Oceanicaulis* which could not be isolated in this study from the bathypelagic samples (Table S9). Finally, the remaining genera represented 20% of the cultures. Then, these results revealed which genera are commonly isolated from distant stations with contrasted environmental conditions, depths and seasons covering 6 years of temporal range.

In a parallel study, we compared the isolates from each station with 16S rRNA sequences obtained through Illumina HTS of environmental DNA (16S iTAGs, hereafter) from two marine circumnavigations

(*Tara* Oceans (Karsenti et al., 2011) and Malaspina Expedition (Duarte, 2015)), to investigate whether our isolates have identical matches with environmental 16S iTAGs. Despite the global comparison is out of the scope of this study, and a detailed description of these analyses are explained in Chapter 2, here we present the biogeographic distribution of the abundant top12 iOTUs, those with more than 20 isolates (Table S2). To do so, we show the relative abundances of the denoised zOTUs (zero-radius OTUs, i.e. OTUs defined at 100% sequence similarity), obtained from the 16S iTAGs that matched at 100% similarity with these top12 iOTUs. We were aware that different iOTUs could match with the same zOTUs, and therefore the biogeography results presented here for the different zOTU represent the sum of the abundances of the top12 abundant iOTUs with other less abundant/rare iOTUs (Table S10). The top12 iOTUs matching at 100% with zOTUs represented the 48.3% of the total isolates (754 out of 1561), and only one of the zOTUs matched with two iOTUs, a top12 iOTU and a less abundant/rare iOTUs, the latter representing 6 out 754 (0.8 %) of the isolates included in the top12 iOTUs. Besides, both iOTUs matching to the same zOTUs affiliated with the same genus and, thus, the abundance presented would correspond to different species or ecotypes within a genus. Thereby, in the photic layer, the abundant top12 iOTUs, or in this case, their respective zOTUs matches, represented an average abundance of 16S iTAGs (at 100% similarity) lower than 1%, regardless of their geographic region (Figure 9B). This percentage increased around 1-2% of the reads in the mesopelagic layer in specific regions such as the Indian and South Pacific Ocean. However, our isolates exhibited higher abundances in the bathypelagic layer, in almost all oceanographic regions, especially for iOTU1, iOTU3 and iOTU4 affiliating with *Alteromonas, Erythrobacter* and *Halomonas* spp. (Figure 9B). Overall, these iOTUs accounted a total average proportion of reads of 0.3 and 1.1% in the photic layer in two independent datasets (*Tara* and Malaspina respectively), 2.7% in the mesopelagic layer, and up to 7.8% in the bathypelagic, indicating that the commonly found isolates are more abundant in the deeper layers of the ocean.

Novelty of the isolates of the MARINHET collection

The percentages of similarity between the strains and their Closest Cultured Match (CCM) and Closest Environmental Match (CEM) were extracted and compared with the 97% and 99% identity thresholds to explore the possible novelty of our culture collection. The results showed that most of the isolates were similar to previously published cultured bacterial species, but also to environmental sequences obtained using molecular techniques (Figure 10A). Therefore, most of the isolates were previously known microorganisms. However, we detected three 100% identical strains in their partial 16S rRNA gene that had a percentage of identity, both for CCM and CEM, below the threshold, at around 94%. One of the strains was isolated from surface samples of the North Atlantic Ocean (ISS653), whereas the other two were isolated from two mesopelagic samples of the Pacific Ocean (ISS1889, ISS2026). Further analyses with the complete 16S rRNA gene of ISS653 indicated that they could be candidates for a new species or even a new genus according to the thresholds proposed by Yarza et al. (2014). The three databases consulted (National Center for Biotechnology Information (NCBI), Ribosomal Data Base Project (RDP) and SILVA) showed different BLASTn results (Table S11). Nevertheless, the Living Tree Project (LTP) database, which contains the accepted type species of each genus, displayed a 93.5% similarity with *Mesonia mobilis*. The phylogenetic tree constructed (Figure 10B) also supported its novelty as our isolates had less than 93.8% of similarity with the cultivated reference genomes of the *Mesonia* genus.

Table 3: Origin of the novel *Mesonia* strains ISS653 and ISS1889 and phenotypic and genetic main differences. Differences extracted from a total of 158 tests (Supplementary Information Chapter 1). Within protein codifying sequences we included a list of some interesting proteins that were unique for one of the strains.

	ISS653	ISS1889
Origin		
Station	ST 151	ST 102
Ocean	North Atlantic Ocean	South Pacific Ocean
Depth (m)	5	475.6
In situ Temperature (°C)	17.3	9.2
Phisiological differences		
Maximum temperature for	37	30
Tween-80 hydrolysis	weak	-
Phenylacetate assimilation	+	-
Acid from (API50CH/E)		
arbutin	-	weak
2-ketoglutarate	-	weak
Cellular fatty acids		
iso-$C_{15:0}$ 2OH	11.2	14.5
iso-$C_{17:0}$ 3OH	9.2	14.6
iso-$C_{17:1}$ w9*c*	5.2	10.5
$C_{15:1}$ w5*c*	3.6	Traces (<1%)
$C_{18:1}$ w9*c*	3.9	-
Genomic differences		
Genome size (bp)	4275762	4283636
G+C content (mol%)	34.9	34.9
RNAs	45	45
Protein codifying sequences		
Total	4030	4015
Chaperones	GroEL, GroES, ClpB	-
Cobalt-zinc-cadmium resistance	CusA, CzcA, CzcD	-
Mercury resistance	-	MerA, MerT

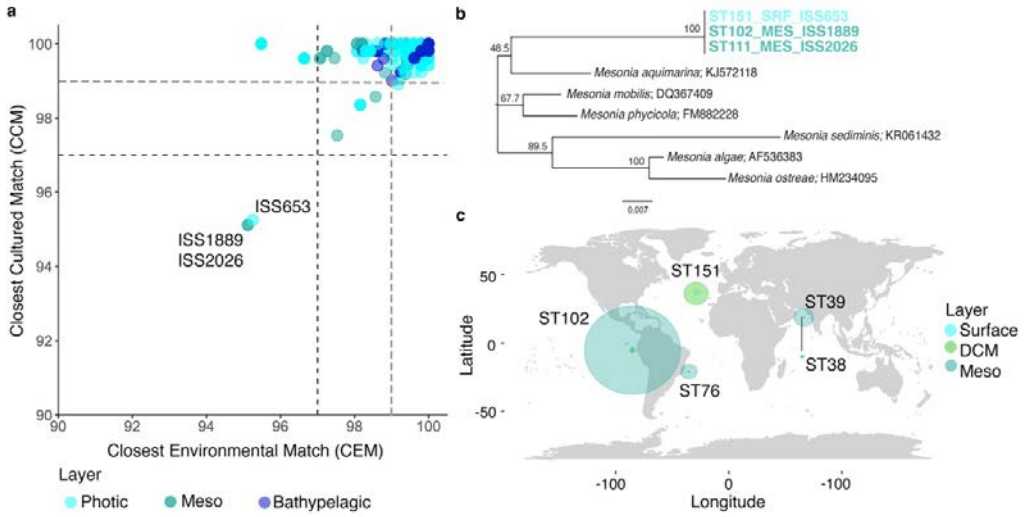

Figure 10: Potential novel isolates. **(A)** Percentages of similarity between the Closest Cultured Match (CCM) and the Closest Environmental Match (CEM) of all the 16S rRNA gene sequences. Horizontal and vertical lines represent the typical cut-off value of 97% (black dashed lines) and 99% (grey dashed lines) commonly used for "species" delineation. **(B)** Neighbour Joining tree of the putative *Mesonia* isolates. The numbers in the nodes represent bootstrap percentages > 45%, calculated from 1000 replicates. Putative new isolates are written in bold letters and color indicates origin of the isolates. **(C)** Read recruitment of ISS653 and ISS1889 in 5 *Tara* Oceans stations. They include the stations where the isolates were retrieved (ST151 and ST102) and some distant stations for the sake of comparison (ST39, ST38, ST76). ST38 is located near ST39 (Latitude 19º 2.24' N, Longitude 64º 29.24' E), but its location in the plot was slightly modified for its correct visualization. Size of the circles are the sum of the abundance of reads from both genomes recruited in each station and layer (x10k). SRF, surface isolates; DCM, deep chlorophyll maximum; Meso, mesopelagic isolates.

Genomes of the strains ISS653 and ISS1889 were fully sequenced and characterized to formally describe a novel species, *Mesonia oceanica* (Lucena et al., 2020b). As a detailed description of the novel species is already given in Lucena et al. (2020b), here we only focused in some interesting phenotypic differences among these two strains and their distribution pattern in marine metagenomes from five stations, including ST102 and ST151, that were the ones in which ISS1889 and ISS653 were respectively isolated. Only a few phenotypic differences could be found among both strains (Table 3). The most important phenotypic trait was the difference in their maximum growth temperatures, being 37ºC for ISS653, isolated from surface waters in ST151, and 30ºC for ISS1889, isolated from the mesopelagic layer in ST102. Genomic comparisons of both strains revealed an average nucleotide identity (ANI) of 99.9%, which indicated that the two strains were almost identical genetic clones. The G+C content and the number of RNAs present were equal in both strains. They slightly differ in the size and the number of protein codifying sequences (Table 3). However, we identified a pool of unique genes for each strain, 33 were only found in ISS653, whereas 6 were unique in ISS1889 (Table S12). Among them, we found interesting some proteins that may confer specific advantages and/or adaptation (Table 3). For example, ISS653 contains some chaperones, GroEL, GroES and ClpB, that may be related with its wider range of growth temperatures. In addition, we detected some resistant mechanisms to toxic heavy metals. Resistance genes to cadmium-zinc-cobalt were detected in ISS653, while mercury resistance genes were observed in ISS1889. Nevertheless, we found that both isolates presented the same distribution patterns (Figure 10C and Figure S2). Thus, these two new strains displayed higher abundances in the mesopelagic waters regardless of the station analysed, but especially in ST102 where ISS1889 was retrieved (Figure 10C). Strain ISS653 was isolated from surface waters of the ST151 but it was in deeper layers where its abundance was also higher (Figure S3).

Discussion

We have elaborated an extensive marine heterotrophic bacterial culture collection with 1561 isolates covering different oceanographic regions, depths, seasons and years. We used an standard marine medium to reach the heterotrophic fraction of the community that could be comparable between layers (photic and aphotic) and across oceans, rather than using other specific media for increasing the novelty on isolates in the deep ocean. Even though we could not fully address distribution patterns along complete latitudinal gradients for all depth layers studied or some seasonal/temporal changes, we could explore the phylogenetic diversity of the MARINHET culture collection and analyse the potential differences between depths. The alpha-diversity metrics were slightly higher in the photic layer, but not significant differences between layers were found. On the other hand, rank abundance curves from different depths showed that the fraction of the heterotrophic isolates retrieved were composed by a few abundant iOTUs with a large number of representatives and many rare low abundant iOTUs, which, in this case, is consistent with many other previous findings based on prokaryotic amplicon 16S iTAGs from environmental samples (Sogin et al., 2006; Pedrós-Alió, 2012). For instance, the 7 most abundant iOTUs (99%) accounted for 41% of the total isolates and similar proportions were found in each layer. Hence, 30% of the bathypelagic isolates, 47% of the mesopelagic isolates, and 43% of the photic isolates affiliated with these seven most abundant iOTUs.

The comparison between those isolates coming from different depths allowed us to detect certain level of vertical connectivity among the heterotrophic culturable community. The significant overlap found between photic, mesopelagic and bathypelagic strains suggest that these heterotrophic bacteria are well adapted to different temperatures, light and pressure. Moreover, they probably have versatile metabolisms to respond to different environments and nutrient availability. These characteristics may make these bacteria more prone to successfully face such long vertical and horizontal dispersion (Jones and Lennon, 2010). In addition, genomic comparison between cultured isolates and uncultured genomes retrieved by single amplified genomes (SAGs) from marine environments revealed that the genomes of the cultures had larger sizes, suggesting a predominant copiotrophic lifestyle (Swan et al., 2013). One possible explanation supporting the high proportion of identical 16S rRNA gene sequences between isolates of photic and aphotic layers, up to 58.9%, would be that these bacteria have the capacity to attach and grow on particles in the photic layers and after sinking to the deep ocean, they still retain the capability for further growth. Certainly, a recent study claimed that the particle colonization process that takes place in the photic layers determines the composition of deeper layers and especially bathypelagic communities, and thus, photic and deep-ocean prokaryotic communities are strongly connected via sinking particles (Mestre et al., 2018). Moreover, the attachment to particles and its presence in the deep ocean has been described at least for *Alteromonas* (Acinas et al., 1999; Crespo et al., 2013; Salazar et al., 2015; Mestre et al., 2018), *Erythrobacter* and *Halomonas* (Mestre et al., 2018).

Those mentioned genera are also the most abundant and commonly isolated in all depths of our dataset together with *Marinobacter*. These genera have been detected in other culture-dependent and culture-independent studies from a wide variety of marine environments, including coastal, shelf, and open ocean waters (Selje et al., 2005; Zeng et al., 2012; Lekunberri et al., 2014; Tonon et al., 2014; Salazar et al., 2015, 2016; Crespo et al., 2016) corroborating their ubiquity. *Alteromonas* and *Erythrobacter* presented the highest number of isolates. *Alteromonas* is among the most common culturable heterotrophic bacteria living in open marine waters all around the world, as it has been isolated from a wide

variety of marine environments (Baumann et al., 1972; Fuhrman and Davis, 1997; García-Martínez et al., 2002; Gärtner et al., 2011; Kai et al., 2017). In addition, this genus is thought to be one of the most significant contributors of the dissolved organic carbon (DOC) consumption and nutrient mineralization in the upper ocean (Pedler et al., 2014). *Erythrobacter* strains are aerobic chemoorganotrophs, and some species contain *bacteriochlorophyll a*, responsible for the aerobic anoxygenic phototrophic (AAP) metabolism (Tonon et al., 2014).

Despite these findings, one of the remaining questions, mainly in relation to the commonly isolated bacteria, is to what extent these strains match with environmental 16S rRNA genes from HTS sequencing of the whole bacterial community. The comparison of the top abundant iOTUs sequences with 16S iTAGs confirmed that these common iOTUs matched at 100% identity with environmental sequences at different extent, being rare at the surface but with increasing representation in the deep ocean, especially in the bathypelagic (in detail in Chapter 2 of this thesis).

On the other hand, even though the isolation of novel strains was a secondary objective, thanks to the large isolation effort done, we managed to isolate three strains, 100% similar among them in their partial 16S rRNA gene, that presented less than 95% of similarity in their 16S sequence to any previously described bacterial species. There are several well-accepted criteria for the classification of bacteria into species and one of them is based in the 16S rRNA gene sequence identity threshold at around 98.7-99% (Acinas et al., 2003; Stackebrandt and Ebers, 2006; Yarza et al., 2014). Two of these strains, ISS653 and ISS1889, had been fully characterized recently and their genome has been sequenced by the Spanish Culture Collection of Type Strains (CECT) to formally describe a novel species, *Mesonia oceanica* (Lucena et al., 2020b). Members of this taxon are mainly retrieved from a variety of marine environments, sometimes associated with eukaryotic organisms, such as algae (Nedashkovskaya et al., 2006). Interestingly, isolate ISS653 was obtained from the surface North Atlantic waters whereas ISS1889 was retrieved from mesopelagic waters of the Pacific Ocean, and the biogeography analysis in some vertical profiles suggest that this putative novel species is not locally restricted and it has preference for deeper layers. Curiously, even though we could not detect any significant difference in the distribution of these two strains and genetically they seem almost clonal, we cannot discard that their minimal genetic differences and phenotypic plasticity may provide adaptation advantages (of growth rate and tolerance to metals) under particular environmental conditions not observed in our marine samples.

Finally, as we stated at the beginning of this study, cultures are important because they allow to retrieve novel bacterial taxa and complete genomes, but most importantly, they enable to test hypotheses that emerge from genomic data. However, the isolation of a great battery of strains from different oceanographic regions and depths by traditional culture techniques, as we presented, does not guarantee the retrieval of many new bacterial species or taxa detected by HTS techniques, which greatly outnumber those accessible by cultivation. One of the challenges that marine microbial ecologists still face is the innovation in the isolation methods for the retrieval of axenic cultures of those uncultured taxa. In the last years, new isolation approaches had been developed to improve the recovery of bacteria under laboratory conditions like microfluidics (Ma et al., 2014; Boitard et al., 2015), cultivation chips (Ingham et al., 2007; Hesselman et al., 2012; Gao et al., 2013), microcolony cultivation techniques (Ferrari et al., 2005), manipulation of single cells (Ben-Dov et al., 2009; Park et al., 2011), and high-throughput cultivation techniques named "culturomics" (Giovannoni and Stingl, 2007; Lagier et al., 2012). Nevertheless, all these strategies are usually expensive and include inherent trial-and-error approaches. In this manner, metagenomic, metatranscriptomic and metaproteomic data, which had increased our knowledge of

the microorganisms present in marine ecosystems and allowed to predict their metabolic capabilities, would provide essential information to design different isolation strategies and allow the retrieval of environmental bacteria. Accordingly, understanding the microbial complexity of the marine ecosystems would be possible if combined culture-dependent and culture-independent studies start to be the rule among marine microbial ecologists.

Conclusions

In summary, culturing remains an important tool in microbial ecology, helping to map the diversity of marine communities. We are aware that our study is restricted to those heterotrophic marine bacteria that can grow in standard culture conditions, and that we are missing many other fundamental microbial populations that do not grow easily in standard marine media. Nevertheless, given the important isolation effort done and the number of oceanic regions and depths covered in different years, we were able to enhance our knowledge of the taxonomy, phylogenetic diversity and distribution of the targeted bacteria. Equally to those HTS studies of ribosomal genes targeting the whole marine prokaryotic community, the culturable marine heterotrophic bacteria isolated presented few abundant taxa and a tail of rare and low abundant iOTUs. We detected that half of the total isolates were shared in the three different depth realms, reinforcing the already introduced idea of vertical connectivity between the photic and the deep ocean probably through sinking particles. In addition, we identified *Alteromonas* and *Erythrobacter* genera to be the most abundant and commonly isolated heterotrophic bacteria from more than 80% of the studied samples and from all layers. Finally, we found three strains belonging to a new species of the genus *Mesonia*. Overall, this study highlights the relevance of complementary studies with focus on marine isolated bacteria to provide a more comprehensive view of marine microbial diversity. Furthermore, our MARINHET culture collection represents a valuable resource for future genome sequencing projects and potential physiological experiments involving marine isolates.

Chapter 2

Contribution of marine cultured heterotrophic bacteria to the dark ocean microbial diversity

Abstract

It is a long-standing observation that traditional culture techniques only isolate a small fraction of the microbial diversity, which mainly belong to the so-called rare biosphere, however, this paradigm has not been fully explored in the deep ocean. In this study, we examined the fraction of heterotrophic bacterial diversity from the photic and the deep-ocean layers that could be retrieved by culture-dependent techniques at large scale. Therefore, we used a large collection of heterotrophic marine cultured bacteria (MARINHET_v2) of a total of 2003 isolates covering different oceanographic regions and depths. The partial 16S rRNA gene (average 500 bp of the V3-V4 region) of this culture dataset was then compared with denoised OTUs (zOTUs) identified from amplicon 16S rRNA gene Illumina TAGs (16S iTAGs) from *Tara* Oceans and Malaspina 2010 Expedition, and to metagenomic OTUs (mOTUs) extracted from metagenomic TAGs (miTAGs) from Malaspina 2010 Expedition, including overall surface, mesopelagic and bathypelagic datasets. Comparative analyses of amplicon 16S iTAGs and miTAGs revealed similar patterns for photic and aphotic layers including a higher proportion of isolates identical to zOTUs or mOTUs in the deeper layers, although significant differences displayed in their abundances. Based on amplicon 16S iTAGs, a 28% of the reads of the microbial community at the bathypelagic ocean could be retrieved by heterotrophic isolated bacteria. Also, these isolates seem to contribute significantly to the bacterial communities attached to particles recruiting 21% of the zOTUs and 45% of the reads of the largest plankton size fraction (>20 µm) in the bathypelagic. These results of a higher proportion of isolated zOTUs described in the largest size particles specially in the bathypelagic, are in line with the idea that slow sinking particles in the bathypelagic realm act as resource-rich habitats, which are suitable for isolating heterotrophic bacteria with a copiotroph lifestyle and represent an important fraction of the total prokaryotic diversity inhabiting the deep ocean.

Introduction

Microbial ecology has been long challenged by the fact that only a small fraction of the natural bacterial community can be cultivated, a phenomenon that has been traditionally called "the great plate count anomaly" (Staley and Konopka, 1985). The recovered proportion of cells using selective media and standard plating techniques when compared to both microscopy analyses by direct staining and 16S rRNA sequencing, only represented among 0.001-1% of the community (Kogure et al., 1979; Staley and Konopka, 1985; Amann et al., 1995). This phenomenon lead to the known paradigm that "less than 1% of the bacterial community can be cultured" (Razumov, 1932; Jannasch, 1958; Staley and Konopka, 1985; Eguchi and Ishida, 1990). Culture-independent techniques, such as high-throughput sequencing (HTS) of ribosomal genes (Sogin et al., 2006; Zinger et al., 2011), and genomic centric studies with the exploration of single amplified genomes (SAGs) (Stepanauskas and Sieracki, 2007) and the analyses of metagenome-assembled genomes (MAGs) (Parks et al., 2017) have provided great advances in the description of microbial diversity at an unprecedent scale (Parks et al., 2018; Pedrós-Alió et al., 2018). Nonetheless, culture-dependent isolation is still a valuable resource to decipher a fraction of the diversity not targeted by molecular/omics approaches and, therefore, necessary to fully understand the holistic view of microbial diversity (Eloe et al., 2011a; Gärtner et al., 2011; Kai et al., 2017; Sanz-Sáez et al., 2020). Moreover, isolates are so much needed to perform physiological experiments and test ecological hypothesis (Giovannoni and Stingl, 2007; Gutleben et al., 2017). The seminal studies on marine prokaryotic diversity comparison between culture-dependent and culture-independent techniques, revealed, in most cases, that there is a few overlap between isolates and environmental sequences (Floyd et al., 2005; Zeng et al., 2012; Lekunberri et al., 2014), and that most bacterial strains growing under laboratory conditions belong to the rare biosphere (Shade et al., 2012; Crespo et al., 2016). Nevertheless, these studies are scarce, and mostly focused on the photic ocean (0-200 m) (Eilers et al., 2000; Zeng et al., 2012; Lekunberri et al., 2014; Crespo et al., 2016), or in specific deep-ocean ecosystems like hydrothermal vents (Harmsen et al., 1997; Huber et al., 2002; Hirayama et al., 2007), hypersaline deep-sea basins (Sass et al., 2001), or deep-sea sediments (Gutierrez et al., 2015; Chen et al., 2016), leaving the mesopelagic and bathypelagic waters less explored. Recently, some studies have compared the proportion of prokaryotes that can be isolated in different ecosystems, including marine environments, using genomes of isolates public available against amplicon 16S rRNA gene Illumina TAGs (16S iTAGs), metagenomes and metatranscriptomic public datasets (Lloyd et al., 2018; Steen et al., 2019; Martiny, 2019). However, they used different genetic thresholds and methods for such comparisons generating contrasted results, and the deep ocean was poorly examined. Therefore, the long-standing observation that traditional culture techniques only retrieve a small fraction of the microbial diversity in marine environments needs to be revised and properly tested in the pelagic deep ocean.

Previous studies targeting the diversity of microbial communities from the bathypelagic ocean across the tropical and temperate oceans described the most abundant operational taxonomic units (OTUs) (Salazar et al., 2016). Interestingly, some of these OTUs affiliated with well know heterotrophic cultured genera including *Alteromonas*, *Alcanivorax* or *Halomonas* (Salazar et al., 2016; Kai et al., 2017; Sanz-Sáez et al., 2020). In addition, some deep-ocean marine bacteria such as the genera *Alteromonas* have been defined as copiotrophs (López-Pérez et al., 2012) which may favor their isolation in pure culture. Moreover, given that the bathypelagic realm is constituted mainly by slow sinking particles (Herndl and Reinthaler, 2013), which are resource-rich habitats for microbes (Bochdansky et al., 2016), this copiotrophic lifestyle could be the rule for an important fraction of the bacteria inhabiting the deep

ocean. Taking these facts altogether, it led us to hypothesize that maybe higher proportions of bacteria dwelling in the deep ocean could be retrieved under laboratory conditions. Thus, in this study we examined the fraction of heterotrophic microbial diversity that could be retrieved by isolation in both the photic and deep oceans. We aimed to: (i) rank the isolated bacteria among members of the abundant or rare biosphere, (iii) test if the "great plate count anomaly" applies in the deep ocean, and (iii) analyze weather the isolated heterotrophic bacteria are more prone to be associated to free-living or to particle-attached bacterial communities at different depths. To that end, we compared a large collection of heterotrophic cultured bacteria (MARINHET_v2), covering a wide range of oceanographic regions and depths, including the photic, the mesopelagic and the bathypelagic ocean (Sanz-Sáez et al., 2020), with its corresponding environmental sequences identified from amplicon 16S iTAGs and metagenomic 16S TAGs (miTAGs) from two global oceanographic expeditions, *Tara* Oceans 2009 (Karsenti et al., 2011) and Malaspina Circumnavigation Expedition 2010 (Duarte, 2015). Hence, we present here the first attempt to combine results from culture-dependent and high-throughput sequencing techniques, from the same samples and others worldwide distributed, and to test the culturability of natural bacterial communities at different oceanic depths, including the underexplored mesopelagic and bathypelagic layers.

Materials and methods

Seawater sampling collection and filtration in *Tara* Oceans and Malaspina Expedition

During the Malaspina Expedition, surface seawater (3 m) was collected using 20 L Niskin bottles, prefiltered through a 200 µm mesh to remove large plankton, and finally, between 6-15 L were sequentially filtered through a 20 µm mesh, a 47 mm polycarbonate membrane filter of 3 µm (Merck Millipore, Darmstadt, Germany, Isopore polycarbonate) and a 0.2 µm (Merck Millipore, Express Plus) filter pore size with a peristaltic pump. For these surface samples we focused on the 0.2–3 µm fraction, which represents mostly free-living bacteria (Ruiz González et al., 2019). On the other hand, for the bathypelagic samples (approximately 4000 m depth), 120 L of seawater were also sequentially pre-filtered through a 200 µm and a 20 µm mesh to remove large plankton. Further filtering was done by pumping water serially through 142 mm polycarbonate membrane filters of 0.8 µm (Merk Millipore, Darmstadt, Germany, Isopore polycarbonate) and 0.2 µm (Merck Millipore, Express Plus) pore size with a peristaltic pump (Masterflex, EW-77410-10). In this case two different size fractions were analyzed representing the free-living (0.2-0.8 µm) and the particle-attached (0.8-20 µm) microbial communities (Salazar et al., 2016). Additionally, samples from eight vertical profiles were also collected and were size fractioned using five different size pore filters in order to obtain microbial communities of different size fractions (Mestre et al., 2018). Surface water from these profiles was sampled with Niskin bottles, whereas water from the other depths were sampled with Niskin bottles attached to a rosette sampling system. Thus, seawater was pre-filtered by a 200 µm mesh to remove large plankton, and then 10 L were sequentially filtered through 20-, 5.0-, 3.0-, 0.8-, and 0.2 µm pore-size filters, all 47 mm polycarbonate filters (20 µm pore-size filter from GE Water and Process Technologies and the rest of the filters from Millipore), using a peristaltic pump, resulting in five different size fractions (0.2–0.8, 0.8–3.0, 3.0–5.0, 5.0–20, and 20–200 µm). Filters from all surface, bathypelagic and vertical profile samples were then flash-frozen in liquid nitrogen and stored at -80 ºC until DNA extraction.

In *Tara* Oceans surface and mesopelagic seawater was also collected using Niskin bottles attached to a rosette sampling system, prefiltered by 200 µm and 20 µm meshes to remove large plankton and filtered by 3 µm or 1.6 µm pore size filters. In this dataset we targeted mostly the free-living bacterial communities compressed within the 0.2 and 1.6 or 3 µm (Sunagawa et al., 2015; Ibarbalz et al., 2019). Filters were then flash-frozen in liquid nitrogen and stored at -80 ºC until DNA extraction. Further sampling strategy and methodology are described in Pesant et al. (2015).

Samples for Illumina amplicon 16S rRNA sequencing

Different datasets containing Illumina amplicon 16S rRNA (16S iTAGs) gene sequences were analyzed in order to compare them with partial 16S rRNA sequences from isolates obtained by traditional culture techniques. The fist dataset from Malaspina Expedition comprised a total of 124 surface samples and 41 bathypelagic samples distributed across the temperate and tropical world's oceans (Figure 11). A second dataset, also from Malaspina Expedition, included eight stations with vertical profiles and different microbial size fractions (Figure 11). At each of these stations four depths were sampled corresponding to surface (3 m), deep chlorophyll maximum (DCM, 48-150 m), mesopelagic (250-670 m), and bathypelagic waters (3105-4000 m). The third dataset from *Tara* Oceans 2009 and *Tara* Oceans Polar Circle 2013 expeditions was formed by 80 surface samples and 39 mesopelagic samples which covered the major oceanic provinces including the polar circles (Figure 11A, B).

Samples for Illumina metagenome sequencing

Illumina metagenomic sequencing was performed with 11 vertical profiles collected during the Malaspina Expedition. For this study, we used 10 surface and 11 bathypelagic samples from the vertical profiles globally distributed across the world's oceans. Bathypelagic samples were collected mainly at the depth of 4000 m, although a few samples were taken at shallower depths, all within the bathypelagic realm (average depth: 3731 m ± 495; standard deviation) in the tropical and subtropical oceans (Figure 11C).

Samples for heterotrophic bacterial isolation

A total of 14 photic-layer stations and 11 deep-ocean stations (including 4 from the mesopelagic in areas with oxygen minimum zone (OMZ) regions and 7 from the bathypelagic) were collected from four different oceanographic expeditions including the Malaspina Expedition, *Tara* Oceans (*Tara* Oceans 2009 and *Tara* Oceans Polar Circle 2013), ATP09 (Lara et al., 2013) and MIFASOL, and also from the Blanes Bay Microbial Observatory BBMO[4], covering then wide latitudinal and oceanographic regions (Figure 11). Details regarding samples collection has been previously described in Sanz-Sáez et al. (2020). Geographical coordinates of stations, sampled depth, *in situ* temperature, number of sequenced isolates, total prokaryotes cell abundances, and count of cfu/ml are listed in Table S13. Prokaryote cell abundance was determined using flow cytometry (BD FACSCalibur) of SYBR Green I stained samples (Gasol and del Giorgio, 2000).

[4]http://www.icm.csic.es/bio/projects/icmicrobis/bbmo

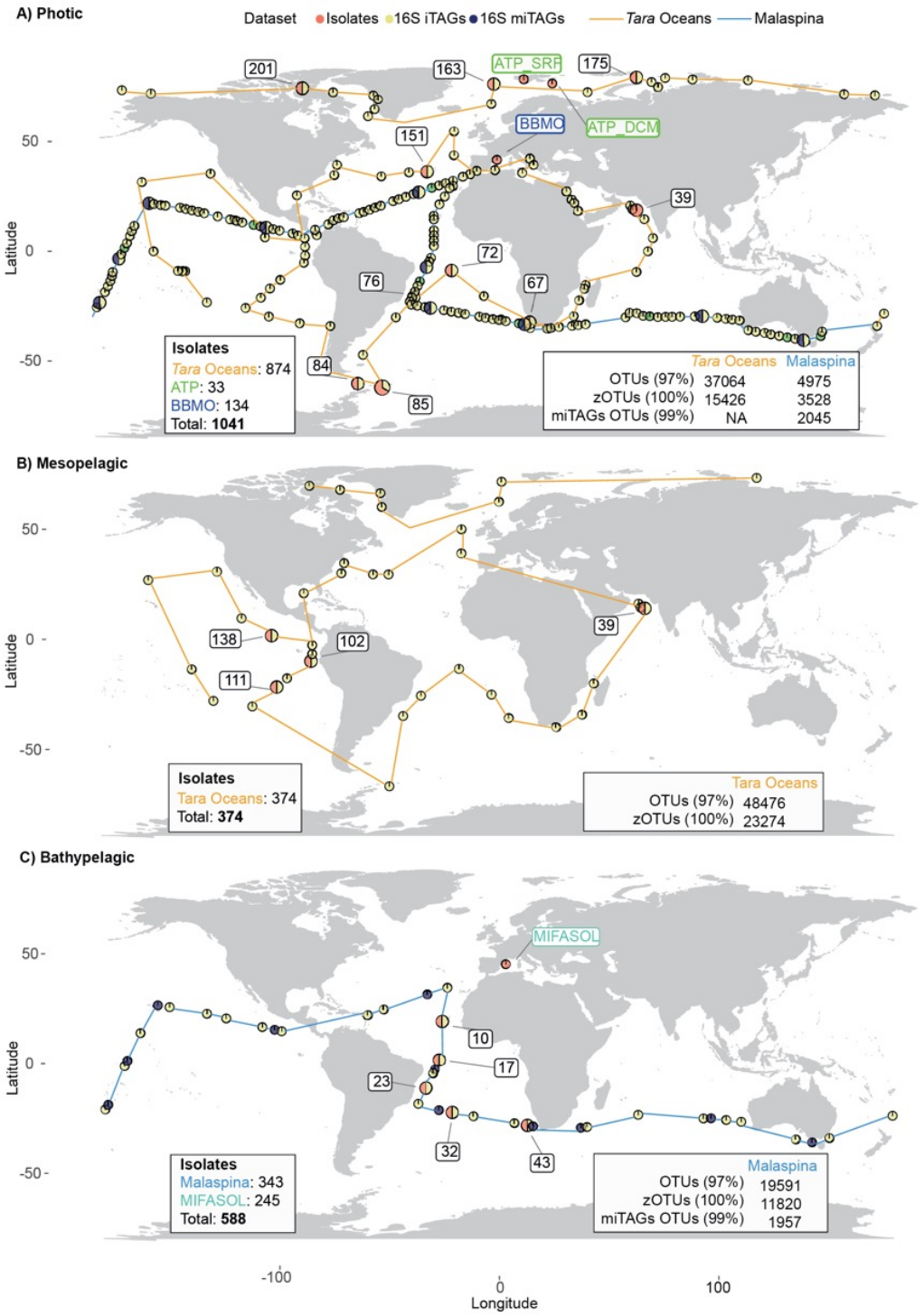

Figure 11: World map showing the distribution of the samples used in this study per layer. **(A)** Photic. Labeled samples correspond to those stations where isolates were obtained from: *Tara* Oceans (39, 67, 72, 76, 84, 85, 151, 163, 175,201), Blanes Bay Microbial Observatory (BBMO), and ATP Arctic cruise (ATP_SRF, ATP_DCM). Stations colored green correspond to amplicon 16S iTAGs from eight vertical profiles with five different microbial size fractions collected from the Malaspina Expedition. **(B)** Mesopelagic. Labeled samples correspond to those stations where isolates were obtained from the *Tara* Oceans (39, 102, 111, 138). **(C)** Bathypelagic. Labeled samples correspond to those stations where isolates were obtained from: Malaspina (10, 17, 23, 32, 43) and MIFASOL. Circles connected with a blue line show the distribution of the samples obtained from the Malaspina Expedition, while circles connected with an orange line show those from the *Tara* Oceans. Each pie chart shows the presence or absence of samples from the different datasets: orange, isolates; light yellow, 16S iTAGs; dark blue, 16S miTAGs.

Isolation and cultivation of heterotrophic bacteria

Isolates were mainly retrieved by plating 100 µl of undiluted and 10x diluted seawater from the 14 photic-layer, 4 mesopelagic and 7 bathypelagic samples. Additionally, isolates from the bathypelagic sample collected in the NW Mediterranean during the MIFASOL campaign were obtained using also different culturing strategies: (i) enrichment in Zobell broth (5 g peptone, 1 g yeast extract in 750 ml of 30 kDa filtered seawater and 250 ml of Milli-Q water) of the seawater previously concentrated by tangential flow filtration, subsequently diluted and plated, (ii) direct plating of 0.8 and 0.2 µm pore size filters into plates, (iii) after resuspension of 0.8 and 0.2 µm pore size filters in 20 ml of 30 kDa filtered seawater, dilution and subsequent plating, and (iv) after enrichment of the 0.8 and 0.2 µm pore size filters in 10 ml of Zobell broth, diluted and plated. All samples were plated in triplicates in nutrient rich media including Zobell agar, Marine Agar 2216 and modified Marine Agar, where disodium phosphate was autoclaved apart from the rest of the medium and added as a separate solution before solidification. Detailed composition of the media has been previously described (Zobell, 1941; Sanz-Sáez et al., 2020). Our medium culturing strategy was only focused to retrieve heterotrophic marine bacteria that could grow easily under laboratory conditions (nutrient rich media, standard oxygen concentrations and atmospheric pressure).

Photic-layer samples from temperate oceans and mesopelagic samples were incubated at room temperature (RT), while photic-layer samples from the Artic and Southern Oceans and bathypelagic samples were incubated at their *in situ* temperature, which ranged from approximately 4ºC (photic-layer samples in the Artic and Southern Ocean, and 4000 m depth in the Atlantic Ocean) to approximately 12ºC (NW Mediterranean at 2000 m depth) (Table S13 and Table S14), but also at RT in order to assure bacterial recovery from all stations. In all cases, triplicates of each temperature condition and dilution were incubated in the dark until no more colonies appeared (10-30 days). The number of cfu/ml growing in the plates was calculated counting all the colonies present on the agar plates at the end of the incubation times. For calculation at least two replicates per incubation condition and dilution were considered.

A total of 2003 bacterial isolates (MARINHET_v2 culture collection) were randomly selected for DNA amplification and partial sequencing of their 16S rRNA gene. Similar number of isolates were selected from photic layers (1041; average: 70.6 isolates per station) and from the deep ocean (962; average: 67.6 isolates per station). Colonies were streaked on agar plates in duplicate to ensure their purity and avoid contamination. The isolates were stored with glycerol 25% in cryovials at -80ºC.

DNA extraction and amplification

Available DNA from isolates used for template in Polymerase Chain Reaction (PCR) was extracted from 200 µL of liquid cultures placed in 96 well plates, diluted 1:4 and heated (95ºC, 15 min) to cause cell lyses. The partial 16S rRNA gene sequences were PCR amplified using bacterial primers 358F (5'-CCT ACG GGA GGC AGC AG-3') (Muyzer et al., 1993) and 907Rmod (5'-CCG TCA ATT CMT TTG AGT TT-3') (Sánchez et al., 2007). Detailed PCR mix preparation and reactions are described in Sanz-Sáez et al. (2020). Purification and OneShot Sanger sequencing of the partial 16S rRNA gene products was performed by Genoscreen (Lille, France) with primer 358F. ChromasPro 2.1.8 software (Technelysium) was used for manual cleaning and quality control of the sequences.

The DNA from the samples of the three different datasets described for Illumina sequencing was extracted with a phenol-chloroform protocol (as described in Massana et al. (1997); Salazar et al. (2016); Alberti et al. (2017)). Prokaryotic barcodes for each of the datasets were generated by amplifying the V4 and V5 hypervariable regions of the 16S rRNA gene using primers 515F-Y (5'-GTGYCAGCMGCCGCGGTAA-3') and 926R (5'-CCGYCAATTYMTTTRAGTTT-3') described in Parada et al. (2016). Sequencing was performed in an Illumina MiSeq platform (iTAGs) using 2x250 bp paired-end approach at Genoscope for the Malaspina Bathypelagic, the *Tara* Oceans and the *Tara* Polar Oceans datasets, and at the Research and Testing Laboratory facility [5] for the Malaspina Surface and Malaspina vertical profiles datasets, from now on called the Malaspina size fractions dataset.

Metagenomic DNA was extracted as described before (Logares et al., 2014). DNA was sequenced using pair-end Illumina sequencing technology (Illumina, USA) and high-quality (HQ) reads were obtained using MOCAT (v. 1.2) (Kultima et al., 2012) as previously described (Sunagawa et al., 2015).

Sequence data processing

Amplicon 16S rRNA iTAGs:

The obtained amplicons were processed through the bioinformatic pipeline described in the SushiLab Amplicon Recipes github repository[6]. Briefly, pair-end reads were merged at a minimum 90% of identity alignment, and those with ≤ 1 expected errors were selected (quality filtering). Primer matching was performed with CUTADAPT v.1.9.1. Dereplication, OTU clustering at 97% (UPARSE algorithm) and zOTU denoting at 100% similarity (UNOISE algorithm) were performed with USEARCH v.10.0.240 (Edgar, 2010). OTUs and zOTUs were taxonomically annotated against the SILVA database v.132 (2018) with the LCA approach. Finally, OTUs and zOTUs were quantified to obtain OTU-abundance tables. Non-prokaryotic OTUs (eukaryotes, chloroplast and mitochondria) were removed, whereas singletons (OTUs appearing only once) were maintained. Computing analyses were run at the MARBITS bioinformatics platform at the *Institut de Ciències del Mar* and at the Euler scientific compute cluster of the ETH Zürich University.

This procedure was applied individually for the: (i) 41 Malaspina Bathypelagic samples, (ii) 124 Malaspina Surface samples, (iii) 119 *Tara* Oceans and *Tara* Polar Oceans Surface and Mesopelagic samples, and (iv) 155 samples from the Malaspina size fractions dataset. Hence, four different OTU-abundance tables were obtained after applying the pipeline. To allow comparisons between samples, each OTU table was randomly sampled down to lowest sampling effort using the function *rrarefy.perm* with 1000 permutations from the R package *EcolUtils* (Salazar, 2018a). A summary of the total number of reads per dataset, sample with the lowest number of reads and total OTUs or zOTUs before and after rarification is described in Table S15.

16S rRNA metagenomic sequences

Prokaryotic communities were also analyzed using miTAGs, i.e. 16S gene's fragments extracted from Illumina-derived metagenomes (Logares et al., 2014). The HQ reads corresponding to 16S genes were detected using HMMER[7] v.3.0. All reads detected as part of an rRNA gene with length >100 bp were

[5] https://rtlgenomics.com
[6] https://github.com/SushiLab/Amplicon_Recipes
[7] www.hmmer.org

aligned against the SILVA database (v.132) (Pruesse et al., 2012) using USEARCH v10.0.240 (Edgar, 2010). The alignment was done using a 99% similarity cut-off and thus the SILVA database was pre-clustered, also at a 99% similarity, before the alignment. All the ambiguous hits (i.e. reads with a successful hit to more than one sequence of the reference database) were excluded. In this way the miTAGs were binned into mOTUs (metagenomic OTUs): each sequence of the pre-clustered SILVA database that recruited a read was considered as a mOTU and their abundances were obtained as the number of reads that aligned to this sequence. A mOTU table was constructed containing the number of reads that belonged to each mOTU in each of the samples. All these processes were developed using the R package *mtagger*[8].Once the final mOTU table was available, it was separated into two corresponding to Malaspina Surface and Malaspina Bathypelagic. In order to avoid artifacts due to the uneven sequencing effort among samples, these mOTU tables were rarefied to the minimum number of total miTAGs per sample (Table S16) using the *rrarefy* function in the *vegan* package(Oksanen et al., 2018) within R Statistical Software (core team, 2017). This process was repeated 100 times and the mean number of reads (rounded to integers) from the 100 rarefactions was used. The taxonomy from the SILVA sequences was used as the taxonomical classification of the mOTUs.

Comparison between OTUs, zOTUs or mOTUs and cultured isolates

Primers used to obtain the 16S rRNA genes of the isolates were different from the ones used to obtain the 16S rRNA iTAGs, although both amplified the V4 and V5 hypervariable region of the 16S rRNA gene, and therefore, comparisons between both datasets could be performed selecting this common region (Figure S1). For mOTUs, comparisons were also possible as the complete 16S rRNA sequences were extracted from SILVA database (v.132) and the V4 band V5 region was covered.

All isolates sequences were compared to zOTUs and to mOTUs sequences at 100% similarity, but isolates were also compared with OTUs and mOTUs at 97% similarity for comparative purposes with other studies by running global alignment using the *usearch_global* option from the USEARCH v10.0.240 (Edgar, 2010). The results were filtered by coverage of the alignment at 100% and at their maximum percentage of identity for those comparisons at 99-97% similarity. We are aware that these comparisons sometimes resulted in more than one hit per isolate, even after filtering by coverage. Nevertheless, all datasets presented similar proportions of isolates with more than one match and the % of reads relucted by those extra hits was minor, making comparisons between datasets possible (Figure S9).

Data analysis

All data treatment and analysis were conducted with the R Statistical Software (core team, 2017) using the version v.3.4.3 and the following packages: *vegan* (Oksanen et al., 2018), *ape* (Paradis et al., 2004), *EcolUtils* (Salazar, 2018a), *stats* (core team, 2017), *tydiverse* (Wickham, 2019). For each studied dataset we calculated the mean abundance and relative abundance of OTUs, zOTUs and mOTUs across samples in order to rank them. Moreover, we calculated the mean percentage of reads (iTAGs or miTAGs), OTUs, zOTUs or mOTUs of the bacterial community that matched at 97% or at 100% similarity with the 16S rRNA sequences of the strains isolated by traditional culture techniques. These percentages

[8]https://github.com/GuillemSalazar/mtagger

were calculated from the rarefied OTU-abundance tables. In order to see the differences between the proportion of isolated reads of each dataset, the non-parametric Kruskal-Wallis test was applied followed by the post hoc pairwise Wilcox test to see the differences between pairs of datasets. To asses significance, the statistical analyses were set to a conservative alpha value of 0.05.

Furthermore, with the Malaspina size fractions iTAGs dataset we wanted to elucidate if our isolates were more prone to be found in the bacterial communities found in the free-living fraction (0.2-0.8 µm) or in particle-attached fraction, considering that this last one can be divided into four different size-fractions (0.8-3 µm, 3-5 µm, 5-20 µm and 20-200 µm). Hence, we also calculated the percentage of reads or iTAGs, OTUs and zOTUs that matched at 100% similarity with our isolates. The differences between size fractions were also detected with the non-parametric Kruskal-Wallis test followed by the post hoc pairwise Wilcox test to see the differences between pairs of datasets. Again, the statistical analyses were set to a conservative alpha value of 0.05 to asses significance.

Nucleotide accession numbers

The 16S rRNA gene sequences from the MARINHET isolates culture collection were deposited in GenBank under accession numbers MH731309 - MH732621 and MK658870-MK659428.

Amplicon 16S rRNA iTAGs from the Malaspina Surface dataset are available in Ruiz González et al. (2019), from the Malaspina Bathypelagic dataset in Salazar et al. (2016), and from Malaspina size fractions in Mestre et al. (2018). Amplicon 16S rRNA iTAGs from the *Tara* Oceans and the *Tara* Oceans Polar Circle are available in the MARBITS bioinformatics platform at the *Institut de Ciències del Mar*.

Metagenomic 16S rRNA from Malaspina Surface and Malaspina Bathypelagic datasets are also available in the MARBITS bioinformatics platform at the *Institut de Ciències del Mar*.

Results

Abundance comparison between cultured heterotrophic bacteria and flow cytometry counts from seawater samples from different oceanic regions and depths

We calculated the percentage of isolated cells in 10 photic-layer, 3 mesopelagic and 7 bathypelagic stations where plate colony count (cfu/ml) and flow cytometry values -as a measure of total cell counts- (cells/ml) were available (Figure 12). For this comparative analyses, flow cytometry counts only included the abundance of all heterotrophic bacterial cells excluding the photosynthetic *Cyanobacteria* as Synechococcus and *Prochlorococcus* since those taxa were not targeted with the media nor incubations conditions selected. The percentage of recovery of heterotrophic bacterial cells ranged from 0.01 to 1.3% in photic-layers samples, and from 0.9 to 2.5% in the mesopelagic, while percentages for bathypelagic samples were between 0.08% and 3.5%, indicating a higher fraction of culturability in some samples in the deeper layers of the ocean (Figure 12). Additionally, if we look at the mean percentage of isolated cells per layer, we detected a higher percentage in the mesopelagic and bathypelagic samples (1.3%) in comparison with the photic layer (0.3%). However, significative differences were not found (P-value

> 0.05) between photic, mesopelagic and bathypelagic samples.

Comparative analyses between 16S rRNA sequences from isolates and 16S amplicon iTAGs from photic and aphotic *Tara* Oceans and Malaspina Expedition datasets

A total of 2003 heterotrophic isolates from our marine culture collection (MARINHET_v2) were compared with two global amplicon 16S iTAGs datasets with a total of 38,700 zOTUs from *Tara* Oceans (15,426 *Tara* Surface and 23,274 *Tara* Mesopelagic) and 15,348 zOTUs from Malaspina Expedition (3,528 Malaspina Surface and 11,820 Malaspina Bathypelagic) (Figure 11). We determined the mean percentage of zOTUs, as well as the mean percentage of reads that were 100% identical in their 16S rRNA sequences to our isolates MARINHET_v2 culture collection (hereafter referred as isolated zOTUs) in the *Tara* Surface, *Tara* Mesopelagic, Malaspina Surface and Malaspina Bathypelagic datasets. These comparisons were done at three different levels: (i) comparing each 16S iTAGs dataset from *Tara* Oceans and Malaspina Expedition with all the isolates sequences, (ii) comparing separately the photic and aphotic *Tara* Oceans and Malaspina Expedition 16S iTAGs datasets with only photic or aphotic isolates sequences, and (iii) comparing the 16S iTAGs and the partial 16S rRNA gene isolates sequences retrieved from the same seawater stations. In Table S17 we present a summary of the results from the three levels of comparisons obtained from rarefied zOTU-abundance tables. Results from the comparisons between OTUs (97%) and isolates at 97% sequence similarity are also included in Table S17, but as similar patterns were obtained for 100% comparisons, here we will only describe those from the zOTUs compared at 100% similarity with isolates, as is the strictest possible comparison.

The highest mean number of zOTUs that were 100% identical to the partial 16S rRNA gene of the heterotrophic isolates was found in the Malaspina Surface dataset (4.5%), followed by the Malaspina Bathypelagic (2.4%), *Tara* Surface (2.3%) and *Tara* Mesopelagic datasets (1.7%) (Figure 13A). Even though those percentages of isolated zOTUs do not seem to vary greatly between datasets, significative differences were found among them (Kruskal-Wallis, P-value < 0.01, Figure 13A). Similarly, when the comparisons of the datasets were done separately for photic and aphotic isolates, the highest mean abundance of isolated zOTUs was found in Malaspina Surface (2.7% photic, 4.1% aphotic) and the lowest percentages in *Tara* Mesopelagic (0.9% photic, 1.1% aphotic) (Table S17). Although the fraction of isolated zOTUs were rather small and ranged between 0.1% to 4.5% of the total zOTUs (Figure 13A and Table S17), the percentage of reads within these isolated zOTUs increased significantly in the deep ocean (Figure 13B). Thus, around 1.6-4.9% of the sequences were 100% identical to our isolates in the photic ocean, this value increases up to 8.5% in the mesopelagic ocean, and increases even more, up to 27.9%, in the bathypelagic ocean. In this case, the differences between datasets were also statistically significant (Kruskal-Wallis test, P-value < 0.01, Figure 13B). Differences were also found when comparing all the isolates together or comparing separately with photic and aphotic isolates (Figure S10 and Table S17).

On the other hand, there is a fraction of isolates that did not match with any zOTU regardless the dataset (Table S18). Indeed, about 11% of our heterotrophic isolates did not match any of the *Tara* Oceans zOTUs, whereas it increased up to 18% in the Malaspina Bathypelagic zOTUs and up to 28% in the surface samples from Malaspina (Figure S9A). The taxonomic classification at the family level of all the isolates revealed some interesting differences between those that were identical to zOTUs and those that did not match any zOTU (Figure S11). We found some families that were identified in

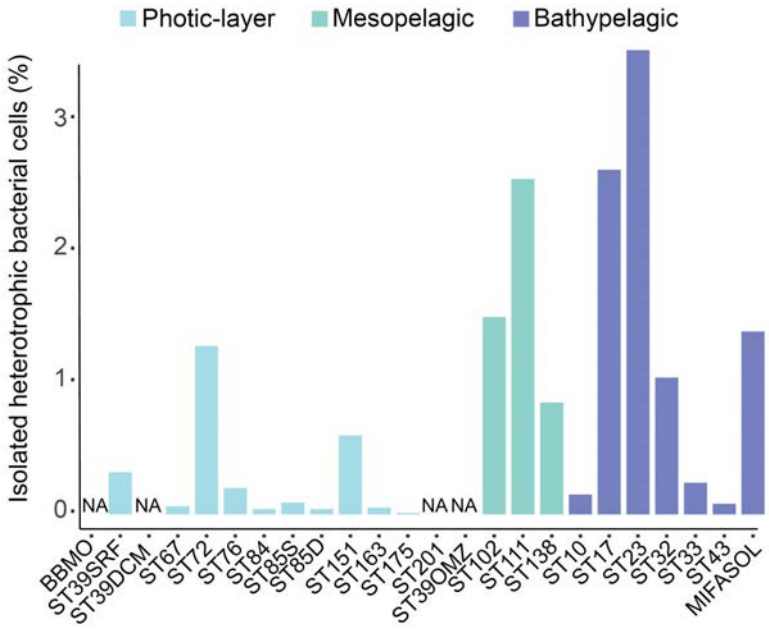

Figure 12: Testing the great plate count anomaly. For each station, heterotrophic bacterial cell counts were estimated using both traditional culture techniques (cfu/ml) and flow cytometry- as a measure of total cell counts- (cells/ml). The percentages represent the pool of heterotrophic bacterial cells that could be retrieved by culturing in some photic, mesopelagic stations that belonged to oxygen minimum zones (OMZ) regions, and bathypelagic stations. Color indicates the depth of the sample: cyan, photic-layer; bluegreen, mesopelagic; and dark-blue, bathypelagic. No significative differences were found between layers (P-value > 0.05).

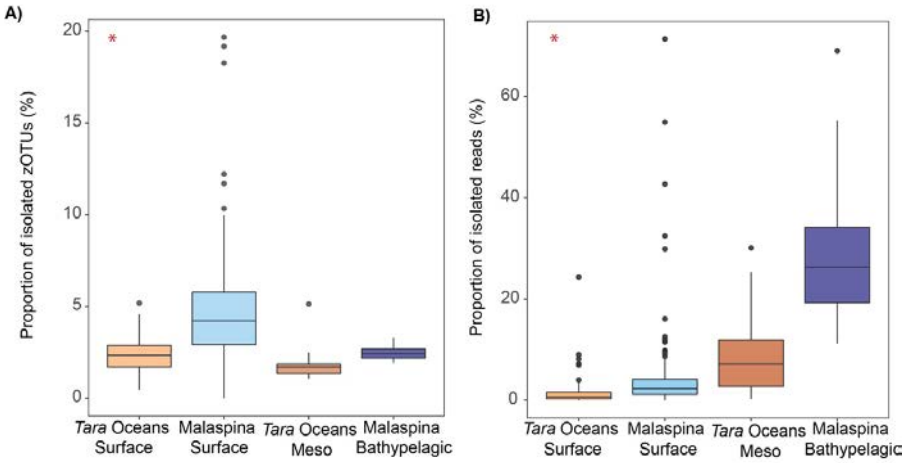

Figure 13: (A) Proportion of isolated zOTUs (zOTUs matching isolates at 100 % similarity). **(B)** Proportion of isolated reads (16S iTAGs reads matching isolates at 100 % similarity). Values are extracted from the mean abundance of reads or zOTUs in each dataset from rarified zOTU-abundance tables. Outliers are indicated with grey circles. If significant differences are found between all datasets it is indicated inside boxplots with an asterisk (Kruskal-Wallis , P-value < 2.2 e-16).

Tara Oceans Surface and Mesopelagic datasets but not in the Malaspina Expedition Surface or Bathypelagic datasets, such as *Tistrellaceae*, *Nitrincolaceae* or *Colwelliaceae*. Contrarily, *Kangiellaceae* was found in both Malaspina Expedition datasets but not in *Tara* Oceans. Interestingly, we found some families that included isolates that did not match with zOTUs in any of the 16S iTAGs amplicons datasets, such as *Dermabacteraceae*, *Balneolaceae* or *Psychromonadaceae* (Figure S11). These results indicated that despite the high sequencing effort applied for amplicon 16S iTAGs to marine *Tara* Oceans and Malaspina Expedition samples, there was still a proportion of the isolates, which may belong to the rare biosphere, that could not be detected by this sequencing technique.

Heterotrophic isolates matched with some of the most abundant zOTUs of the deep ocean

To test weather isolated zOTUs belonged to the abundant or rare biosphere, rank abundance plots were performed for each of the studied *Tara* Oceans and Malaspina Expedition datasets with the zOTUs mean abundances extracted from rarefied zOTU-abundance tables (Figure 14). The rank abundances were plotted either based on the global comparisons of all isolates (Figure 14) or by comparing photic and aphotic isolates separately against photic 16S iTAGs datasets, or photic and aphotic isolates separately versus mesopelagic and bathypelagic 16S iTAGs datasets (Figure S12) giving similar results for all of them. Each rank abundance plot showed a similar pattern of few abundant zOTUs (relative abundance > 1%) and a long tail of rare or low-abundant zOTUs (relative abundance <0.01%). We coloured those zOTUs that were 100% similar with at least one isolate (referred here as isolated zOTUs) to see differences between depths (Figure 14). In photic layers, we did not detect any isolated zOTUs within the abundant taxa of the *Tara* Surface dataset (Figure 14A), whereas only one isolated zOTU belonging to the abundant biosphere, taxonomically related to *Sulfitobacter* and presenting 1.65% of the reads, was found in the Malaspina Surface dataset (Figure 14B). The rest of the isolated zOTUs in these two large-scale photic datasets reside in the mid-abundant biosphere (38 isolated zOTUs in Malaspina and three isolated zOTUs in *Tara*) or in the rare biosphere (54 in Malaspina and 151 in *Tara*). In the mesopelagic layer (*Tara* Mesopelagic) we only found one isolated zOTU classified into the abundant biosphere associated to *Alteromonas* with a total of 1.1% reads, although an increased number of isolated zOTUs with medium abundances (87 zOTUs) emerged (Figure 14C). Interestingly, in the bathypelagic ocean, the most abundant taxon by 16S iTAGs matched at 100% identity with our isolates. This was related to *Sulfitobacter* and represented a total of 4.64% of the iTAGs reads. In total, seven isolated zOTUs belonged to the abundant biosphere in Malaspina Bathypelagic, affiliating to the genera *Sulfitobacter*, *Halomonas*, *Erythrobacter*, *Alteromonas* and *Sphingobium*. These abundant isolated zOTUs together with a big proportion of isolated zOTUs classified among the medium biosphere (78 zOTUs), and the ones belonging to the rare biosphere (52), recruited a total of 28% of the environmental 16S iTAGs from the temperate and tropical global deep oceans (Figure 14D). Thus, it seems that a higher proportion of abundant zOTUs could be retrieved by culture-dependent techniques when depth increased, especially in the bathypelagic layer.

Taking together all the datasets, we found that the zOTUs that were 100% identical to our isolates affiliated mostly to classes *Alphaproteobacteria* (33.5% zOTUs) and *Gammaproteobacteria* (40.8% zOTUs) followed by phyla *Bacteroidetes* (18.2%), *Actinobacteria* (5.8%) and very few to *Firmicutes* (1.5%) (Figure 14 and Table S19). We noticed that despite finding relatively similar proportions of isolated zOTUs belonging to *Alphaproteobacteria* and *Gammaproteobacteria* in all ocean layers, the proportion of reads within

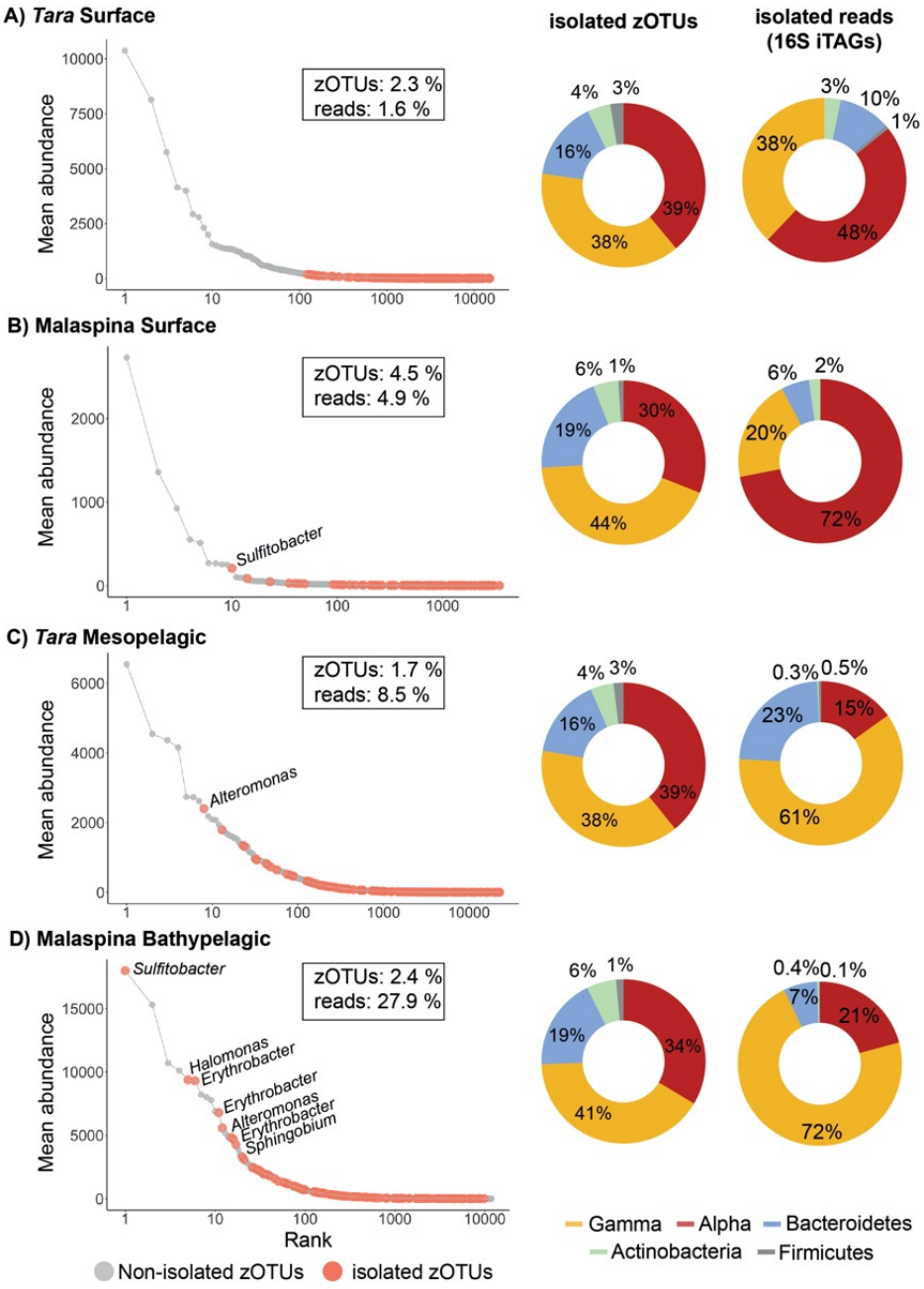

Figure 14: Rank plots showing the identified isolated zOTUs in different datasets. **(A)** *Tara* Surface **(B)** Malaspina Surface. **(C)** *Tara* Mesopelagic **(D)** Malaspina Bathypelagic. Colour of the dots indicates the isolated zOTUs (zOTUs 100 % identical to at least one isolate) in the comparisons made with all photic and aphotic isolates (global): grey, non-isolated zOTUs; orange, isolated zOTUs. Taxonomic affiliation is indicated for the abundant zOTUs identical to isolates (isolated zOTUs, >1 % abundance) in each rank plot. Donut-like charts describe the taxonomic affiliation at the phylum/class level of the isolated zOTUs (left donut-like charts) or the recruited 16S iTAGs (right donut-like charts): yellow, *Gammaproteobacteria*; red, *Alphaproteobacteria*; blue, *Bacteroidetes*; green, *Actinobacteria*; and grey, *Firmicutes*. Percentages are indicated inside donut-like charts.

these isolated zOTUs of these classes differed between photic and aphotic layers; affiliating mainly to *Alphaproteobacteria* (48-72%) in the photic ocean, and dominating the *Gammaproteobacteria* (61-72% of the reads) in the deeper layers (Figure 14).

Increase of isolated zOTUs with depth and in bigger plankton size fractions

In this study, we also compared our isolates with samples from five different plankton sizes fractions (related to free-living bacterial communities or those attached to particles of different sizes ranging from 0.8 to 200 µm) from eight detailed vertical profiles from the Malaspina Expedition (Mestre et al., 2018), which included surface, deep chlorophyll maximum (DCM), mesopelagic and bathypelagic layers. We wanted to elucidate if the heterotrophic isolates recovered preferred a specific lifestyle (free-living vs. particle-attached) and find out possible differences between ocean layers. In this manner, the 16S rRNA sequences of the MARINHET_v2 were compared to a total of 6,274 zOTUs identified in those fractionated vertical profiles. Our analysis revealed that our isolates were present across all size fractions and depths although proportions of isolated zOTUs varied between them (Figure 15 and Figure S13). First, when focusing in the differences between layers (surface, DCM, mesopelagic and bathypelagic), we confirmed again higher mean abundances of isolated zOTUs in the bathypelagic. The 16.3% of isolated zOTUs from the bathypelagic (Figure 15A) recruited an average of 40% of reads (Figure 15B). Among surface, DCM and mesopelagic samples, a similar number of isolated zOTUs was detected (approximately 8% average) and also similar proportions of isolated reads were identified (approx. 20% average) (Figure 15). Interestingly, the DCM was the layer with the smaller number of isolated zOTUs (approx. 6% average) and reads recruited (approx. 17.3% average). Given the notable proportion of isolated zOTUs in the bathypelagic samples, statistically significant differences were found between this deeper layer and the other depths in the different size fractions, at least in the proportion of isolated zOTUs (Kruskal-Wallis test, P-values <0.01), but not among the surface, DCM or mesopelagic samples (Figure S13 and Figure S14).

On the other hand, the proportion of isolated zOTUs across different plankton size fractions, revealed that our isolates were prevalent in the larger plankton size fraction associated to particles (≥3.0 µm) (Figure 15). In the photic samples (surface and DCM) higher mean abundances of isolated zOTUs and reads (10% and 26% on average, respectively) were recovered in those larger size fractions (3 µm, 5 µm and 20 µm pore filters), but not in the free-living bacterial communities (0.2 µm, 2.9% average zOTUs, and 1.5% average reads) or in the smallest particles (0.8 µm, 6% average zOTUs, and 12% average reads) (Table S20 and Table S21). In the mesopelagic samples the number of isolated zOTUs (7%) in the 0.8 µm size fraction recruited up to 23% reads, but not in the free-living bacteria size fraction (0.2 µm), that still presented lower values of isolates zOTUs and reads recruited (4% of reads). Finally, in the bathypelagic samples all plankton size fractions presented similar percentages of isolated zOTUs, including the free-living bacteria (average 16%), and also uniform proportions of reads (40% on average). Besides, the highest values were found in the largest particles (>20 µm, 21% of zOTUs and 45% of reads). In surface, DCM and mesopelagic samples, significant differences were detected between the isolated zOTUs of free-living bacteria and the communities attached to particles (Kruskal-Wallis test, P-value < 0.05), but not among the isolated zOTUs attached to particles of different sizes (Kruskal-Wallis test, P-value > 0.05) (Figure 15 and Figure S14). In the bathypelagic, only significative differences were found in the proportion of isolated zOTUs between the larger size particles and the other fractions (Figure

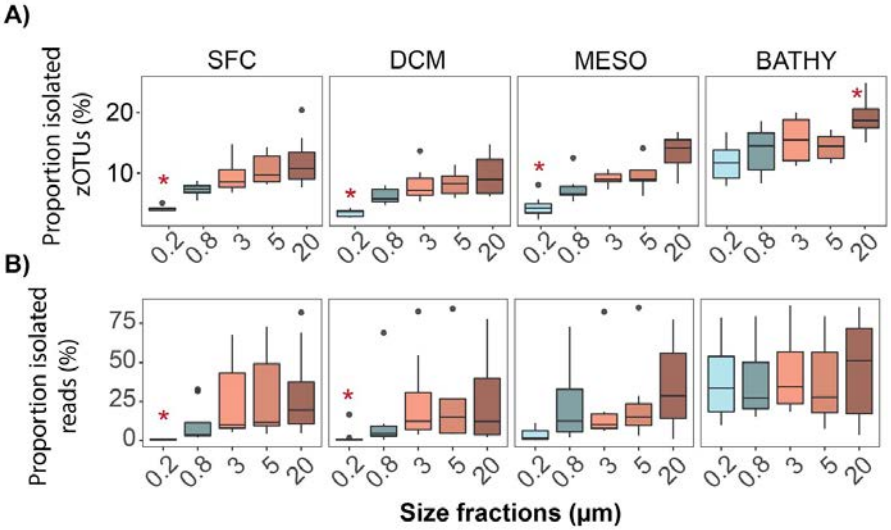

Figure 15: Boxplots from Malaspina Expedition profiles with different size fraction showing: **(A)** the percentages of isolated zOTUs per size fraction and depth, and **(B)** the percentage of reads (16S iTAGs) that were 100 % identical to at least one isolate. SFC: surface, DCM: deep chlorophyll maximum, MESO: mesopelagic, and BATHY: bathypelagic. 0.2 µm: free-living bacteria, 0.8 µm: bacteria attached to small particles, and 3-20 µm: bacteria attached to big particles. Significative differences between size fractions in each layer are indicated by red asterisks (Kruskal-Wallis, P.value <0.05).

15).

Do the metagenomics miTAGs and amplicons iTAGs datasets show the same pattern?

A total of 36,215 16S metagenomic fragments (miTAGs, (Logares et al., 2014)) were extracted from 10 surface metagenomes (2,045 miTAGs OTUs (hereafter mOTUs)) and 11 bathypelagic metagenomes (1,957 mOTUs) from same vertical profiles from Malaspina Expedition. We first compared the zOTUs from the previous described 16S iTAGs datasets to the mOTUs in order to elucidate which proportion of the prokaryotic community detected by amplicon TAGs are represented (at 100% identity) among the metagenomic datasets (Figure 16 and Figure S15). We observed an average of 36.2% mOTUs and 45.9% of their reads 100% similar to zOTUs, values that increased up to 69.4% and 78.2% of the mOTUs and reads, respectively, when comparisons were done at 99% similarity (Figure S15 and Table S22). We also revealed that zOTUs identical to mOTUs were present along all rank abundance plots covering abundant and rare members of the biosphere (Figure S15). Conversely, those zOTUs that did not match with any mOTU, represent 68% and 93% of the Malaspina Surface and Bathypelagic zOTUs, and 18.1% and 58.6% of the total 16S amplicon reads, respectively (Table S22). Furthermore, taxonomic differences were found between prokaryotic communities detected by mOTUs and zOTUs regardless the proportion of reads or OTUs shared. Taking together all the data from zOTUs and mOTUs, we identified 39 prokaryotic classes in Malaspina Surface and 88 in Malaspina Bathypelagic (Figure 16). Some classes were found in both mOTUs and zOTUs, being 16 (41%) present in Malaspina Surface and 29 (33%) in Malaspina Bathypelagic. On the other hand, we found that those mOTUs that did not match with any zOTU belong mainly to the same classes, and only 6 (Malaspina Surface) and 3 (Malaspina Bathypelagic) were only detected among the mOTUs data (Figure 16). In Malaspina Surface these unique mOTUs included for example *Leptospirae* or *Parcubacteria*, whereas in the bathypelagic we observed *Saccharimonadia*, *Halanaerobiia* and AT-s3-28. As we mentioned, we also noticed zOTUs that

were not identified among the mOTUs (Figure 16). These classified into classes that were recognized among the mOTUs, but also other that were unique for the amplicon data, 12 (30.7%) in Malaspina Surface such as SAR 406 clade, *Mollicutes*, *Nitrospinia* or *Bacilli*, and 23 (23.1%) in Malaspina Bathypelagic including *Deinococci*, *Ignavibacteria* or *Fusobacteriia* (Figure 16).

Besides, we found a moderate correlation between species richness (Chao-1) of the mOTUs and zOTUs of the Malaspina Surface dataset (Figure S16). Taken together these differences between mOTUs and zOTUs, it can be pointed out that less than 50% of OTUs or reads overlapped between both approaches and therefore, comparison should be taken with caution. Nevertheless, we compared the mOTUs of Malaspina Expedition with our isolates sequences at 100% (most strict comparison) and at 97% similarity (usual similarity threshold for species delineation, (Yarza et al., 2014)) in order to detect which mOTUs were identical to isolates (hereafter isolated mOTUs). Some patterns were coherent with the comparisons done against zOTUs (Figure 17 and Figure S17). As in the amplicon 16S iTAGs datasets, higher mean proportions were found in the bathypelagic dataset, with 3.4% of isolated mOTUs and 6.1% miTAGs reads (Figure S17) mainly when comparisons are done at 97% sequence similarity. Significative differences were found between photic and aphotic isolated mOTUs (Kruskal-Wallis test, P-value < 0.05). Nevertheless, as observed, the proportion of isolated mOTUs was greatly reduced compared to the amplicons datasets (isolated zOTUs). These reduction is more noticeable in the comparisons at 100% similarity where isolated mOTUs only represent 0.27% and 1% of the photic and bathypelagic datasets, respectively, including 0.13% and 0.74% of the total reads (Figure 17 and Figure S17). Notably, these proportions of isolated mOTUs and reads derived from an average 15% of the isolates, as 90% and 79% of the isolates did not match with any mOTU (at 100% similarity) in Malaspina Surface and Bathypelagic, respectively (Figure S9). These low proportion of isolates contrasts with what we obtained in the amplicon datasets where, as we previously mentioned, around 82% of the isolates match with at least one zOTU. The rank abundance plots of the mOTUs detected for surface and bathypelagic samples (Figure 17) show that only in the comparisons at 97% similarity we are able to identify one isolated mOTU belonging to the abundant biosphere affiliating to *Nocardioides* genus and representing 1.38% of the total reads in the bathypelagic. The rest of the isolated mOTUs belong to the mid-abundant or rare biosphere. Curiously, it is in the bathypelagic samples where a higher proportion of isolated mOTUs are classified as mid-abundant for both thresholds of comparison (Table S23). These patterns resemble those obtained also for the amplicon 16S iTAGs even though the number of miTAGs reads, and consequently, the proportion of isolated mOTUs from the total prokaryotic community had been considerably diminished.

Discussion

The precise meaning of the traditional paradigm that only 1% of microbes are culturable is difficult to interpret as recently stated by Martiny (2019). Our comparative analyses between the number of colonies retrieved in pure culture and the cell abundances calculated with culture-independent techniques such as flow cytometry just gives us information about the degree of culturability of the samples. Our results for mesopelagic and bathypelagic samples reflected a certain degree of variability regardless of depth, being the average percentage of culturability between 1.5 to 3.5%. This variability has been also highlighted in recent metadata analysis studies where they examined culture-dependent surveys covering a wide variety of environments, from lakes, seawater, soils or sediments, and also human hosts asso-

Figure 16: Taxonomic differences at the class level between mOTUs matched at 100 % with zOTUs, mOTUs that did not match any zOTUs and zOTUs that did not match with any mOTU. **(A)** Malaspina Surface. **(B)** Malaspina Bathypelagic. Size of the dots indicate the percentage (%) of mOTUs or zOTUs affiliating to each class from the total of mOTUs or zOTUs, respectively, identified in each dataset. Name classes colored in blue belong to the Archaea domain, in black to the Bacteria domain.

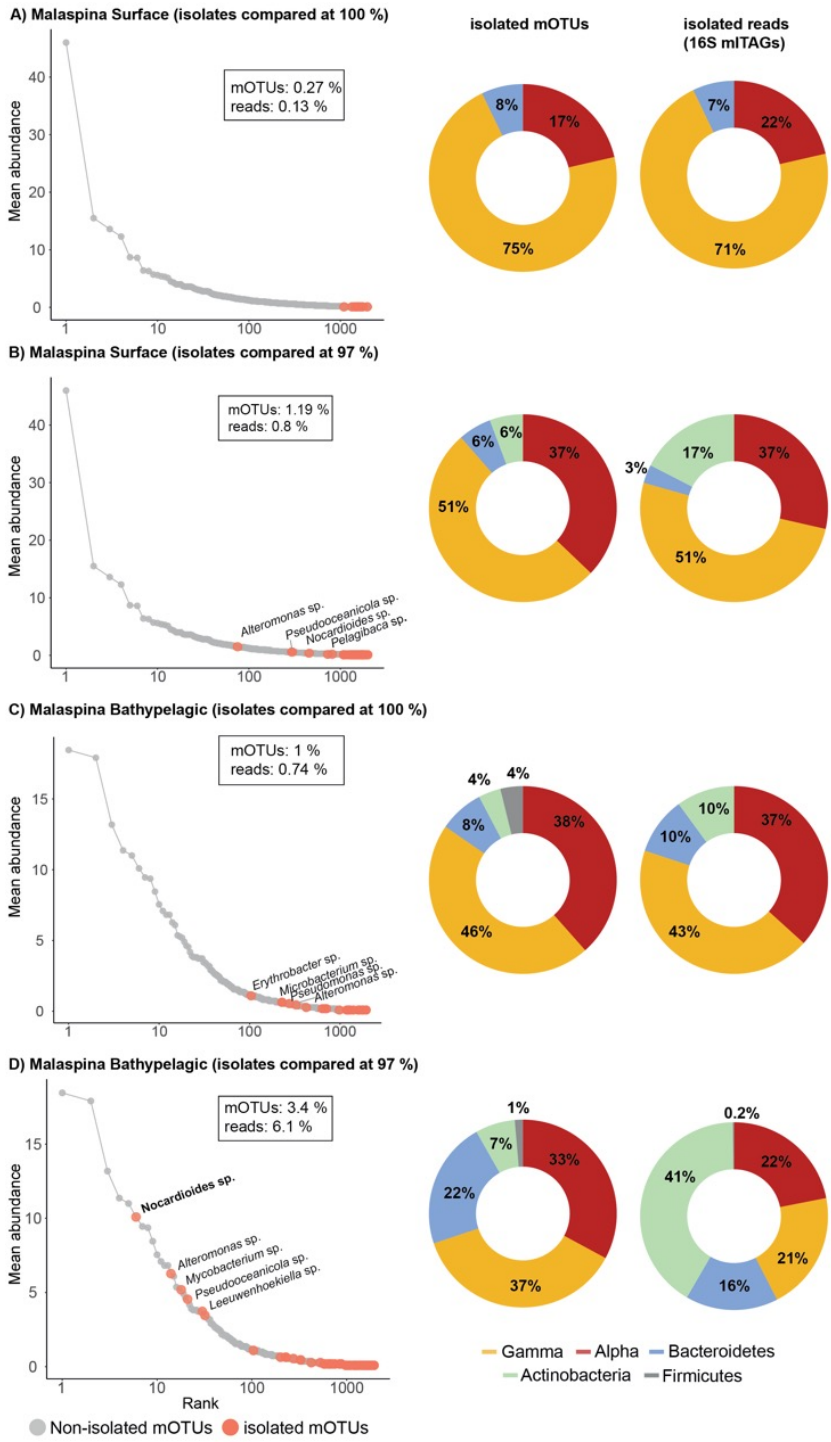

Figure 17: Rank plots showing the identified isolated mOTUs in surface and bathypelagic Malaspina Expedition datasets. **(A,B)** Malaspina Surface dataset comparing mOTUs and isolates at 100% and 97% sequence similarity. **(C,D)** Malaspina Bathypelagic dataset comparing mOTUs and isolates at 100% and 97% sequence similarity. Color of the dots indicates those mOTUs that were recruited by isolates: grey, non-isolated mOTUs; orange, isolated mOTUs. Taxonomic affiliation of the abundant isolated mOTUs (>1% reads, bold) and the top 4 mid-abundant mOTUs (1-0.01% reads, italics) is indicated in each plot. Donut-like charts describe the taxonomic affiliation at the phylum/class level of the isolated mOTUs (left donut-like charts) or the recruited 16S miTAGs (right donut-like charts): yellow, *Gammaproteobacteria*; red, *Alphaproteobacteria*; blue, *Bacteroidetes*; green, *Actinobacteria*; and grey, *Firmicutes*. Percentages are indicated inside donut-like charts.

ciated communities (Lloyd et al., 2018; Steen et al., 2019). Despite they concluded that the median of culturable cells was around 0.5%, they remarked this mentioned variability in different marine environments including deep-sea brine pools, artic sea-ice and their surrounding waters, anoxic waters of the Baltic sea or deep-sea sediments (Lloyd et al., 2018). However, none of the studies reviewed included open water samples deeper than 200 m depth. Thus, we can confirm that this variance can also be found across the mesopelagic and the bathypelagic ocean samples, concluding that, for some samples, more than 1% of cells and up to 3% of the cells (Figure 12) could be cultivated in the deep ocean and therefore the traditional paradigm should not be literally interpreted.

However, further comparisons with HTS data, such as amplicon 16S iTAGs or metagenomic iTAGs, is needed in order to decipher which fraction of the community is being cultivated and how abundant are the heterotrophic isolated bacteria in natural communities. The rank abundance curves of the zOTUs and mOTUs are represented by some high-abundant and moderately abundant taxa but many of them are present in very low abundances, what is so called the rare biosphere (Sogin et al., 2006; Pedrós-Alió, 2012). We confirmed this structure for both *Tara* Oceans and Malaspina Expedition marine datasets covering surface, mesopelagic and bathypelagic layers, in agreement with previously reported studies (Sunagawa et al., 2015; Crespo et al., 2016; Salazar et al., 2016). The common statement that culture-dependent studies mainly capture members of the rare biosphere seems to be the rule for free-living bacteria in the photic layer as an average 81% of the isolated zOTUs in *Tara* and Malaspina photic datasets recruited less than 1% of the total 16S rRNA sequences at 100% identity (Figure 13), confirming that the isolated zOTUs in those datasets were mostly members of the rare-biosphere. However, we found that 2.4% of the isolated zOTUs at the bathypelagic ocean recruited a significant fraction of the total microbial community in the deep realm identified by amplicon 16S iTAGs, up to 28% of the reads (Figure 13B). One could argue that these differences between photic and aphotic samples are due to the isolation strategy used in this study. We focused only on the heterotrophic marine bacteria and we are aware that the photosynthetic bacterial community in the surface samples is not captured with our isolation strategy. Therefore, when removing from the abundance tables those zOTU affiliated with *Cyanobacteria*, we obtained similar patterns with a slight increase, being on average up to 8%, 10% and 29% of the amplicon 16S iTAGs reads identical to isolates in surface, mesopelagic and bathypelagic datasets, respectively (Table S24).

Also in this study, the results from amplicon 16S iTAGs were contrasted to those comparisons between isolates and mOTUs defined from metagenomic data, which may usually represent a less biased view of the natural prokaryotic communities (Steen et al., 2019). A good correlation between primer-amplified 16S rRNA gene sequences and metagenomic data has been previously reported for *Tara* Oceans (Ibarbalz et al., 2019) and Malaspina Expedition datasets (Salazar, 2018b). Indeed, similar patterns were observed in this study between 16S iTAGs and miTAGs as both pointed out the differences of isolated zOTUs/mOTUs and reads recruited between surface and bathypelagic samples, and the higher proportions of isolated zOTUs or mOTUs in the bathypelagic samples. These similarities between HTS data were deciphered even though a weak correlation was identified in our study in Chao-1 diversity index estimates for the Malaspina Surface dataset ($r=0.6$, P-value $=0.07$). Nonetheless, we clearly see a reduction in the proportion of reads identical to isolates in the metagenomic dataset, which decreased from the mean 28% observed by amplicon data to 6.1% or 1% of the miTAGs at 97% or at 100% similarity, respectively. These differences between datasets could be explained by different reasons. First, it is known that the primers chosen to amplify the 16S rRNA sequences can give different results de-

pending on the targeted region of the gene (Parada et al., 2016; Sambo et al., 2018). Second, it can exist a mismatch between the sequence of the primers and the 16S rRNA gene of certain organisms. Certainly, in the Malaspina Surface and Bathypelagic datasets we found some taxonomic classes that were only detected by mOTUs, but at the same time, we could also recognize some classes only found among the zOTUs. Third, some organisms may be present in such low abundances in the natural environment that primers do not have the chance to match to their DNA and amplify their 16S rRNA gene. Finally, another reason explaining such differences is because PCR demanding sequencing techniques are skewed towards cultured organisms (Lloyd et al., 2018; Steen et al., 2019) and this has increased the percentage of isolated zOTUs and reads found compared to metagenomic data. Also, our mOTUs were defined considering those miTAGs with unambiguous hits against reference SILVA database (v.132) and therefore all the ambiguous hits (i.e. reads with a successful hit to more than one sequence of the reference database) were excluded, which may reduce the % of final reads used in the mapping against our culture collection.

Our results show that most members of the bacterial communities studied, especially for the free-living bacteria, did not match any of our isolates regardless of the depth, thus remaining part of the uncultured diversity. These degree of uncultured taxa have been similarly described previously for marine ecosystems (Selje et al., 2005; Crespo et al., 2016) and other environments (Lloyd et al., 2018) concluding that most part of the microbial diversity on Earth remains uncultured. However, around 37% of Bacteria and 34% of Archaea (Lloyd et al., 2018) are identical to cultured strains when compared to primer-amplified 16S rRNA gene at 97% similarity, which is a conservative genetic threshold to define species (Yarza et al., 2014). Nevertheless, to the best of our knowledge, comparisons between isolates and HTS data retrieved from the same samples was not previously tested at a large scale for deeper layers, including the mesopelagic and the bathypelagic ocean.

On the other hand, an interesting result that emerged in this study from amplicon data was that the recruitments of isolated zOTUs were higher in the largest size fractions (>20 µm) of the Malaspina Bathypelagic with about 45% of the total reads recruited (Figure 15B). This finding indicated that most of the heterotrophic isolates from our MARINHET_v2 collection have a preference to live attached to particles. It is already well know that marine microbial communities attached to particles are very different from the free-living ones (Acinas et al., 1999; Eloe et al., 2011b; Salazar et al., 2015; Mestre et al., 2018). A previous study of Mestre and collaborators (Mestre et al., 2018), who analyzed the same Malaspina plankton size fraction dataset included here, determined the importance of sinking particles to promote vertical connectivity in the ocean microbiome. Here, we noticed similar proportions of recruited reads by isolates in all size fractions when comparing zOTUs and isolates in the bathypelagic samples, suggesting that our heterotrophic isolates are well represented in the bathypelagic realm in all plankton size fractions. These similarities between size fractions could be due to the presence of bathypelagic isolated bacteria with dual lifestyles, those living in large particles and those who are part of the free-living bacterial community. Furthermore, the isolated zOTUs that were abundant (>1% relative abundance) in both free-living and particle-attached bacterial communities in the bathypelagic samples were also abundant in the largest size fractions of the surface layers, but belonging to the mid-abundant and rare biosphere in the free-living fraction (Figure S18 and Table S25). Surface bacteria will act then as a seed bank for bathypelagic communities, an hypothesis that has been also proposed in other studies (Gibbons et al., 2013; Sebastián et al., 2017; Ruiz-González et al., 2020). It seems that our isolates belonged to the bacterial community that prefers a lifestyle linked to particles. In surface layers, they

may live attached to particles and sink to deeper layers where they grow and finally become part of the planktonic community. These will explain also why when we compare separately our photic and aphotic isolates to the different datasets, same patterns arise (Figure S5). We are therefore retrieving in pure culture isolates that belong to a community attached to particles, where nutrients are abundant, and we are missing most free-living bacteria adapted to oligotrophic conditions. Indeed, genomic comparison between cultured isolates and uncultured genomes retrieved by single amplified genomes (SAGs) from marine environments revealed that the genomes of the cultures had larger sizes, suggesting a predominant copiotrophic lifestyle (Swan et al., 2013). Therefore, our results of a higher proportion of isolated zOTUs described in the largest size particles, especially in the bathypelagic, are coherent with the hypothesis suggesting that the bathypelagic realm is constituted mainly by slow sinking particles (Herndl and Reinthaler, 2013), which are resource-rich habitats for microbes (Bochdansky et al., 2016), and therefore that a copiotroph lifestyle could be the rule for an important fraction of the heterotrophic bacteria inhabiting the deep ocean.

Conclusions

As far as we know, this is the first large-scale study combining results from culture-dependent and culture-independent techniques from the same water samples and from different oceanographic regions and depths, including the mesopelagic and the bathypelagic layers of the ocean. The comparison between isolates and HTS data from the free-living fraction confirmed that heterotrophic cultures are mainly found in very low abundances in the photic ocean, supporting the traditional paradigm that isolation techniques retrieve members of the rare biosphere. However, our results also confirmed that the axiom that less than 1% of bacteria can be retrieved by isolation should be revised as variability can be found among samples when including the mesopelagic and the bathypelagic ocean. This study revealed, based on 16S amplicon iTAGs, that a significant fraction of bacterial diversity can be retrieved by isolation, especially in the bathypelagic and associated to particles. Although amplicon iTAGs and metagenomic miTAGs data reflect similar patterns, the bacterial isolates represent a much lower proportion of the microbial community based on metagenomics. Future isolation efforts as dilution-to-extinction (Rappé et al., 2002; Selje et al., 2005), microdroplets or cultivation chips (Ingham et al., 2007), methods combining cell sorting with isotopic labelling (Thrash, 2020), or the use of metagenomic and metatranscriptomic data to predict the metabolic requirements of certain bacterial groups (Gutleben et al., 2017) may help us bring at the laboratory, aside the rare taxa, other some abundant key taxa in natural communities.

Chapter 3

Degradation capacity and biogeography of marine heterotrophic methylmercury detoxifying bacteria

Abstract

Bacterial transformation of inorganic divalent mercury (Hg^{+2}) and methylmercury (MeHg) is performed by the *mer* operon and is known to be harbored in diverse marine bacterial taxa. However, little is known about the tolerance capacity and phenotypic characterization of marine bacteria displaying *merA* or *merAB* genes. In this study, we performed a functional screening of these genes in 293 marine bacterial cultures from a large bacterial culture collection, belonging to different *Alteromonas* and *Marinobacter* spp., which could potentially have the *mer* operon and presented a broad geographical distribution, covering different oceanic depths and latitudinal gradients. About 25 % of the isolates presented the *merA* and only 8.8 % presented both *merAB* genes. On the basis of their 16S rRNA sequence, the Minimum Inhibitory Concentration (MIC) for $HgCl_2$ and MeHg was determined for 43 of them displaying a wide tolerance variability even within the same operational taxonomic unit (OTU). Strain ISS312 presented the highest tolerance, with a MIC of 70 µM and 10 µM for $HgCl_2$ and MeHg, respectively. This strain, affiliated to *Alteromonas mediterranea*, presented a 13 h longer lag-phase and a slightly slower growth rate when growing with MeHg, but reached a similar cell concentration and carrying capacity compared to the control grown without MeHg. Additionally, it achieved a degradation efficiency of 98.2 % in 24 h when grown at 5 µM of MeHg. The complete genome sequence of the strain ISS312 allowed us to explore its biogeographic distribution against global bathypelagic metagenomes, indicating also that it was a widely distributed bacteria across all the bathypelagic ocean. In summary, our functional screening analyses helped to understand the capacity of marine bacteria to detoxify MeHg and to propose *Alteromonas* sp. ISS312 as a promising marine bacterium to be implemented in future marine MeHg bioremediation processes.

Introduction

Mercury (Hg) is one of the most toxic, widespread and worrisome contaminant (Miller and Clarkson, 1973; Clarkson, 1993), which is emitted to the atmosphere by natural sources, such as volcanoes and rock weathering, but also by anthropogenic activities. Indeed, it has been estimated that anthropogenic Hg emissions have enriched present-day atmosphere by a factor of 2.6 relative to 1840 levels, and by a factor of 7.5 relative to natural levels (Amos et al., 2013). Both elemental (Hg^0) and inorganic divalent Hg (Hg^{+2}) can be deposited to land and oceans by wet and dry depositions (Enrico et al., 2016; Saiz-Lopez et al., 2018). Mercury in the ocean can be then volatilized back again to the atmosphere as Hg^0 in surface waters (Mason and Sheu, 2002), but it can be methylated (Monperrus et al., 2007; Lehnherr et al., 2011; Podar et al., 2015; Gionfriddo et al., 2016; Munson et al., 2018) forming the highly toxic methylmercury (MeHg), characterized by its neurotoxicity (Bakir et al., 1973; Harada, 1995; Clarkson and Magos, 2006). MeHg levels in the oceans vary with depth, and usually, measures are being reported low in open ocean surface waters, maximal in intermediate layers, especially in regions of low-oxygen and near or below the thermoclines (up to 1000 m depth), and low and relatively constant in deeper waters (>1000 m depth) (Cossa et al., 2009; Sunderland et al., 2009; Mason et al., 2012). It has been reported that Hg^{+2} methylation occurs both in oxic and sub-oxic layers of the water column (Cossa et al., 2009; Malcolm et al., 2010; Lehnherr et al., 2011; Hammerschmidt and Bowman, 2012; Blum et al., 2013) mainly associated with the microbial remineralization of sinking particulate organic matter (Cossa et al., 2009; Sunderland et al., 2009; Lamborg et al., 2016). MeHg bioaccumulates and biomagnifies in aquatic food webs (Mason et al., 2012; Harding et al., 2018; UN Environment, 2019), and as a consequence, humans are exposed to this neurotoxicant mainly through fish and seafood consumption (Mergler et al., 2007; Karagas et al., 2012; Mason et al., 2012).

The high toxic effect of Hg and the rising levels since the industrial era, estimated as an increase of 450 % of Hg in the atmosphere (UN Environment, 2019), makes the study of its biogeochemical cycle a major concern to the scientific community and also to all the governments around the world. The Minamata convention, held in 2013 and entering into force in August 2017 (Eriksen and Perrez, 2014), has been created to regulate the use of Hg and its releases to the environment in order to reduce its current levels. One of the biological alternatives to the removal of Hg in contaminated sites is the use of Hg resistant bacteria as a bioremediation strategy. These bacteria harbor in their genomes the *mer* operon, that can be composed by different sets of genes (Barkay et al., 2003; Boyd and Barkay, 2012). However, the operon key genes are *merA* and *merB*. The first one codifies a mercuric reductase, a cytosolic flavin disulfide oxidoreductase that uses NAD(P)H as a reductant, and is responsible for the transformation of Hg^{+2} to the less harmful and volatile Hg^0 (Barkay et al., 2003). The *merB* gene encodes an organomercurial lyase enzyme that confers resistance to the organic MeHg form. It is the responsible for its demethylation, transferring a proton to a C-H bound in alkyl-mercurials, releasing Hg^{+2} that will be then reduced to Hg^0 by *merA* (Barkay et al., 2003). These machineries have been found in numerous microorganisms including aerobic and anaerobic species, although demethylation appears to be predominantly accomplished by aerobic organisms (Oremland et al., 1991; Boyd and Barkay, 2012). As MeHg degradation is thought to be predominantly a microbial mediated process in water (Barkay et al., 2003), combined with the fact that different concentrations of MeHg can be found through the ocean water column (Heimbürger et al., 2010; Hammerschmidt and Bowman, 2012; Wang et al., 2012; Gworek et al., 2016), isolation of Hg resistant bacteria in aquatic ecosystems and the assessment of their tolerance to different concentrations of MeHg appears to be an interesting starting

point for bioremediation strategies. Most of the studies focusing on the isolation of marine Hg resistant bacteria have targeted coastal seawaters (Nakamura et al., 2000; De and Ramaiah, 2007; Zhang et al., 2012), sediments (De and Ramaiah, 2007; Jayaprakashvel et al., 2015), mangroves and estuaries (Chiu et al., 2007; Deng and Wang, 2012; Oyetibo et al., 2015), hydrothermal vents (Vetriani et al., 2005), or highly contaminated sites (Nakamura et al., 2001; Lima de Silva et al., 2012). Nonetheless, to the best of our knowledge, a culture-dependent study aiming to detect marine resistant bacteria from different depths, including the bathypelagic ocean, as well as different oceanographic regions, has not been addressed before. In this study, we performed a functional screening of the *merA* and *merB* genes in 293 marine bacterial cultures selected from a large heterotrophic culture collection (MARINHET) (Sanz-Sáez et al., 2020) with the aim to detect marine bacteria codifying *mer* genes, and assess their tolerance to different concentrations of inorganic and organic Hg to further be implemented in future MeHg bioremediation processes.

Materials and methods

Selection of marine strains for *merAB* functional screening

A total of 2003 marine strains were previously isolated from a wide variety of oceanographic regions and depths. Detailed sampling, isolation procedures and partial sequencing of the 16S rRNA gene have been already described in Sanz-Sáez et al. (2020). This heterotrophic marine bacterial culture collection, called MARINHET, was the basis for the functional screening of Hg resistant bacteria. First, a preliminary search of marine genera codifying both *merA* and *merB* in all available finished genomes in the Integrated Microbial Genomes (IMG) database (March 2016) of the Joint Genome Institute (JGI) was done by: (i) searching the functional annotation of *merA* (as mercuric reductase or mercuric ion reductase) and *merB* (organomercurial lyase or alkylmercury lyase), and (ii) using the KEGG orthologs (Kanehisa and Goto, 2000) K00520 for *merA* and K00221 for *merB*. Secondly, we downloaded the 16S rRNA gene sequences of those genomes in order to BLASTn (Altschul et al., 1990) them against the partial 16S rRNA sequences of our isolates and to obtain a list of putative candidates for Hg bioremediation. From those, we focused on two genera, *Alteromonas* and *Marinobacter*, that were highly abundant in our MARINHET cultured collection. Therefore, 293 isolates from both *Alteromonas* (247) and *Marinobacter* (46) genera were finally selected for functional screening of the Hg resistance genes *merA* and *merB*. These strains were isolated from a variety of oceanographic regions and depths covering both the photic and aphotic regions of the water column. Information about the origin of the samples where isolates were retrieved is summarized in Table 4.

Primers design for *merA* and *merB* genes

Sequences of the *merA* and *merB* genes were downloaded from IMG/JGI. First, we looked for "mercuric reductase" or "mercuric ion reductase" (*merA*) and "organomercurial lyase" or "alkylmercury lyase" (*merB*) genes in the published finished genomes of 22 *Alteromonas* and 6 *Marinobacter* (IMG database in 2016). Then, we downloaded the nucleotide sequences of 15 *merA* and 1 *merB* genes present in those *Alteromonas* genomes, as well as 9 *merA* and 1 *merB* genes from *Marinobacter* (Table S26). Specific

Table 4: Characteristics of the marine seawater samples from five different cruises used for the isolation of bacteria screened for Hg resistance genes. ATP09: Arctic Tipping Points cruise in 2009; BBMO: Blanes Bay Microbial Observatory.

Station	Oceanic location	Latitude	Longitude	Depth (m)	N° of screened isolates
Tara Oceans					
ST 39	Indian Ocean	18° 35.2′ N	66° 28.22′ E	25	44
ST 67	South Atlantic	32° 17.31′ S	17° 12.22′ E	5	20
ST 72	South Atlantic	8° 46.44′ S	17° 54.36′ W	5	15
ST 76	South Atlantic	20° 56.7′ S	35° 10.49′ W	5	31
ST 85	Southern Ocean	62° 2.19′ S	49° 31.44′ W	5.9	1
ST 85	Southern Ocean	62° 2.19′ S	49° 31.44′ W	87.4	6
ST 151	North Atlantic	36° 10.17′ N	29° 1.23′ W	5	31
ATP09					
AR_1	Arctic Ocean	78° 20.00′ N	15° 00.00′ E	2	3
AR_2	Arctic Ocean	76° 28.65′ N	28° 00.62′ E	25	4
Malaspina					
ST 10	North Atlantic	21° 33.36′ N	23°26′ W	4002	11
ST 17	South Atlantic	3° 1.48′ S	27° 19.48′ W	4002	3
ST 23	South Atlantic	15° 49.48′ S	33° 24.36′ W	4003	13
ST 32	South Atlantic	26° 56.8′ S	21° 24′ W	3200	14
ST 33	South Atlantic	27° 33.2′ S	18° 5.4′ W	3904	2
ST 43	South Atlantic	32° 48.8′ S	12° 46.2′ E	4000	3
MIFASOL					
ST 8	NW Mediterranean	40° 38.41′ N	2° 50′ E	2000	85
BBMO					
IBSURF	NW Mediterranean	41° 40′ N	2° 48′ E	5	7

Table 5: Summary of the primers sequences designed for the screening of *merA* and *merA+merB (merAB)* genes. Some of the primers display degenerated bases (see in bold): R: A or G; S: G or C.

Primer	Primer sequence	Position	PCR product length (bp)
merA Alteromonas			
Forward	5′ GTG CAA GGG TCA CCA TAA TA 3′	116	1162
Reverse	5′ CCT TCG TGT GCT AAG ATT TG 3′	1227	
merA Marinobacter			
Forward	5′ CTG CAC ATT **R**C**S** GTG ATT GG 3′	16	1368
Reverse	5′ TTG CTT **S**AC **R**TC CTT GGT 5′	1383	
merA+merB Alteromonas			
Forward	5′ AGG CAT CTT GCA TAC CAT TT 3′	64	2066
Reverse	5′ CTC CTT GCC CTG TTG AGT AT 3′	2129	
merA+merB Marinobacter			
Forward	5′ CCG TCC AAA ATC ATG ATC CG 3′	147	1700
Reverse	5′ ACC GAC ACT AAA GAG ACG AC 3′	1847	

primers pairs were designed separately for: (i) *merA* of *Alteromonas*, (ii) *merA* of *Marinobacter*, (iii) *merA* + *merB* of *Alteromonas*, and (iv) *merA* + *merB* of *Marinobacter*. Both sets of *merA* primers were selected after alignments of the sequences in Clustal Omega[9] in order to check for the most conserved areas of the gene. For *Alteromonas*, as different *merA* gene copies could be present within the same genome, the designed primers only covered the gene/s copies with the highest similarity between them and between the taxa, since evidence of horizontal gene transfer has been detected in *mer* genes, usually in plasmids, transposons or genomic islands (Ivars-Martinez et al., 2008) (Table S26). On the other hand, as *Marinobacter merA* sequences differ greatly from species to species and most of the *Marinobacter* isolates present in our culture collection affiliated to *Marinobacter hydrocarbonoclasticus* or *Marinobacter salarius*, primers were designed to cover the *merA* sequence variants of these species (Table S26). Otherwise, as *merB* genes were only found in *Alteromonas mediterranea* DE strain (CP003917) and *Marinobacter aquaeolei* VT8 strain (NC_008740) (also named *Marinobacter hydrocarbonoclasticus* VT8), *merA* + *merB* sets of primers were designed with the online tool Primer-BLAST of the National Center for Biotechnology Information[10]. The input sequence for the generation of these primers was a concatenate nucleotide sequence of the *merA* and *merB* genes that were co-localized one next to the other in the *mer* operon (Figure S19). All the designed primers had to meet the following requirements: optimum PCR product size of 1200 bp for *merA* or 2100 bp for *merA* + *merB* (referred hereafter as *merAB*), annealing temperature around 57 ºC, primers length 20 bp, and 50 % GC content. In Table 5 we summarize the sequence of the different sets of primers used in this study.

DNA extraction and PCR conditions

The primers previously designed were used for the screening of *merA* and *merAB* Hg resistance genes in the 293 selected strains. DNA of all the strains was extracted from 48 h liquid cultures grown in Zobell broth medium (i.e. 5 g peptone, 1 g yeast extract in 750 ml of 30 kDa filtered seawater and 250 ml of Milli-Q water) using the DNeasy Blood and Tissue kit (Qiagen) following the manufacturer's recommendations. Each PCR reaction with a final volume of 25 µl contained: 2 µl of template DNA, 0.5 µl of each deoxynucleotide triphosphate at a concentration of 10 µM, 0.75 µl of $MgCl_2$ 1.5 mM, 0.5 µl of each primer at a concentration of 10 µM, 0.125 µl of Taq DNA polymerase (Invitrogen), 2.5 µl of PCR

[9]https://www.ebi.ac.uk/Tools/msa/clustalo/
[10]https://www.ncbi.nlm.nih.gov/tools/primer-blast

buffer supplied by the manufacturer (Invitrogen, Paisley, UK) and Milli-Q water up to the final volume. Reactions were carried out in a Biorad thermocycler using the following program: initial denaturation at 94 ºC for 5 min, followed by 30 cycles of 1 min at 94 ºC, 1 min at 55 ºC and 2 min at 72 ºC, and a final extension step of 10 min at 72 ºC. The PCR products were verified and quantified by agarose gel electrophoresis with a standard low DNA mass ladder (Invitrogen). Purification and OneShot Sanger sequencing of *merA* and *merAB* gene products was performed by Genoscreen (Lille, France) with both forward and reverse primers. Geneious software v.11.0.5 (Kearse et al., 2012) was used for manual cleaning and quality control of the sequences.

Phylogenetic analyses of the 16S rRNA genes of isolates containing *merA* or *merAB* genes and of the amino acid sequences of the targeted genes

A phylogeny of the isolates screened by PCR for *merA* and *merAB* genes was inferred from their partial 16S rRNA sequences in order to detect a possible clustering between all the positive strains. The closest sequence to each isolate 16S rRNA gene in SILVA v.132 database was found and collected using BLASTn (Altschul et al., 1990). Alignment of the isolates and reference sequences was performed with MUSCLE from the Geneious software v.11.0.5 (Kearse et al., 2012). The alignment was trimmed to the common 16S rRNA gene fragment covered by both sets of sequences. Phylogeny was constructed using maximum-likelihood inference with RAXML-NG 0.9.0 (Kozlov et al., 2019) and the GTR evolutionary model with optimization in the among-site rate heterogeneity model and the proportion of invariant sites (GTR+G+I), and 100 bootstrap replicates. In the same way a phylogenetic tree was constructed with the partial 16S rRNA sequences of the positive isolates only. In this tree the closest match in SILVA v.132 database was also included. Presence of *merA* and *merAB* genes, origin of the strains, plus their MIC to $HgCl_2$ and MeHg were added with ITOL (Letunic and Bork, 2019). On the other hand, phylogenetic trees were also constructed with the amino acid sequences of the amplified *merA* and *merAB* genes. As reference sequences we included the best BLASTn hits against UniProtKB sequences retrieved with KEGG identifiers K00520 (*merA*) and K00221(*merB*). Sequences were aligned with ClustalW of the Genious software v.11.0.5 (Kearse et al., 2012) with the Gonnet substitution matrix and default gap extension and opening penalties as described previously (Boyd and Barkay, 2012). For *merA*, the dihydrolipoamide dehydrogenase protein sequences from *Magnetospirillum magneticum* AMB-1 (WP_011386317.1) and *Pseudomonas fluorescens* Pf0-1 (WP_011336663.1) served as outgroups. Likewise, we included, only in the alignment, the *merA* sequence of *Streptomyces lividans* (P30341), which served to trim the N-terminal region of the aligned sequences such that only the core domain of *merA*, corresponding to positions 1–464 of this *Streptomyces lividans* remained (Barkay et al., 2010). In the case of *merAB*, outgroup sequences were not included. Phylogenetic trees were constructed using maximum-likelihood inference with RAXML-NG 0.9.0 (Kozlov et al., 2019) and the LG evolutionary model with optimization in the among-site rate heterogeneity model and the proportion of invariant sites (LG+G+I), and 100 bootstrap replicates.

Minimum inhibitory concentration experiments

To assess the ability of the marine strains to tolerate different concentrations of inorganic Hg (mercury(II) chloride, $HgCl_2$) and organic Hg (methylmercury chloride, CH_3HgCl) and thus, to test the activity of *merA*

and *merB* genes respectively, a Minimum Inhibitory Concentration (MIC) assay was designed based on previous studies (Ivars-Martinez et al., 2008; Wiegand et al., 2008). Stock solution of $HgCl_2$ was prepared at 500 µM with autoclaved Milli-Q water. Liquid cultures of the strains growing in Zobell broth with an optical density (O.D. at 600nm) of 0.1 were placed in 24-well plates and inoculated with $HgCl_2$ to reach final concentrations at 5 µM, 10 µM, 20 µM, 25 µM and 50 µM. In specific cases growth was observed in all $HgCl_2$ concentrations and further MIC assays were done increasing the final concentrations at 50 µM, 60 µM, 70 µM, 80 µM, 90 µM and 100 µM. On the other hand, a stock solution of CH_3HgCl was prepared at 50 µM. In these case, 24-well plates were inoculated with the stock solution to reach final concentrations at 2.5 µM, 5 µM, 10 µM, 15 µM and 20 µM. In all plates a positive control (0 µM, liquid culture of the strain without CH_3HgCl) and a negative control (liquid culture of a non-resistant marine strain of the genus *Erythrobacter*) were included in the assays. Plates were sealed with parafilm and incubated at room temperature (RT, approximately 20 ºC) in the dark for 72 h. Visual lectures and O.D. lectures at 600 nm were done in a 24 h period using an automatic plate reader (Infinite ® M200, Tecan) and data was collected using the Magellan™ Data Analysis Software (Tecan Diagnostics®).

Growth curves

Growth curves were performed to characterize the growth rates of the most tolerant bacteria to different concentrations of MeHg. We prepared 200 ml of different liquid cultures in Zobell broth supplemented with $CH_(3)HgCl$ at final concentrations of 0 µM (positive control), 1 µM, 2.5 µM and 5 µM. Each concentration was prepared in triplicate. The initial O.D. at 600 nm of the cultures was 0.05 in order to assure enough concentration of cells for growing. Samples for O.D. measurements and for bacterial cell counts were taken approximately every two hours. O.D. was measured at 600 nm with a spectrophotometer (Varian Cary® 100 UV-Vis) and cells were stained with 4',6'-diamidino-2-phenylindole (DAPI) and counted with an automated microscope Zeiss Axio Imager Z2M (Zeder and Pernthaler, 2009; Zeder et al., 2011) using the automated image analysis software ACME Tool[11]. Predicted growth curves based on O.D. observations and kinetics values (growth rates (μmax), carrying capacity (k) and lag phase time) were calculated with R package *growthcurver* v.0.3.0 (Sprouffske and Wagner, 2016) and GrowthRates v.4.3 software (Hall et al., 2014). For graphical representation, replicates of the different growth curves experiments at several MeHg concentrations were averaged. Hence, mean O.D. and standard deviation was calculated for each time point of the curves.

Measurement of the biotic and abiotic degradation of MeHg

In order to characterize the MeHg degradation rates of our bacteria, either caused by the action of the *merA* and *merB* genes, or by other abiotic processes we measured the MeHg concentrations along time in different liquid cultures. A first set of 2 ml samples was taken during the 1 µM and 5 µM growth curves of the most tolerant strain, *Alteromonas* sp. ISS312, at times 0, 6, 12, 24 and 48 h. A second set of samples was taken from 6-well plates experiments, which included one well of a liquid culture from the ISS312 strain in Zobell broth (initial O.D. at 600 nm of 0.05) supplemented with 5 µM CH_3HgCl and incubated at RT during 72 h in the dark, as well as two different controls to detect the possible abiotic degradation of MeHg: (i) medium control (CH_3HgCl at concentrations of 1 µM and 5 µM with no strain

[11] www.technobiology.ch

added), and (ii) killed control (CH_3HgCl at concentrations of 1 µM and 5 µM added to autoclaved liquid cultures of the strain). Three replicates of 2 ml samples from each well, containing the MeHg treatment and the different controls, were taken at times 0 and 72 h. Once samples were collected, they were immediately frozen at -80 ºC to preserve concentrations of MeHg at the time of sampling. Hg species determination was conducted by direct derivatization of the culture samples with sodium tetraethyl borate and injection into a hyphenated system consisting of a gas chromatograph coupled to an atomic fluorescence detector via pyrolysis (GC-pyro-AFS) as previously described elsewhere (Berzas Nevado et al., 2011). Briefly, 2 ml samples were used for derivatization. The pH of the extracts was adjusted to 3.9 by adding 5 ml of 0.1 M acetic acid-sodium acetate buffer and ammonia (20 %) if necessary. Then, 2 ml of hexane and 250 µl of sodium tetraethyl borate (6 %, w/v) were added and the mixture was manually shaken for 5 min. The sample was centrifuged for 5 min at 600 g. The organic layer was transferred to a chromatographic glass vial and stored at -18 ºC until analysis. When Hg species were not detectable, the organic layer was preconcentrated under a gentle stream of nitrogen to a low volume (50-100 µL) just before the measurement. The procedural detection limits, after pre-concentration, were 0.19 and 0.23 nM for MeHg and inorganic Hg, respectively.

TEM-EDX samples preparation and observation

For TEM observation of the morphology and ultrastructure of strain ISS312 when growing with MeHg, the isolate was grown during 24 h in Zobell broth supplemented with CH_3HgCl at final concentrations of 0 µM (positive control), and 5 µM with shaking in the dark. The overnight culture was centrifuged at 1000 g during 15 min and supernatant was discarded. The pellet was fixed with paraformaldehyde 2 % final concentration during 30 min at RT. After fixation, pellet was processed as previously described (Lee et al., 2008) to finally obtain thin sections of the samples that were examined by using TEM (JEM-1400 plus, JEOL). Visualizations were done by the microscopy service of the *Universitat Autònoma de Barcelona*[12].

ISS312 genome sequencing

DNA from strain ISS312 was extracted using the DNeasy Blood and Tissue kit (Qiagen), following the manufacturer's recommendations. Genome library was prepared with a CeleroTM DNA-seq library system and sequencing was performed with paired-ended 300 bp long reads by IGA-Tech with a MiSeq Illumina machine. The sequence data was filtered to remove the adapters and the unpaired reads with *cutadapt* v.1.16. and the quality was assessed before and after with *fastaqc* v.0.11.7. The clean data was used to do the assembly with Spades v.3.12.0 and then optimized with the same program. QUAST v.5.0.1 (Gurevich et al., 2013) and ALE were used to assess the quality of the assemblies and the best scores were selected. K-mer 121, 125 and 127 were selected for the optimized-combined assembly and the quality was assessed again. The annotation was done with Prokka v1.13 (Seemann, 2014) and the completeness of the genome was checked with CheckM v1.0.18 (Parks et al., 2015). In order to search for plasmids within our contigs, we used the database PLSDB (Galata et al., 2019).

[12]https://sct.uab.cat/microscopia/en/content/inici

Fragment recruitment analysis of the genome ISS312 and selected bathypelagic metagenome-assembled genomes (MAGs) in the bathypelagic metagenomes

The abundance of ISS312 across the global bathypelagic ocean was assessed, thanks to its complete genome, through Fragment Recruitment Analysis (FRA), by mapping the metagenomic reads of 58 Malaspina samples from 32 stations (Duarte, 2015), including free-living (0.2-0.8 µm) and particle-attached (0.8-20µm) bacterial communities, with BLASTn v2.7.1+ (Altschul et al., 1990) with the following alignment parameters: *-perc_identity* 70, *-evalue* 0.0001. Only those reads with more than 90 % coverage and mapping identity equal to or higher than 95 % were kept for analysis. In order to remove possible false mapping hits to the conserved regions of rRNA genes, reads aligning to the regions annotated as ribosomal genes were not considered for the analysis. Read counts from mapped reads from each metagenome were corrected by their sequencing depth to make them comparable through samples.

A total of 76 metagenome-assembled genomes (MAGs) from the bathypelagic ocean (Acinas et al., 2019) were screened to extract those affiliating to *Alteromonas* and *Marinobacter* genera and codifying *merA* and *merB* genes. The abundance of these MAGs across the Malaspina bathypelagic dataset was done as described above, but with a previous subsampling step to the shallower sequencing depth (4,175,346 read pairs) with *bbtools reformat.sh* v.38.08 (https://sourceforge.net/projects/bbmap/). Here, when reads could map with the same probability to any of the genomes (same e-value, same alignment length and identity), they were assigned at random. The average nucleotide identity (ANI) of ISS312 against MAG-0289 was calculated with fastANI v1.2 as both were related to the *Alteromonas mediterranea* specie.

Non-parametric Kruskal-Wallis test, from the *stats* package of the R Statistical Software (core team, 2017) was applied followed by the post hoc pairwise Wilcox test to see the differences between FRA results in different oceanographic regions and between free-living and particle-attached bacterial communities . To asses significance, the statistical analyses were set to a conservative alpha value of 0.05.

Results

Functional screening of the *merA* and *merB* genes in *Alteromonas* and *Marinobacter* strains

We detected from IMG/JGI database 20 different bacterial taxa that matched at the genera level with the taxonomic assignation of the isolates within the MARINHET culture collection, and therefore, that could be putative candidates for Hg bioremediation (Table S27). The comparative analyses between the 16S rRNA gene sequences of these 20 genera containing the targeted *merAB* genes and the partial 16S rRNA sequences of our isolates revealed a total of 352 strains that were, at least, 99 % identical to one of the putative candidates genera. These comprised 7 genera (*Alteromonas*, *Marinobacter*, *Idiomarina*, *Pseudomonas*, *Micrococcus*, *Zunongwangia* and *Bacillus*) (Table S28). We selected 247 strains affiliating to *Alteromonas* sp. and 46 strains to *Marinobacter* sp. (Table S29). These two genera were chosen for *merAB* functional screening for three main reasons: (i) they were found highly abundant in our MARINHET culture collection, (ii) they are among the most common culturable heterotrophic bacteria living in open marine waters all around the world, as they have been isolated from a wide variety of marine environments (Baumann et al., 1972; Eilers et al., 2000; Floyd et al., 2005; Gärtner et al., 2011;

Table 6: Summary of the PCR screening results for *merA* and *merAB* in *Alteromonas* and *Marinobacter* strains. Photic includes surface and deep chlorophyll maximum (DCM) isolates, while aphotic includes bathypelagic isolates.

Layer	Nº of tested strains	Positives PCR for		Total strains with *merA* and/or *merAB*
		merA	*merAB*	
Alteromonas				
Photic	129	13 (10%)	1 (0.7%)	33 (13.3%)
Aphotic	118	19 (16.1%)	3 (2.5%)	
Marinobacter				
Photic	33	30 (90.9%)	20 (60.6%)	41 (89.1%)
Aphotic	13	10 (76.6%)	2 (15.4%)	

Handley and Lloyd, 2013; Lekunberri et al., 2014; Kai et al., 2017), and in the case of *Alteromonas* it is one of the most ubiquitous cultured taxa in the ocean (Sanz-Sáez et al., 2020), and (iii) it has been already described that species of those genera harbor in their genomes the *mer* operon (Ivars-Martinez et al., 2008; Singer et al., 2011; López-Pérez et al., 2012; Handley and Lloyd, 2013; Fontanez et al., 2015).

The functional screening of the *merA* and *merB* genes from the 247 *Alteromonas* and 46 *Marinobacter* strains revealed that 13 % (32 out of 247) and 87 % (40 out of 46) of the strains presented only *merA*, respectively, while only 1.6 % (4 out of 247), and 47.8 % (22 out of 46) presented both *merA* and *merB* genes (*merAB*) (Table 6). These results showed that *Marinobacter* displayed a higher proportion of *merAB* genes than *Alteromonas*. The phylogenetic analyses of the 16S rRNA gene from of all these positive strains harboring *merA* and *merAB* genes, together with the rest of the screened strains, displayed some clustering patterns (Figure 18). First, most of the isolated *Alteromonas* strains with *merA* were related to *Alteromonas* australica and *Alteromonas mediterranea*, and only strains affiliating to this last species presented both genes *merA* and *merB* (merAB). We also detected one strain with *merAB* genes affiliated to *Alteromonas macleodii*. Secondly, *Marinobacter* strains displaying both genes were related to *Marinobacter hydrocarbonoclasticus*, *Marinobacter salarius* and some uncultured *Marinobacter* strains. A closer view to the amino acids *merA* phylogeny, which theoretically could include the different gene sequences variants covered by the primers designed, unveiled that all the *Alteromonas merA* sequences grouped into two sister clades with the reference *Alteromonas mediterranea merA* sequence (Figure 19A), while for *Marinobacter* the *merA* genes grouped into three different clusters together with *Marinobacter salarius*, *Marinobacter hydrocarbonoclasticus* and the *Marinobacter* sp. Arc7-DN-1 strain, showing therefore a genetic heterogeneity among their *merA* genes copies (Figure 19A). In the case of the *merAB* phylogeny, as expected due to the primers used, all the sequences clustered with *Alteromonas mediterranea* and *Marinobacter hydrocarbonoclasticus* (Figure 19B).

Mercury tolerance of marine bacterial isolates

MIC experiments were addressed in a selection of 74 isolates that presented *merA* and/or *merAB* genes. This selection was based on a clustering of the isolates 16S rRNA gene sequences at 99 % sequence similarity to define operational taxonomic units (OTUs). The clustering grouped the isolates into 7 OTUs, and representatives of the different OTUs were randomly selected for MIC experiments (Table S30). First, we tested the tolerance for inorganic mercury ($HgCl_2$) and 43 isolates (19 *Alteromonas* and 24

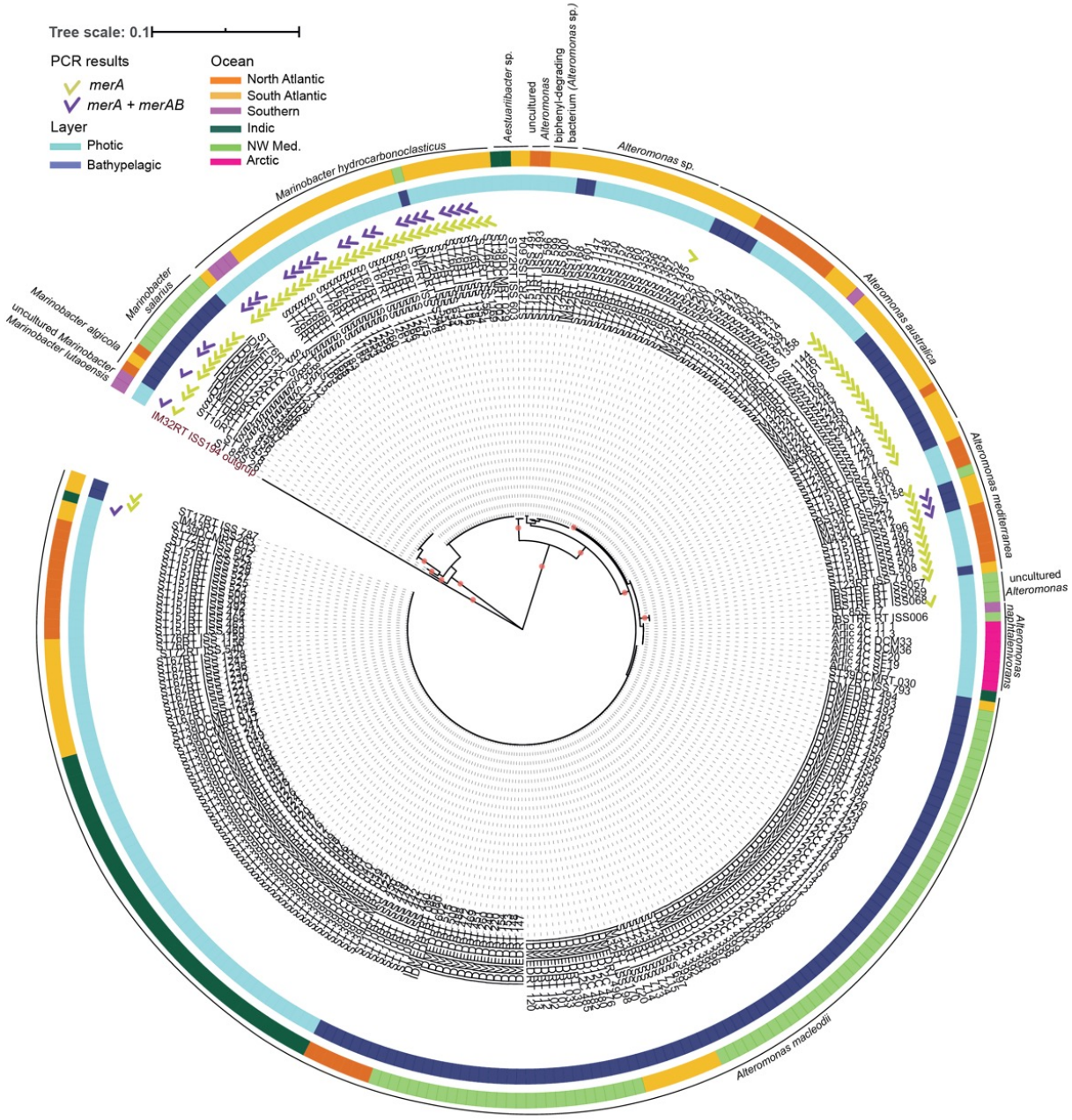

Figure 18: Phylogenetic tree with the 16S rRNA gene sequences of all the isolates PCR screened for *merA* and *merAB* genes. Presence or absence of both *merA* and *merAB* genes is indicated by tick symbols. Depth (first color strip) and oceanographic location (second color strip) where the isolates were retrieved are indicated by colors. Bootstrap values > 75 are indicated by red circles in the tree nodes. For graphic representation reference 16S rRNA sequences were removed from original tree, and names of these references sequences had been placed surrounding tree labels.

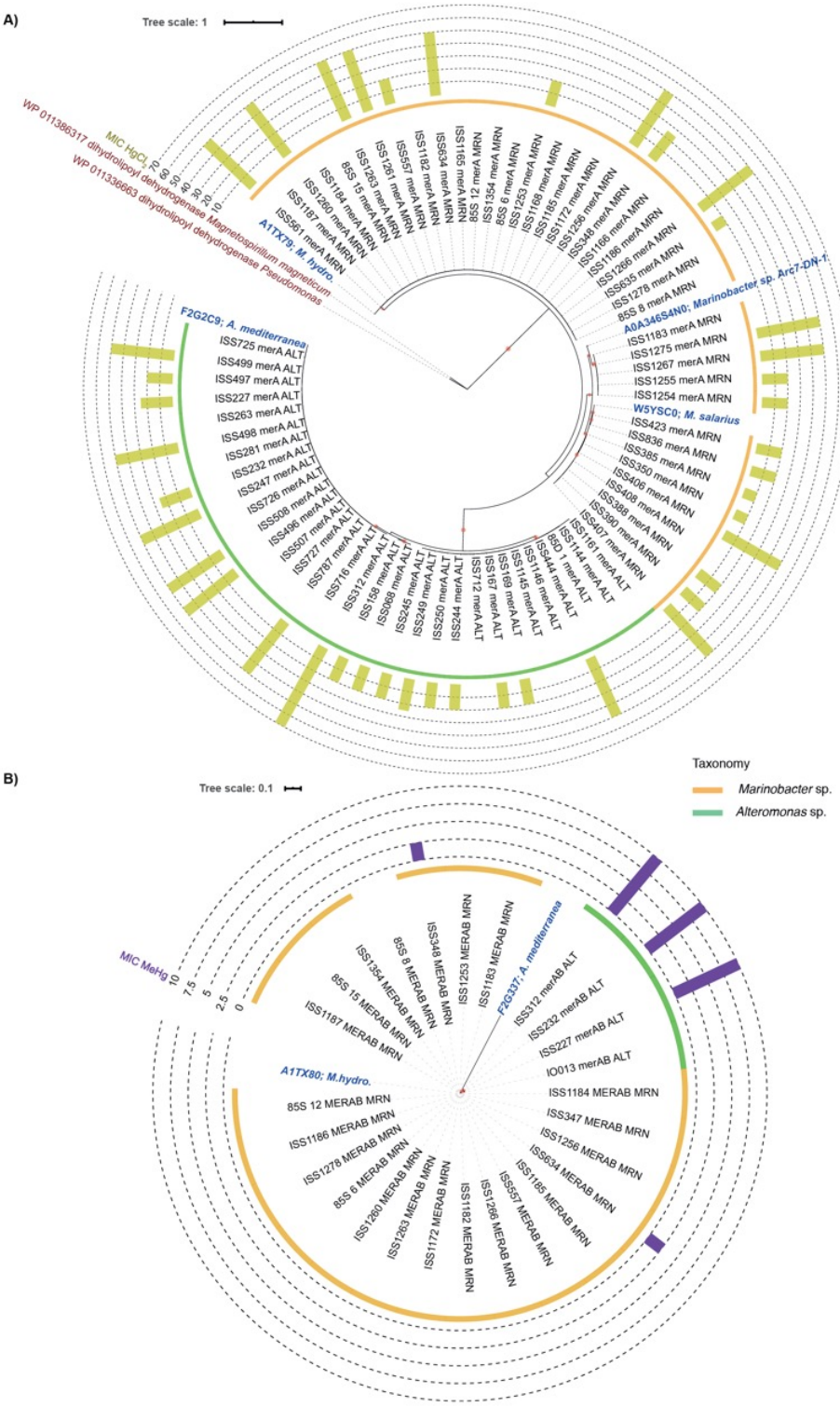

Figure 19: Phylogenetic trees inferred with the *mer* amino acid sequences. **(A)** *merA*. Sequences of dihydrolipoyl dehydrogenase act as outgroup of the tree. **(B)** *merAB*. Color strip in both trees indicates taxonomy of the sequences: *Alteromonas* sp., green; *Marinobacter* sp., orange. Reference sequences in each tree are indicate in bold blue: *A. mediterranea*, *Alteromonas mediterranea; M. hydro., Marinobacter hydrocarbonoclasticus; M. salarius, Marinobacter salarius.* MIC results for the tested strains against $HgCl_2$ or MeHg is indicated by bars in each tree. Bootstrap values >75% are indicate by red circles in the tree nodes.

Marinobacter) displayed different levels of tolerance. MIC values ranged generally from 10 to 50 µM and different tolerances were found within a same phylogenetic cluster based on 16S rRNA genes (Figure 20). Around 50 % of the *Alteromonas* and *Marinobacter* strains tested presented a MIC of 20 µM and one of the isolates stood out as it presented a tolerance to $HgCl_2$ up to 70 µM (Table S31). We noticed that MIC tolerance values also varied among isolates within OTUs at 99 % similarity in their 16S rRNA gene. For example, within the same cluster of *Alteromonas mediterranea* some strains presented a MIC of 20 µM, while the isolate that presented the highest tolerance (70 µM, ISS312) also belonged to the same *Alteromonas* species (Figure 20). The same occurred among *Marinobacter* isolates, where members of the *Marinobacter hydrocarbonoclasticus* cluster presented MIC values ranging from 10 to 50 µM (Table S31 and Figure 20). In addition, variability was also observed in the amino acids *merA* phylogeny. *Alteromonas mediterranea merA* sequences clustering together presented MIC values ranging from 20 µM and 70 µM, those of the *Marinobacter* sp. Arc7-DN-1 cluster varied between 20 µM and 50 µM, while those of the *merA* sequences affiliating to *Marinobacter salarius* and *Marinobacter hydrocarbonoclasticus* ranged from 10 µM to 50 µM (Figure 19A).

On the other hand, in order to test the tolerance to methylmercury (MeHg), 3 strains affiliating to *Alteromonas mediterranea* that codified for the *merB* gene and presented a MIC for $HgCl_2$ above 20 µM were selected. Remarkably, these strains presented a high tolerance to MeHg, growing at concentrations up to 10 µM (Figure 20). Interestingly, despite the fact that for *Marinobacter* sp. a higher number of positive strains were observed for *merAB* genes, the 2 strains encoding the *merB* gene and also having a MIC for $HgCl_2$ above 20 µM did not show a substantial growth above 2.5 µM of MeHg.

Effect of methylmercury on the growth of the highly resistant *Alteromonas* sp. ISS312 strain

Strain ISS312, isolated from South Atlantic bathypelagic waters at 4000 m, affiliated to the species *Alteromonas mediterranea*. It displayed the highest tolerance to both $HgCl_2$ (70 µM) and MeHg (10 µM) and it could be a good candidate for future bioremediation studies in highly contaminated areas with both organic and inorganic mercuric compounds. Consequently, the growth rates of this isolate at different concentrations of MeHg were assessed. Tested concentrations were selected based on MIC results and included: a control without MeHg (0 µM), 1 µM, 2.5 µM and 5 µM MeHg. Growth curves at 0 µM and 1 µM were very similar, as well as between 2.5 µM and 5 µM. We observed that the major difference between growth curves was the length of the lag phase, where bacteria adapt themselves to the growth conditions. Cultures showed a longer lag phase in the highest concentration of MeHg, around 13 h, compared to the control, which started to grow immediately after inoculation (Figure 21A). Lag phase length declined as long as the concentration of MeHg decreased, being 6 h at 2.5 µM and 2 h at 1 µM. However, once the cultures started to grow, their growth rates (μmax) were very similar independently of their initial MeHg concentrations, ranging from 0.10 h^{1} in the control to 0.09 h^{1} at 5 µM. Stationary phase was reached in all concentrations at 80 h, even though at this time cultures at higher concentrations of MeHg seemed to be only entering to the plateau (Figure 21A). In addition, their carrying capacity (k), i.e. the maximum population size of a species, was between 1.6 and 1.9 based in O.D. measures, revealing very similar values between tested concentrations. TEM observations of the ISS312 cultures growing at 0 µM and 5 µM of MeHg also showed similar morphology and ultrastructure of the cells (Figure 21A).

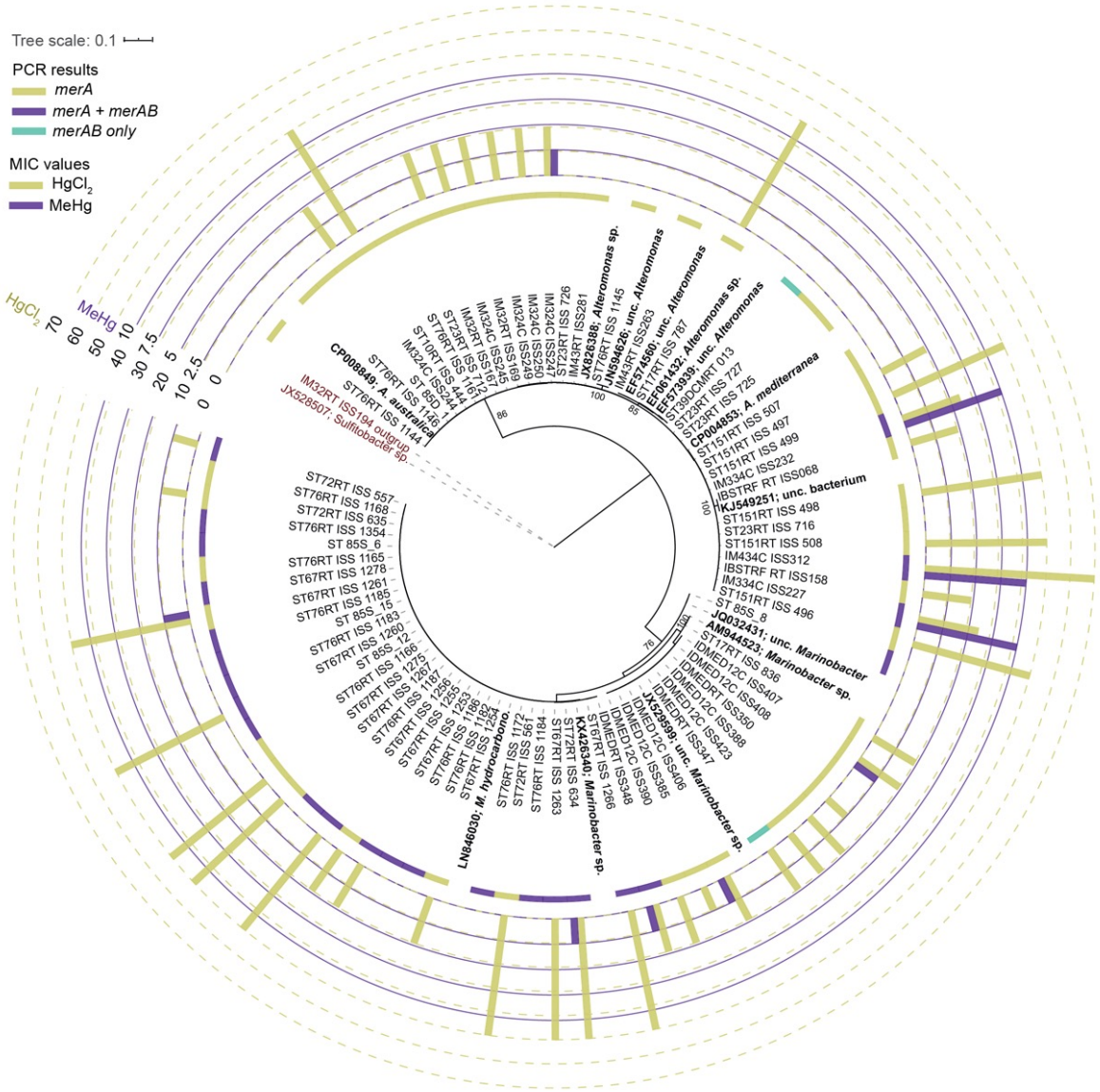

Figure 20: Phylogenetic 16S rRNA tree with the positive strains for *merA* and/or *merAB* genes screening. Colored strip indicate presence or absence of genes based on PCR results: yellow, only *merA*; turquoise, *merAB*; purple, positive for both *merA* and *merAB*. Bars indicate results from the MIC experiments: yellow, $HgCl_2$; purple, MeHg. Tolerance values are in μM. JX52807, *Sulfitobacter* and IM32RT_ISS194 are outgroups of the tree. The numbers in nodes represents bootstrap percentages > 75%. *A. australica*, *Altermonas australica; A. mediterranea, Alteromonas mediterranea*; unc., uncultured; *M. hydrocarbon., Marinobacter hydrocarbonoclasticus.*

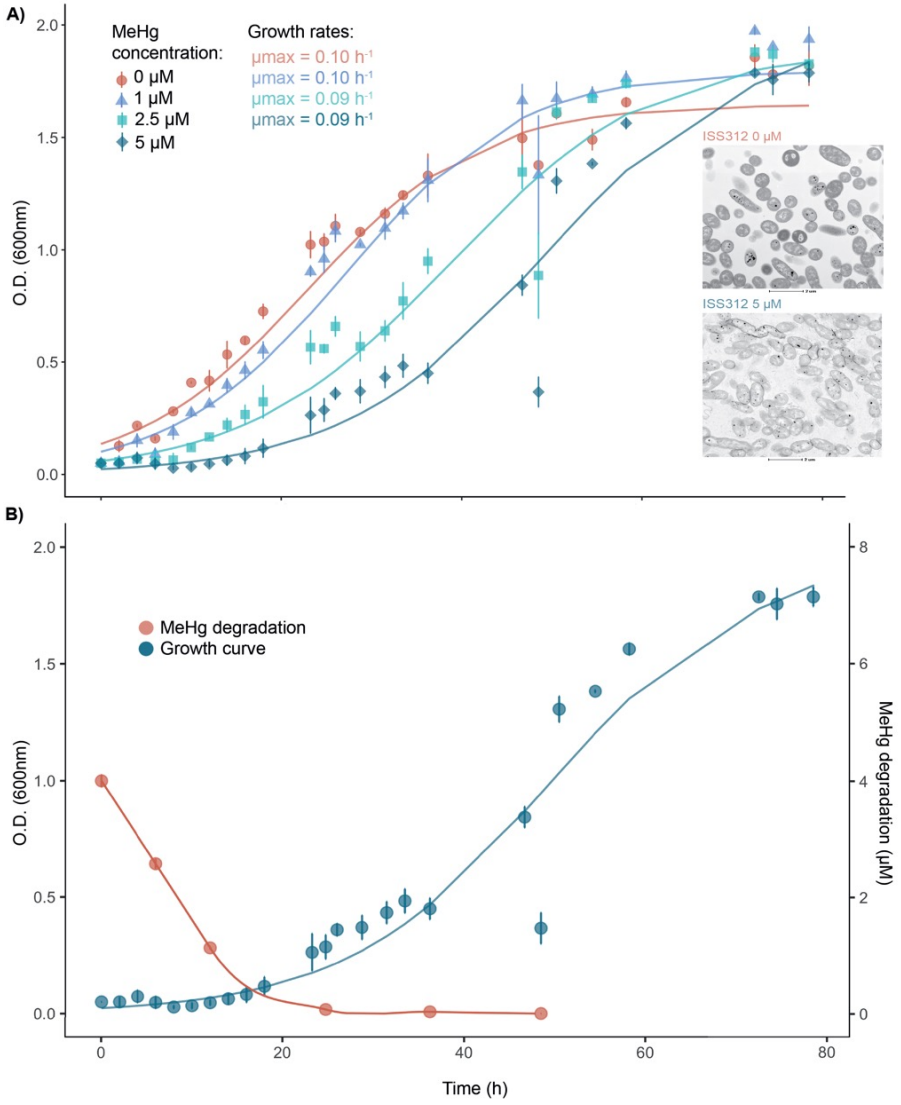

Figure 21: Growth effect of MeHg in strain ISS312. **(A)** Growth kinetics of strain *Alteromonas mediterranea* ISS312 in Zobell broth containing MeHg (control (0 µM), 1 µM, 2.5 µM and 5 µM). *µmax* indicates the growth rate for each MeHg concentration. TEM images of the strain growing at 0 µM and at 5 µM are shown in the right side of the plot. **(B)** MeHg removal by strain ISS312 during the growth curve experiment at 5 µM. Mean and standard deviation from three replicates samples are shown in both graphs.

Mercury-removal capacity of strain ISS312

Samples for the measurement of MeHg concentrations during the growth of the strain ISS312 in experiments at 5 µM were taken at different time points. Samples were also taken for the 1 µM growth curves, but as similar degradation tendencies were observed (Table S32), here we will only describe the results from the 5 µM growth curve. We observed that initial MeHg concentrations did not match perfectly with the expected concentration of 5 µM and, instead, a value of 4 µM was measured at time 0 h (Figure 21B). The cause of this discrepancy was probably due to some abiotic degradation in the MeHg stock bottle used to inoculate cultures even though the bottles were covered to avoid light-mediated degradation. However, we were able to detect a great capacity to remove Hg by strain ISS312. Hence, in samples taken at 6 h and at 12 h, which corresponded to the lag phase of the growth curve, we observed that MeHg concentrations were reduced by 36 % (2.6 µM) and 72 % (1.1 µM), respectively (Figure 21B). Furthermore, at time 24 h, when virtually almost all MeHg was removed (with a concentration of only 0.07 µM, and a removal of 98.2 %) the microorganism began the exponential growth phase. After 48 h, MeHg could not be detected (Figure 21B (Table S32).

Based on the results obtained, we assumed that most of the MeHg was degraded biotically by the *mer* operon encoded in ISS312, although a fraction could also be abiotically removed. In order to check which proportion of MeHg was either biotically or abiotically removed, we took additional samples at times 0 h and 72 h from the strain culture and from different control treatments (a killed control and medium alone) in the presence of MeHg. As expected, in the culture we did not observe MeHg remaining at 72 h. However, we detected certain level of abiotic degradation in the medium and killed controls (Table S33). We found that MeHg concentration was reduced by 25 % in the absence of bacteria, suggesting that three quarters were removed biotically while the rest could be degraded by abiotic processes. Still, most part of the MeHg transformation to Hg^{+2} and then to volatile Hg^{0} is caused by ISS312 strain thanks to *merA* and *merB* genes.

Biogeography of heterotrophic isolated bacteria harboring *merAB* genes

The strains screened for *merAB* genes were isolated from different oceanic regions such as the North Western Mediterranean Sea (92), South (101) and North Atlantic (42), Indian (44), Arctic (7) and Southern (7) Oceans and included isolates from photic (162) and aphotic (131) layers (Table 4). We obtained positive strains (*merA* and/or *merAB*) from the different depths and oceanographic regions, except for the Arctic Ocean (Figure 22 and Table S34). For *Alteromonas* sp., station (ST) 32 from the bathypelagic in the South Atlantic Ocean and ST151 from surface of the North Atlantic were the ones with a greater number of positive strains (Figure 22), while ST76 and ST67 from surface South Atlantic ocean followed by ST8 of the bathypelagic NW Mediterranean were the ones which presented more positive *Marinobacter* strains (Figure 22). However, the highest proportion of total positive strains was found in waters of the South Atlantic with a 47.5 % (48 out of 101) and in the Southern Ocean with a 71 % (5 out of 7). Regarding the presence of the *merAB* genes across the ocean layers, isolates came from both depths and 27 % and 23% of the total surface and the deep ocean isolates screened, respectively, gave positive results. Even though, no significative differences were found between oceans neither between depths (ANOVA, P-value >0.05). Furthermore, despite finding different proportions of positive isolates per oceanographic region and depth, we observed that isolates coming from different samples

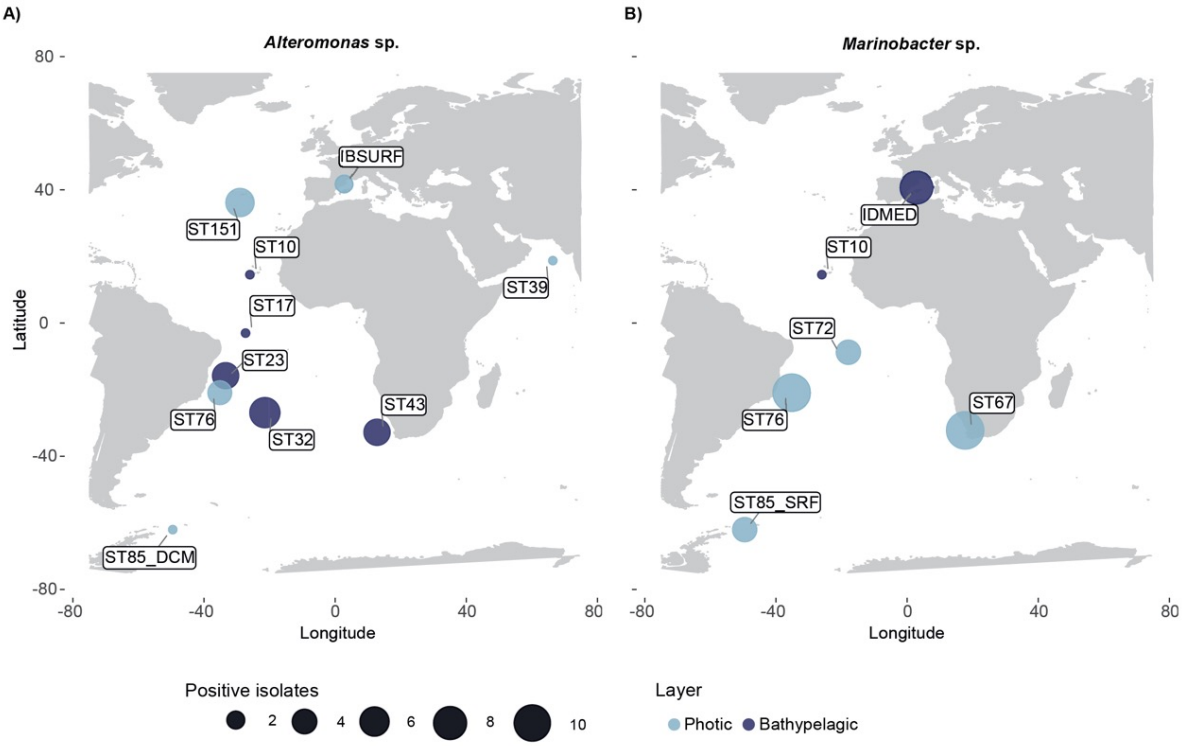

Figure 22: Distribution of the PCR positive strains for *merA* and/or *merAB* genes. **(A)** *Alteromonas* sp. strains **(B)** *Marinobacter* sp. strains. Size of the dots indicates how many strains per station including *merA* and/or *merAB* were identified. Color of the dots indicated the layer (photic or bathypelagic) where the strains were retrieved.

clustered together based on the 16S rRNA gene phylogeny, regardless the presence of *merAB* genes (Figure 18).

Biogeographic distribution of the cultured ISS312 isolate and uncultured *Alteromonas* and *Marinobacter* genomes with *merAB* genes in the bathypelagic ocean

The biogeographic and size fraction distribution of ISS312 genome was assessed in all available bathypelagic metagenomes of the Malaspina Expedition since isolate ISS312 was originally retrieved from bathypelagic waters of the South Atlantic Ocean. These bathypelagic metagenomes allowed also the reconstruction of 76 uncultured microbial genomes referred as metagenome-assembled genomes (MAGs) (Acinas et al., 2019). Out of the 76 MAGs reconstructed, one of them was almost identical to our isolate ISS312 as it showed 99.34 % of average nucleotide identity (ANI), and 99 % of the MAG aligned with the genome of the isolate (Figure S20). Hence, biogeographic and size fraction distribution was also assessed for this MAG (MAG-0289). Interestingly, we found that both genomes, affiliating to *Alteromonas mediterranea* species, were distributed across all the temperate bathypelagic waters, including the Atlantic, the Pacific and the Indian Ocean (Figure 23). Their abundances, according to the data from the fragment recruitment analyses (FRA), varied across ocean basins and we found significative differences between the Pacific and the Brazil basins for ISS312 (P-value= 0.019), and between the Pacific and the Canary basins for ISS312 and MAG-0289 (P-value: 0.011, 0.0024, respectively), suggesting a higher abundance of these bacteria in the Atlantic Ocean. Despite finding these differences between oceans,

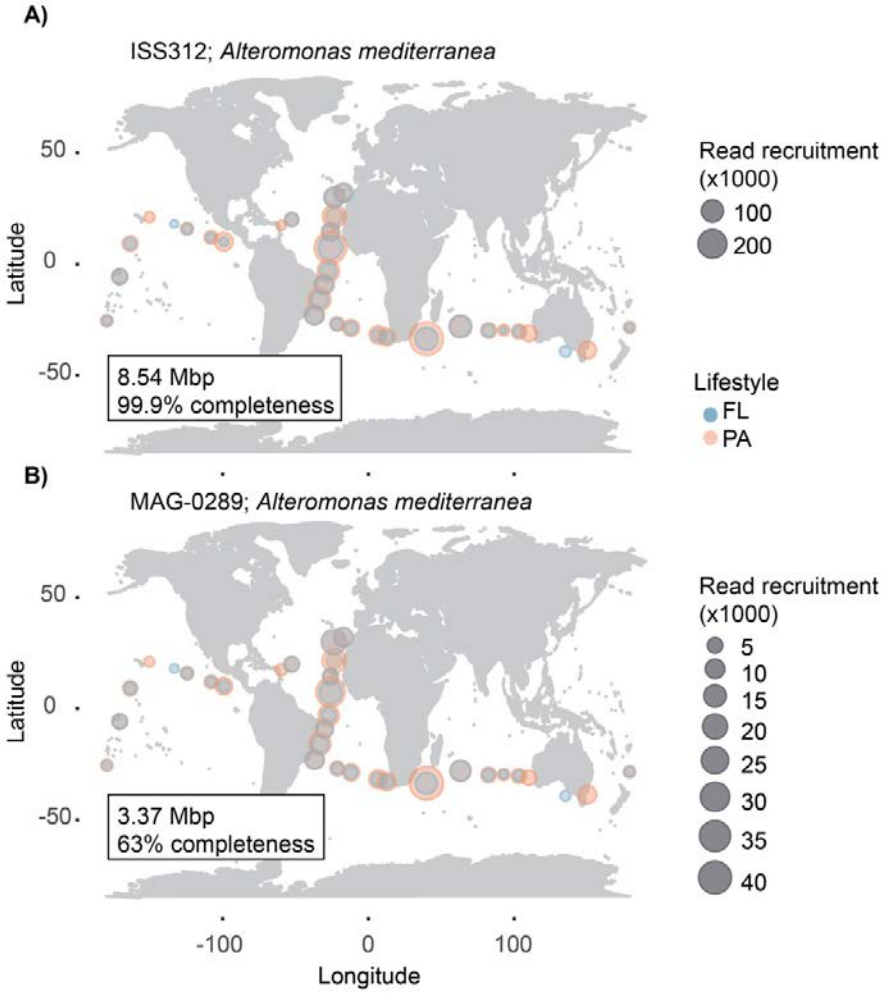

Figure 23: World map showing the distribution of *Alteromonas* ISS312 and MAG-0289 in bathypelagic marine metagenomes. **(A)** *Alteromonas mediterranea* ISS312, **(B)** the metagenome-assembled genome MAG-0289. Size of the dots indicate number of reads (x1000) and color indicate if the reads recruited in the free-living (FL) or in the particle-attached (PA) bacterial communities of the bathypelagic samples.

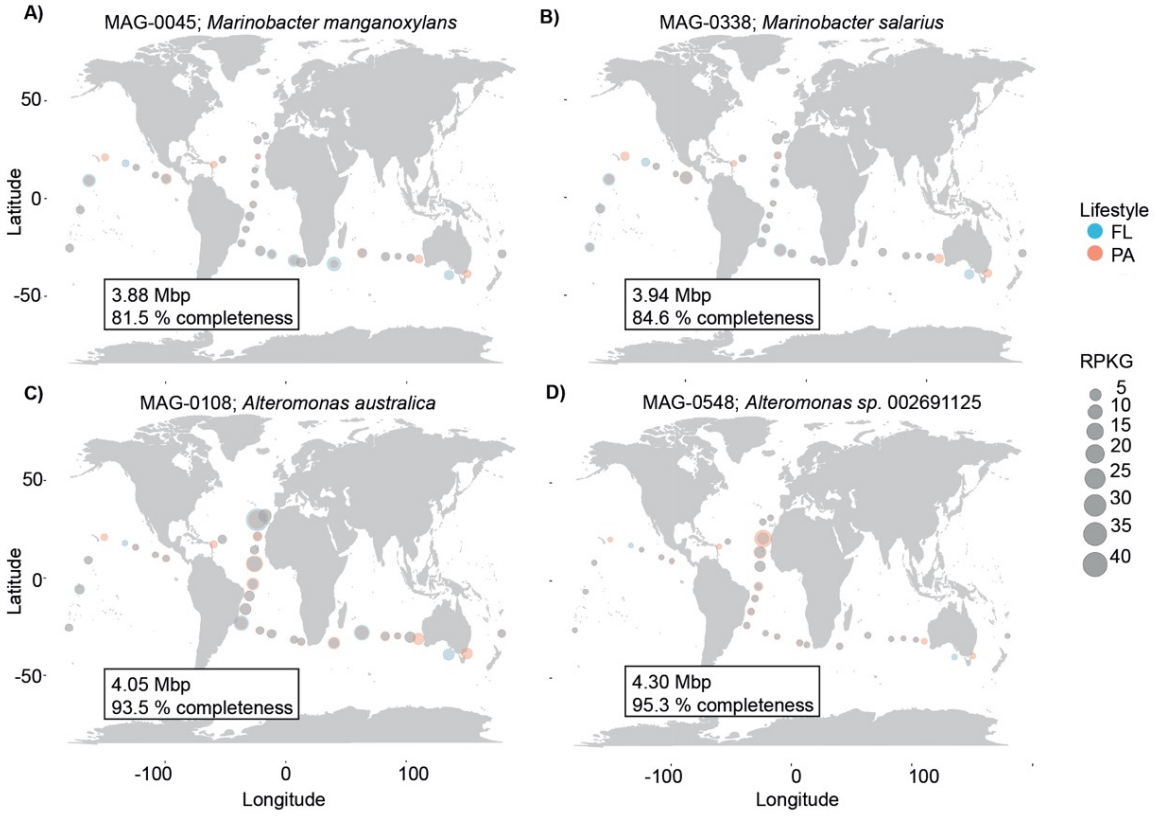

Figure 24: World map showing the distribution of *Alteromonas* and *Marinobacter* MAGs in bathypelagic marine metagenomes. **(A)** MAG-0045, **(B)** MAG-0338, **(C)** MAG-0108, **(D)** MAG-0548. Size of the dots indicate number of reads per kilobase per genome (RPKG) and color indicate if the reads recruited in the free-living (FL, 0.2-0.8 µm) or in the particle-attached (PA; 0.8-20 µm) bacterial communities of the bathypelagic samples.

we did not find significative differences between plankton size fractions, indicating that MAG-0289 and ISS312 isolate were present in both the free-living and the particle-attached bacterial communities (Figure 23). Moreover, together with MAG-0289, we could reconstruct another two MAGs affiliating with *Alteromonas* genus and two with *Marinobacter* genus that codified Hg resistance genes, including *merA* and *merB* genes (Table S35). Reconstructed *Marinobacter* MAGs (MAG-0045, MAG-0338) were evenly distributed across oceans basins, and we did not find significative differences between oceanographic regions. Contrarily, MAGs affiliating to *Alteromonas* genus (MAG-0108, MAG-0548) displayed differences in their distribution patterns in the bathypelagic ocean similarly to isolate ISS312 and MAG-0289 (Figure 24). For example, MAG-0108 presented abundance differences between the Pacific, Brazil and Canary basins (all P-values = 0.012), while MAG-0548 was found at different proportion in all basins and significative differences were found between the Canary, South-Australian, Pacific and Agulhas-Cape basins (all P-values < 0.007). Finally, equal to *Alteromonas mediterranea* ISS312 and MAG-0289, no significative differences were found between bacterial size-fractions for these four MAGs.

Discussion

Large-scale functional screening of *merAB* genes in heterotrophic marine bacteria

Hg resistance genes have been found in multiple Gram-negative and Gram-positive bacteria, and isolates containing the *mer* operon have been retrieved from many environments (Olson et al., 1979; Barkay, 1987; Simbahan et al., 2005; Rasmussen et al., 2008; Zeyaullah et al., 2009; Fashola et al., 2016; Ciok et al., 2018), including marine ecosystems such as coastal seawater and sediments (De et al., 2003), the highly contaminated Minamata Bay (Nakamura et al., 2000), hydrothermal vents (Vetriani et al., 2005), or sea-ice (Møller et al., 2014). Therefore, it was expected to found isolates harboring those genes among the MARINHET bacterial culture collection, which included strains from different depths, such as the surface, the deep chlorophyll maximum and the bathypelagic zone, as well as from diverse oceanographic regions. We focused on the detection of strains codifying the *merA* but also the *merB* gene because only those microorganisms including both genes, would have the potential ability to display a broad-spectrum resistance to Hg and would be good candidates for future bioremediation of MeHg contaminated sites. Isolates affiliating to *Alteromonas* and *Marinobacter* were selected for the screening of the targeted *mer* genes given the high number of strains present in our culture collection and the high probability to find potential candidates within these genera. Specific primers were designed for the correct amplification of *merA* and/or *merB* genes since those general primers described in the literature, mainly for *merA* (Felske et al., 2003; Oregaard and Sørensen, 2007; Sotero-Martins et al., 2008; Fong et al., 2019) did not match with the IMG/JGI public available mercuric reductases of *Alteromonas* and *Marinobacter* genera. In this study, we showed that the primers designed were able to detect at least one of the *merA* sequence variants that could be present in *Alteromonas macleodii*, *Alteromonas mediterranea*, *Alteromonas australica*, *Marinobacter hydrocarbonoclasticus* and *Marinobacter salarius*. Equally, the primers designed to detect both *merA* and *merB* genes in *Alteromonas mediterranea* and *Marinobacter hydrocarbonoclasticus* gave us positive results.

The phylogenetic tree constructed based on the partial 16S rRNA sequences showed that positive strains for *merA* and *merB* genes clustered together with strains which do not harbor these genes (Figure 18). It has been described that some bacterial species codify different sequence variants of the *merA* gene (Boyd and Barkay, 2012; Harada et al., 2019) including the *Alteromonas* genus (Ivars-Martinez et al., 2008). Therefore, it is possible that some of the *Alteromonas* and *Marinobacter* strains within the MARINHET collection tested presented different sequence variants, although we were only able to detect the one targeted by the primers designed. On the other hand, the operon *mer* can be either codified in the chromosome (Math et al., 2012) or in plasmids (Griffin et al., 1987; Barkay et al., 2003), and usually, *mer* genes are components of transposons (Mindlin et al., 2001), and integrons (Osborn et al., 1997; Bass et al., 1999). Thus, it is not surprising to find some strains within the same species without the *mer* operon. Interestingly, the *merA* amino acid phylogenies showed different results than the 16S rRNA phylogeny, wherein *Alteromonas merA* amino acid sequences mostly clustered together with *Alteromonas mediterranea*, suggesting that our primers retrieved very similar sequence variants from different *Alteromonas* species. Contrarily, more variability was found between the *merA* amino acid sequence detected in *Marinobacter* strains, as they clustered into different groups. Despite the fact that our designed primers may be not covering all *merA*/*merB* gene variants, we have obtained an important percentage (25 %) of strains codifying at least one of the addressed genes from different depths and

oceanographic regions. In general, our results showed that the *merB* gene was found in lower proportion than *merA* among both *Alteromonas* and *Marinobacter* strains, being these results consistent with the known fact that the *mer* operon does not always codify for the organomercurial lyase necessary for the detoxification of organic Hg compounds (Barkay et al., 2003; Boyd and Barkay, 2012). In addition, studies targeting the abundance of *mer* genes in the environment reported that *merA* is found to be widely distributed among marine bacteria (Mathema et al., 2011; Boyd and Barkay, 2012; Bowman et al., 2020), while *merB* only has an ecological significance in determined systems where MeHg is present at high concentrations (Schaefer et al., 2004).

Regarding the origin of the Hg resistant isolates, we found a higher proportion of positive strains coming from waters of the Southern Ocean (71 % despite the lower number of strains tested) and the South Atlantic Ocean (48 %), followed by those retrieved from the North Atlantic (17 %) and NW Mediterranean Sea (13 %). However, total Hg concentrations have been recorded to be higher in the Mediterranean and the North Atlantic Ocean compared to the South Atlantic and Southern Oceans, where concentrations were lower (Gworek et al., 2016). Nevertheless, some studies have highlighted the important but also variable levels of MeHg found in Southern and other polar waters compared to open ocean (Cossa et al., 2011; Wang et al., 2012), and isolation of Hg resistance bacteria from those polar waters has been previously reported (Ciok et al., 2018). However, it must be noticed that our results only represent a minor fraction of all potential isolates that may harbor *merAB* genes, since our primers only targeted *Alteromonas* and *Marinobacter* genera, and only specific *merA* gene variants. Thus, correlation between levels of Hg and positive strains cannot be properly done. Nevertheless, some patterns interestingly emerged, such as the presence of both genes in surface and bathypelagic isolates, indicating that these genes may be distributed along different water depths.

Furthermore, biogeographic analysis with *Alteromonas mediterranea* ISS312 genome and the five MAGs harboring *merA* or *merAB* genes assigned to *Alteromonas* and *Marinobacter* (Figure 24), confirmed the wide distribution of these genera, carrying Hg resistance genes across the global bathypelagic ocean, and present in both plankton size fractions analyzed, in the free-living (0.2-0.8 µm) and the particle-attached bacterial community (0.8-20 µm). In fact, *Alteromonas* and *Marinobacter* genera have been described to be able to live associated to particulate organic matter (Ivars-Martinez et al., 2008; Mestre et al., 2018), which forms in surface layers and sinks into the deep ocean. Besides, *merA* and *merB* genes affiliating to *Alteromonas* and *Marinobacter* have been already detected among the microbial communities associated to sinking particles (Fontanez et al., 2015). The enrichment of metal resistance genes in particles would not be surprising as it is known that they are able to concentrate heavy metals (Puig et al., 1999). In addition, MeHg production, presenting its maximum concentrations in deeper waters (200-1000 m depth) (Cossa et al., 2009; Sunderland et al., 2009; Mason et al., 2012) usually near the thermocline, is linked to microbial remineralization of particulate organic matter (Cossa et al., 2009; Sunderland et al., 2009; Heimbürger et al., 2010; Blum et al., 2013; Lamborg et al., 2016). Thereby, it is reasonable to think that bacteria living on those environments may need mechanisms to cope promptly with high concentrations of Hg and that they are present in different depths along the water column.

Towards understanding the tolerance for mercury compounds in *Alteromonas* and *Marinobacter* genera

MIC values heterogeneity within strains belonging to the same OTU (at 99 % similarity), suggested that the level of Hg resistance was isolate specific, and that maybe we retrieved different ecotypes within a same species with different tolerances to Hg. As we mentioned, *mer* genes are commonly found in a wide range of bacteria often encoded in plasmids or associated to transposable elements (Ivars-Martinez et al., 2008). It is therefore likely that the strain-specific Hg resistance patterns of our *Alteromonas* and *Marinobacter* isolates may be the result of horizontal transfer of the Hg resistance genes. This is further supported by the fact that no correlation was observed between the ability to reduce Hg and the taxonomic groups observed in the phylogenetic trees, although negative results should be interpreted with caution. Nevertheless, similar results were obtained in previous studies indicating the isolate-specific resistance to Hg (Nakamura et al., 2001; Møller et al., 2011). Despite these differences between strains, the tolerances found for inorganic Hg were similar to those found in other studies where *Alteromonas* (Morishita et al., 2006; Chiu et al., 2007; Ivars-Martinez et al., 2008; Math et al., 2012) and *Marinobacter* (Vetriani et al., 2005) genera were also isolated. However, to the best of our knowledge, this is the first study addressing the tolerance of *Marinobacter* spp. and *Alteromonas mediterranea* to MeHg. Hence, we found out that a strain affiliating to *Alteromonas mediterranea* (ISS312) presented a MIC to inorganic Hg higher than other strains already published, up to 70 µM, but we also determined that it was able to grow in the presence of MeHg, presenting a MIC up to 10 µM. Moreover, it is noteworthy that the tolerant strains to $HgCl_2$ and/or to MeHg were resistant to much higher concentrations of Hg than those reported in different oceans, which usually range from <0.1 to 10 pM (Lamborg et al., 2014b; Gworek et al., 2016).

Further characterization of ISS312 strain was done through growth curves experiments. The longer lag phase observed in the isolate growing in a medium containing 4 µM of MeHg seems to be a common trait for Hg resistant strains in the presence of the toxic compound, as this behavior has been repeatedly observed in different species of *Pseudomonas* sp., *Alcaligenes* sp. or *Bacillus* sp. (De et al., 2003; De and Ramaiah, 2007; Zheng et al., 2018). Similarly, once this lag phase was surpassed, growth rates of the resistant strains followed the same trend as those observed in growth curves without the toxic compound, an observation recurrently reported (Robinson and Tuovinen, 1984; De and Ramaiah, 2007). Strikingly, it is during the lag phase where we observed a reduction of the MeHg concentration present in the medium up to 72 %. Therefore, it can be hypothesized that during the lag phase in presence of MeHg the *mer* operon machinery is being activated by the MerR protein involved in the regulation of the operon, and *merA* and *merB* proteins are subsequently transcribed to catalyze, first, the demethylation of MeHg to Hg^{+2} (MerB), and then its transformation to Hg^0 (*merA*), which is then volatilized. Once the levels of Hg compounds have dropped considerably, and are no longer toxic, cells can grow normally reaching standard growth rates. Finally, we observed that our strain was able to efficiently remove almost all MeHg present in the medium in approximately 24 h and that very low concentrations of inorganic Hg remained after 48 h, evidencing the high detoxification capability of strain ISS312, comparable to other Hg resistant bacteria characterized for bioremediation strategies (Saranya et al., 2017).

Conclusions

Hg resistant *Alteromonas* and *Marinobacter* strains have been detected in samples from surface and bathypelagic waters including the NW Mediterranean Sea, the North and South Atlantic, the Southern and the Indian Ocean. The highest tolerance to Hg compounds was detected in *Alteromonas mediterranea* strain ISS312, isolated from bathypelagic waters of the South Atlantic Ocean. This strain is a broad-spectrum resistant bacterium, not only able to degrade inorganic Hg (Hg^{+2}), but also organic forms of this compound (MeHg), due to the presence in their genome of both, the mercuric reductase (*merA*) and the organomercurial lyase (*merB*), needed for the complete removal of Hg species from the environment. In addition, we have revealed that this *Alteromonas mediterranea* strain is widely distributed in the global bathypelagic ocean both in the free-living and the particle-attached bacterial community. Given its higher tolerance, the growth rates observed, its efficiency in the removal of MeHg, and its global oceanic distribution, this isolate could be a promising candidate for future MeHg bioremediation studies.

General discussion

GENERAL DISCUSSION

This thesis aimed to unveil the diversity, biogeography and the contribution of heterotrophic marine culturable bacteria as part of the global microbial diversity retrieved from surface, mesopelagic and bathypelagic layers of different oceanographic regions. Moreover, I deciphered the genetic capabilities and tolerance for mercury (Hg) detoxification for some widely distributed marine heterotrophic cultured bacteria. This work has been nourished by various oceanographic expeditions, but it has been mainly possible due to the collection and analyses of samples provided by the *Tara* Oceans 2009 and the Malaspina 2010 Expeditions. In the present thesis, I used a combination of culture-dependent and culture-independent techniques to pursue the main objectives, which allowed to: (i) explore the taxonomic diversity of culturable heterotrophic bacteria and to reveal the most common distributed genera from different oceans and depths, all together confirming the presence of vertical connectivity between photic and aphotic layers, (ii) challenge the known axiom that "less than 1 % of the bacteria can be cultured" in the deep ocean, while unveiling that higher proportions of bacteria, including some abundant members, can be retrieved from the bathypelagic layers in pure culture compared to the photic realm, and (iii) detect marine bacteria with the ability to degrade different forms of Hg, as well as characterize promising strains which could serve as a starting point in the design of future bioremediation strategies. In this general discussion, I will review the importance and challenges of culture-dependent studies and its combination with data obtained from culture-independent studies taking as a reference the results obtained through the different chapters.

The importance of building a large bacterial culture collection: from ecology patters to biotechnology applications

The introduction of molecular methods, and more recently high-throughput sequencing (HTS) techniques, has enabled the scientific community to explore the microbial diversity at an unprecedented scale, enhancing our understanding of marine microbial communities, which was previously based in the knowledge obtained from traditional culture methodologies. For example, Giovannoni et al. (1990) analysed through clone libraries the bacterioplankton community of the Sargasso Sea and discovered that the most abundant oceanic bacteria belong to groups really different from the ones detected in the laboratory by isolation, revealing that the new dominant bacterial groups, at least for the photic ocean, included SAR11. Hence, in this current culture-independent era, bacterial isolation seems to have been almost neglected. However, despite the efforts trying to recover the entire bacterial community from the environment through massive sequencing of the whole DNA of a sample, we rarely retrieve all the bacterial diversity present. An example is the study of Crespo et al. (2016), where despite the high sequencing effort applied (approximately 500,000 final reads per sample) to both surface and deep seawater samples from the North Western Mediterranean in order to study the rare biosphere, around 11-16 % of the bacterial community was still not detected, and the authors predicted that in order to observe the 90 % of the total richness it would be necessary to employ four times the sequencing effort applied in its work. Indeed, several recent studies predicted oceanic prokaryote species richness to an estimate of 10^{10} different OTUs for this ecosystem (Locey and Lennon, 2016; Eguíluz et al., 2019), results that indicate that the currently estimated prokaryote OTUs for surface water (37,000 (Sunagawa

et al., 2015)) and bathypelagic waters (approximately 3,600 (Salazar et al., 2016)) are very conservative and it is probably much larger than described.

In addition, studies where isolates and HTS have been carried out in the same samples show that a high percentage of the isolates could not be found among the 16S rRNA sequences obtained (Floyd et al., 2005; Shade et al., 2012; Zeng et al., 2012; Lekunberri et al., 2014; Stefani et al., 2015; Crespo et al., 2016). Thus, it seems that bacterial isolation from natural samples is still an essential tool to understand the complete picture of the bacterial communities present in natural ecosystems.

The deep ocean, including the mesopelagic and bathypelagic layers, have been studied in the last years by HTS locally (Fuhrman and Davis, 1997; Eloe et al., 2011b; Terahara et al., 2016) and globally thanks to the *Tara* Oceans and Malaspina Expedition (Sunagawa et al., 2015; Salazar et al., 2016). Nevertheless, few studies have focused the study of these ocean layers through culture-dependent techniques (Eloe et al., 2011a; Gärtner et al., 2011; Kai et al., 2017) and to the best of our knowledge, **Chapter 1** of the present thesis is the first attempt to create a large culture collection of heterotrophic bacteria from diverse oceanographic locations, including the NW Mediterranean Sea, the Atlantic, the Indian, the Pacific, the Arctic and the Southern oceans, and covering, when possible, the surface, mesopelagic and bathypelagic layers. This culture collection, called MARINHET, represented between 56.1 % and 70.5 % of the total culturable heterotrophic species present in the different layers studied according to the Good's coverage analyses. Thanks to this culture collection, some ecology patters emerged for the targeted heterotrophic fraction among the culturable bacterial community. First, the role of the sinking particles in the vertical connectivity between the surface and the bathypelagic microbial communities identified by amplicon 16S iTAGs (Mestre et al., 2018; Ruiz-González et al., 2020) was also confirmed for the culturable community, as 58.9 % of the total isolates sequences were 100 % identical between photic and aphotic layers. Among those sequences, genera such as *Alteromonas*, *Cobetia*, *Erythrobacter*, *Leeuwenhoekiella*, *Halomonas*, *Idiomarina*, *Marinobacter* or *Mesonia* were found to be widely distributed along the different depth layers. In addition, the study also provided evidence that there are some genera, mainly copiotrophs, which are widely distributed across oceans and depths. This is the case for *Alteromonas* and *Erythrobacter* genera, which occurred in around 80 % of the stations studied and were retrieved from surface, mesopelagic and bathypelagic layers.

The MARINHET collection also represent a valuable resource for the design of a huge variety of experiments involving isolates, which could include the biochemical and physiological characterization of the strains, testing hypotheses raised from HTS data, or the enquiry of possible biotechnological applications using marine bacterial isolates. The biotechnological potential of the MARINHET cultured bacteria has been addressed in **Chapter 3** of this thesis. I focused in the detection and characterization of heterotrophic marine isolates harbouring the *merA* and *merB* genes, which are part of a Hg detoxification system. I provided evidences of the presence and biogeography of Hg resistance bacteria in different oceanographic regions from distant locations, including the NW Mediterranean Sea and the Southern and the Atlantic oceans. Moreover, I have tested for the first time the methylmercury degradation potential of *Marinobacter* spp. and *Alteromonas mediterranea*. One of the isolated strains, *Alteromonas mediterranea* ISS312, has the ability to degrade 98.2 % of a 5 µM methylmercury solution in 24 h. Thus, the study of marine bacteria for biotechnology applications such as bioremediation reinforces the importance of obtaining bacteria using culture-dependent techniques.

Combining culture-dependent and culture-independent studies at a global scale

Regarding the bacterial fraction that is captured by culture-dependent versus culture-independent techniques, it is usually known that molecular techniques retrieve abundant members of the biosphere, and particularly, HTS can both retrieve abundant and rare members depending on the sequencing effort (Sogin et al., 2006). On the other hand, culture-dependent techniques usually retrieve a small fraction of the community (Kogure et al., 1979; Staley and Konopka, 1985; Amann et al., 1995) and mainly rare members of the biosphere, which are those found in very low abundances in the environment (Pedrós-Alió, 2012; Shade et al., 2012). However, there are always exceptions, like SAR11, one of the most abundant members in oceanic waters, which has been retrieved in pure culture thanks to dilution-to-extinction experiments (Rappé et al., 2002). Therefore, combination of both culture-dependent and culture-independent studies allows having a more complete picture of the bacterial diversity found in specific habitats as mentioned above, and in this line multiple examples can be found through the literature (Harmsen et al., 1997; Huber et al., 2002; Hirayama et al., 2007; Yuan et al., 2015; Lambrechts et al., 2019). However, combining these two approaches does not guarantee neither that we will find a perfect overlap between the diversity uncovered by cultures and the one detected by HTS data. This pattern has also been reflected in **Chapter 2**, where I have reported that between 11-28 % of the isolates were not detected among the zOTUs (100 % identity) of the *Tara* Oceans and Malaspina Expedition datasets, a value that increased up to 80 % when the isolates were compared against metagenomic iTAGs (miTAGs) from surface and bathypelagic samples of the Malaspina Expedition. Therefore, there is still a fraction of the community, belonging to the rare biosphere, that it is uncovered regardless the massive DNA sequencing applied in these two large scale oceanographic circumnavigations where more than 101 millions of 16S rRNA amplicon Illumina TAGs and around 670,000 miTAGs were analysed. Nonetheless, the interesting possibility to obtain data from different methodologies such as amplicon TAGs and metagenomics enables to make comparative studies and revise the well-established paradigm on the culturability of marine bacteria. Accordingly, in **Chapter 2** of this thesis I revised, on one side, the axiom that "less than 1 % of bacteria can be cultured" together with "the great plate count anomaly" phenomenon, and on the other side, I examined if the heterotrophic bacteria retrieved by culturing belonged only to the rare biosphere as commonly stated in diverse studies (Pedrós-Alió, 2012; Shade et al., 2012) or to the abundant biosphere. The general hypothesis that I launched was that a higher proportion of bacteria could be cultured in the bathypelagic ocean using traditional isolation techniques and a standard culture medium, since previous HTS studies (Salazar et al., 2016) identified abundant and cosmopolitan OTUs of the bathypelagic ocean that were related to culture representatives in databases, and as a consequence, the great plate count anomaly maybe does not apply as strong in the bathypelagic as in the surface ocean. The results obtained in this chapter supported this hypothesis. I detected that 27.7 % of the bathypelagic 16S iTAGs were 100 % identical to the 16S rRNA isolates sequences, a percentage significantly different to what we obtained in surface layers, where only a mean percentage of 1.4 % of the reads were identical to the isolated bacteria. These percentages were greatly reduced when I compared the isolates against 16S miTAGs, and I found that 6 % of the bathypelagic sequences and 0.8 % of the surface sequences were identical to isolates, when comparisons were done at 97 % sequence similarity, and were even lower at 100 % sequence similarity. Still, these results support the idea than higher proportions of bacteria can be isolated from deeper waters as metagenomic data usually present a more accurate view of the real abundances in natural

prokaryotic communities (Steen et al., 2019). In addition, I observed in both types of HTS data that some abundant and mid-abundant members of the community were identical to isolates in the bathypelagic layer, for example the *Sulfitobacter* genus. Moreover, I noticed among the bathypelagic samples studied that a higher percentage of the total cell counts could be retrieved in pure culture, with values up to 3.5 %. The differences between the recovery rates in agar plates, even though did not suggest significant differences between layers, reveal that there is certain variability among samples from surface, mesopelagic and bathypelagic and endorses the idea that the known axiom that "less than 1 % of bacteria can be cultivated" should be not literally interpreted as previously stated in global metadata studies (Steen et al., 2019). The results obtained from this study also raised new questions: (i) why could we recover higher proportions of bacteria from bathypelagic layers?, (ii) why can we obtain abundant members of the biosphere in the deeper layers but not from surface waters? The results of **Chapter 2** are in the line to what was suggested by Herndl and Reinthaler (2013), that the bathypelagic realm is constituted mainly by slow sinking particles which are resource-rich habitats for microbes (Bochdansky et al., 2016) and therefore, favouring the growth of copiotroph bacteria that may be more prone to be isolated under laboratory conditions. These ideas could be partly supported by one of the analysis included in **Chapter 2** where 16S iTAGs extracted from vertical profiles including different size fractions had been analysed. The results indicated that the proportion of 16S iTAGs reads that were 100 % identical to our isolates was higher in larger particles for surface layers, and similar proportions were found in the different size fractions of the bathypelagic communities, recording the highest values (45 % of the reads) in the larger particles (20-200 µm) and, thus, certifying that the isolates of the MARINHET collection are mostly copiotrophs and preferred a lifestyle linked to particles.

Technical challenges and future perspectives

Culturing marine bacteria from diverse oceans and depths at high-throughput scale

Given the media and the incubation conditions I used, this study is restricted to a fraction of the heterotrophic marine bacteria, missing many other fundamental groups (such as *Cyanobacteria*), and were not the best option for the isolation of novel bacterial species or genera. Nevertheless, thanks to the great isolation effort done, a new genus (*Thalassocella blandensis* (Lucena et al., 2020a)) and a new species (*Mesonia oceanica* (Lucena et al., 2020b)) were described during this thesis thanks to the MARINHET collection. Moreover, I believe that the culture media and incubation conditions selected were the more suitable given our aim to compare bacteria from different oceans and depths obtained through a standard methodology. Still, some members of the heterotrophic bacterial community may not be easily recovered by isolation and different reasons had been previously proposed (Joint et al., 2010). First, laboratory cultures may destroy the interactions that occur between organisms in the natural environment and the fastest growing species may overwhelm those that divide very slowly, thus leading to an imbalance cell-to-cell communication. It is also possible that some cells produce inhibitory compounds resulting in the inactivation of other microbial cells in the immediate vicinity, or vice versa, some bacteria need the metabolites produced by other organisms to favour its growth (Pande and Kost, 2017). Second, bacteria may be unable to grow on the combination of nutrients provided (Tanaka et al., 2014). We have little knowledge about essential substrates or concentrations of specific metabolites that are needed for the recovery of certain groups of bacteria. In addition, there could be culturable cells that

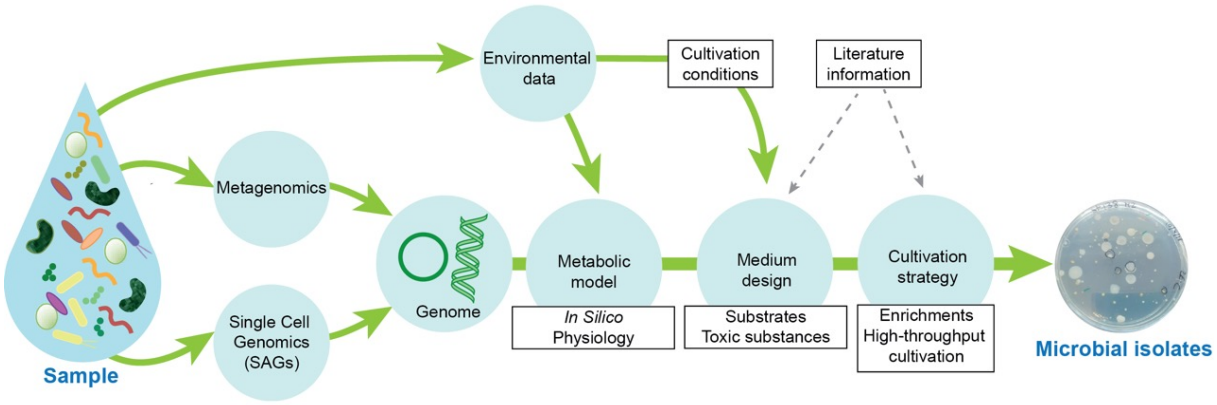

Figure 25: Proposed workflow for integrating omics data into microbial cultivation. Arrows indicate the flow of information. Modified from Gutleben et al. (2017).

fail to grow because they are in a dormancy state and the culture conditions supplied do not favour them (Deming and Baross, 2000; Buerger et al., 2012). Third, the relatively high concentrations of substrate provided in the agar plates may be toxic, particularly for those bacteria that have evolved under oligotrophic conditions such as some free-living bacteria living in open ocean waters (Schut et al., 1997; Lauro et al., 2009). Finally, the stablished incubation conditions may not favour all groups of bacteria. For example, in the bathypelagic ocean some bacteria are adapted to live at high pressure (piezophilic) and low temperatures (psychrophilic) (Bartlett, 1999; Fang et al., 2010) and the recovery of these bacteria would only be possible when the *in situ* pressure and temperatures are maintained in specific incubation chambers. However, in our study we did not maintain the *in situ* pressure, and the incubation at their *in situ* temperature was only established for agar plates of those samples collected from the bathypelagic ocean.

Future studies on this subject can be implemented by the use of other culture media (together with Marine Agar or Zobell agar) which may be, for example, more specific for microaerophilic microorganisms inhabiting in particles, or media mimicking the low concentrations of nutrients in oligotrophic ocean waters. In fact, during the realization of this thesis, I also used a modified Marine Agar 2216 medium, where disodium phosphate was autoclaved apart from the rest of the medium and added as a separate solution before solidification, since Tanaka et al. (2014) showed that toxic hydrogen peroxide could be formed when phosphate is autoclaved together with agar, contributing to growth inhibition and decreasing the number of isolates obtained from environmental samples. With this medium I tried to enhance the recovery of heterotrophic marine bacteria. Comparisons between the isolated OTUs retrieved with this medium and those isolated from Marine Agar and Zobell agar showed that a slightly higher diversity was retrieved from seawater samples when phosphate was autoclaved separately. Nevertheless, I also observed isolated OTUs that were retrieved only in Marine Agar or in Zobell agar and the statistics were not powerful enough to see significative differences between them.

Additionally, stablishing alternative isolation techniques in order to recover other fractions of the culturable microbial community could be also helpful. The metagenomic and metatranscriptomic data available from the *Tara* Oceans and Malaspina Expedition could provide essential information to design different isolation strategies for specific groups, as predictions regarding the ecology, physiology and genetic potential of individual community members are feasible with these data (Gutleben et al.,

Table 7: Summary of some alternative isolation strategies for the recovery of different groups of bacteria besides the heterotrophic community.

Modification	Alternative isolation strategy	Possibility to do it through high-throughput	Reference
Culture media			
	Oligotrophic media	Yes	Jensen *et al.*, 1996; Connon and Giovannoni, 2002
	Customized media (omics based design)	Yes	Tyson *et al.*, 2005; Wurch *et al.*, 2016
Incubation conditions			
	In situ pressure *for deep ocean bacteria*	Yes	Eloe et al., 2011a
Isolation technique			
	Dilution-to-extinction	Yes	Connon and Giovannoni, 2002; Cho and Giovannoni, 2004; Yang *et al.*, 2016
	Cultivation chips	Yes	Nichols *et al.*, 2010; Palma Esposito *et al.*, 2018
	Microfluidics	Yes	Zengler *et al.*, 2002, 2005

2017) (Figure 25). Simultaneously, I consider that the traditional isolation techniques used here are highly time consuming and requires a lot of effort, so moving forward to a more high-throughput format would be important to obtain an even higher number of isolates. Examples of these alternative isolation methodologies, which have been proven to be successful in the isolation of marine bacteria, could include dilution-to-extinction of seawater samples to isolate bacteria in small volumes of low-nutrient media (Connon and Giovannoni, 2002; Cho and Giovannoni, 2004; Yang et al., 2016) or the use of micro-droplet encapsulation in an agarose matrix, which allows the separation of single cells from the natural sample and its further cultivation under low nutrient flux conditions (Zengler et al., 2002, 2005). In Table 7 I have summarized alternative cultivation options that would help in the isolation process of different members of the marine bacterial community. I believe that the implementation of these approaches will improve our knowledge of the culturable bacteria inhabiting the deep ocean, including mesopelagic and bathypelagic waters, and may also enhance the discovery of new bacterial taxa.

The challenge of analysing large-scale marine microbial datasets

A second constraint or challenge that I could identify in this work is the use of samples from a variety of oceanographic expeditions, mainly from the *Tara* Oceans and the Malaspina Expedition, where I obtained samples not only for isolation but also for the massive sequencing of the DNA of the whole bacterial communities. Fortunately, samples for isolation of bacteria were collected, for the most part, using the same sampling protocol and were all equally preserved, which makes comparisons between samples possible. Nevertheless, the sampling protocols used for the DNA collection were slightly different between *Tara* Oceans (Pesant et al., 2015) and Malaspina Expedition (Arrieta et al., 2012). For example, the bathypelagic seawater samples from Malaspina Expedition was filtered through different pore size filters collecting then the free-living (0.2-0.8 µm) and the particle-attached (0.8-20 µm) bacterial communities separately. However, Malaspina samples from surface waters and from vertical profiles with different size fractions were collected using similar protocols as in *Tara* Oceans (0.2-3 µm and 3-20 µm). In *Tara* Oceans, despite at least 7 different plankton size fractions were available (Pesant et al., 2015), I used in this thesis only the prokaryotic fraction comprised between 0.2 µm and 1.6 µM or 3 µM, which includes the free-living bacteria and small picoeukaryotes. In Figure 26 the different size fractions used in this thesis collected in *Tara* Oceans and Malaspina Expedition are illustrated. Later on, the same methodologies were followed for DNA extraction, PCR amplification with the same set of

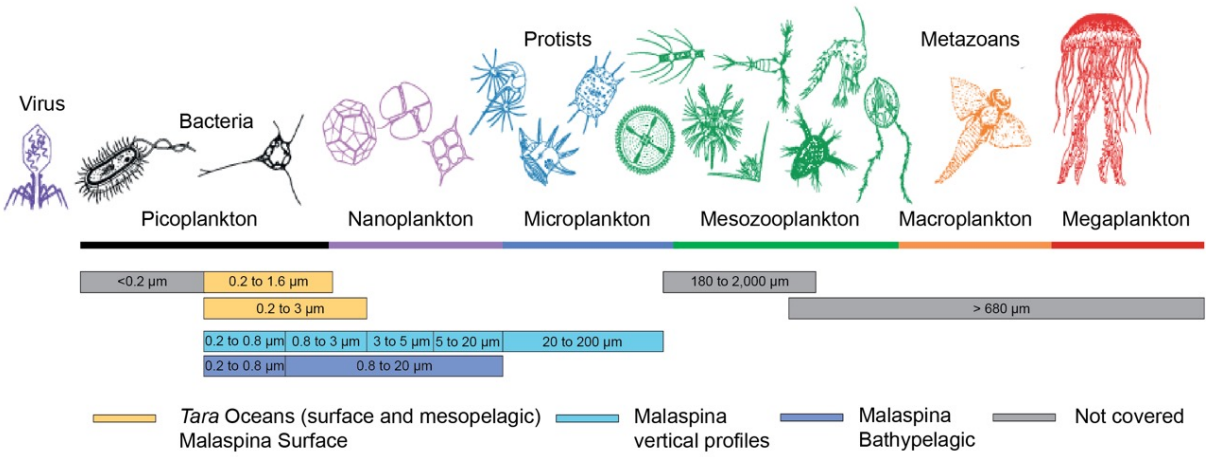

Figure 26: Size fractions samplings in *Tara* Oceans and Malaspina Expedition. Color of the rectangles indicates the circumnavigation expedition dataset that covered the different size fractions. Modified from Sunagawa et al. (2020).

primers, and sequencing of the whole DNA, even though it was performed in different facilities. Hence, our samples are for the most part comparable, but I believe that it is important to follow the exact same sampling protocols (volume, size fractions), laboratory protocols (DNA extraction, PCR amplification) and sequencing using the same platforms and approaches. Ideally, there should be consensus protocols between different research groups within the scientific community so samples from different ocean regions would be more comparable. It is true that large expeditions such as the *Tara* Oceans, which is a consortium of scientists from around the world, succeeded in elaborating consensus and standardized protocols (Pesant et al., 2015) to capture the morphological and genetic diversity of the entire plankton community from small viruses to small zooplankton. Therefore, these standardized protocols were taken as a reference for designing the sampling campaigns of Malaspina Expedition and others used in this thesis (MIFASOL, BBMO) regardless they were at a local or at a global scale. Nonetheless, future improvement and implementation would be desirable in other future expeditions.

Altogether, using samples provided from different oceanographic expeditions had also some advantages because seawater was collected during different seasons and years covering then six years of temporal range. Thus, the results obtained during this thesis, including the vertical connectivity between layers of the heterotrophic cultured bacteria, or the fact that a significant fraction of the deep ocean microbial diversity, mainly copiotrophs associated to particles, can be easily cultured, do not represent a snapshot event but a prevalent phenomenon in oceanic waters. Nonetheless, long term studies in the open ocean waters are scarce and, if initiatives such as the GO-SHIP or the OceanSITES (Boss et al., 2018) can implement standardized biology sampling protocols and can be carried out seasonally, and ideally, at a global scale, it would be a main advance to further understand population dynamics, including seasonal variations, not only in the uncultured fraction of the microbial communities (Fuhrman et al., 2015; Martin-Platero et al., 2018), but also within the culturable heterotrophic microbial communities. This acknowledge would transform our understanding of ocean biology and its impact on Earth systems.

Variability of *mer* genes and wide distribution of cultured marine Hg detoxifiers

The third challenge encountered in my work is the great variability found among the Hg resistance machinery (*mer* operon) especially in the *merA* gene, where different homologs can be found within the same genome and across bacterial lineages (Boyd and Barkay, 2012). In **Chapter 3** of this thesis, I have been able to identify 20 different bacterial taxa in the MARINHET culture collection that could be putative candidates for Hg detoxification, since the presence of *merA* and/or *merB* genes has been previously described in the literature for these isolates. Although designing a set of general primers with the capacity to detect different gene homologs from different marine bacterial taxa is desirable, it would have also implied to design highly degenerated primers with the associated risk, such as a higher probability of false positives. Therefore, I performed this study considering: (i) the high variability mentioned between *merA* homologs that led us to design *merAB* genes for specific gene variants and marine taxa, and (ii) the lack of *mer* sequences in databases, particularly those coming from bathypelagic waters, where no previous studies have focused on the detection of *mer* genes at a global scale.

Due to these two issues I decided to focus on the detection of *merA* and *merB* genes only within *Alteromonas* and *Marinobacter* genera for two main reasons: (i) they are among the most common culturable heterotrophic bacteria living in open marine waters all around the world, as they have been isolated from a wide variety of marine environments (Baumann et al., 1972; Eilers et al., 2000; Floyd et al., 2005; Gärtner et al., 2011; Handley and Lloyd, 2013; Lekunberri et al., 2014; Kai et al., 2017), and (ii) it has been already described that species of those genera harbour in their genomes the *mer* operon (Ivars-Martinez et al., 2008; Singer et al., 2011; López-Pérez et al., 2012; Handley and Lloyd, 2013). Accordingly, in **Chapter 3** I designed specific primers for those two genera and I was able to detect that 25 % of the *Alteromonas* and *Marinobacter* strains harbour *merA* and/or *merB* genes. These results confirmed that bacteria with Hg resistance genes are present in different oceans regions and depths, since positive strains were detected from the Southern and Atlantic oceans and from the NW Mediterranean Sea, as well as from surface and bathypelagic waters. The results of this distribution across oceans was later confirmed thanks to the fragment recruitment analyses (FRA) of the genome of the most tolerant *Alteromonas* strain (ISS312) and reconstructed metagenome-assembled genomes (MAGs) affiliating to *Alteromonas* and *Marinobacter* against the Malaspina bathypelagic metagenomes. Nevertheless, other taxa that codify for *mer* genes may display a different distribution that the ones observed for *Alteromonas* and *Marinobacter*. As a consequence, future studies would require the design of a more general set of primers. These primers should allow the amplification of diverse sequence variants of *merA*, which is the one that shows more variability, but also we would need primers for other *mer* genes including the *merB* gene, responsible for the demethylation of methylmercury. To do that, it would be necessary to have a more accurate database of the different gene homologs codified in the diverse *mer* operons present in the environment. Currently, a study nourished with the Malaspina bathypelagic metagenomes has developed new Hidden Markov Models (HMM) (Rabiner and Juang, 1986) in order to detect *merA* and *merB* sequence variants and aims to study the distribution of both genes across the global bathypelagic ocean (Bravo et al. In preparation). Luckily, the data derived from this study would help to have a more accurate database for *mer* genes and favour the design of different sets of primers to cover the wide diversity of marine Hg detoxifiers.

Furthermore, I believe that the study carried out in **Chapter 3** has set the basis for the design of future experiments in the Hg bioremediation field. First, I have detected a promising candidate, *Alteromonas*

mediterranea ISS312, for future bioremediation studies. Those would include the removal of Hg from contaminated marine sites or residues, using for example silica beads encapsulating the detoxifying bacteria as a biological-based filtration material (Kane et al., 2016). This type of study would previously need a further characterization of the strain in order to know its growing ability in different media, substrates, salinities or temperatures. On the other hand, the experience acquired could be of interest in other polluted environments besides the ocean waters. Marine sediments would be one of the most relevant targets for decontamination as they are typically the ultimate repository for pollutants such as Hg (Covelli et al., 2001; Orani et al., 2020) and a recent EU project, MERCLUB[13] has been funded for this purpose. The culture dependent techniques applied in this thesis could be combined with different omics approaches (metagenomics, metatranscriptomics), as well as Hg stable isotope tracer technologies (Rodríguez Martín-Doimeadios et al., 2003; Bridou et al., 2011), in order to perform a comprehensive screening of marine microorganisms with potential for Hg detoxification and establish the basis for developing a clean-up system transferable for the decontamination of impacted marine sediments.

To summarize, despite some technical and methodological challenges have been identified, the descriptive work based on culture-dependent and culture-independent methodologies applied to the study of marine samples in this thesis provides a significant advance in the knowledge of the diversity of culturable heterotrophic bacteria from less explored areas, which includes mesopelagic and bathypelagic waters, while describing some general ecology patterns. I have been able to confirm the vertical connectivity between photic and aphotic layers among the heterotrophic culturable community, or to revise long-standing observations in microbial ecology like "the great plate count anomaly" in the bathypelagic ocean. I also succeeded in reporting the functional capabilities of specific groups of bacteria focusing in the mercury detoxification processes. At the same time, some of the results obtained serve as a starting point in the design of future experimental procedures that can use marine bacteria for the decontamination of Hg impacted ecosystems.

[13]https://mer-club.eu/

Conclusions

CONCLUSIONS

1. The MARINHET bacterial culture collection allowed the description of the taxonomic, phylogenetic diversity and biogeographic distribution of the culturable heterotrophic marine bacteria from diverse oceanographic regions and depths.
2. The culturable marine heterotrophic bacteria isolated presented a general pattern of few abundant taxa and a tail of rare and low abundant isolated OTUs similarly to those HTS studies of ribosomal genes targeting whole marine microbial communities.
3. Almost 40 % of the isolates recovered were identical in their partial 16S rRNA sequences between photic, mesopelagic and bathypelagic layers, reinforcing the view of vertical connectivity between the photic and aphotic ocean through sinking particles.
4. *Alteromonas* and *Erythrobacter* genera were detected as the most abundant and commonly isolated heterotrophic bacteria using standard marine agar media and incubation conditions, as they were isolated from more than 80 % of the studied samples and from all depths studied.
5. The comparisons between plate and flow cytometry counts showed a great variability between and within ocean layers. Higher recovery rates were detected for the deeper waters (up to 3.5 %), suggesting that the repeated axiom that "less than 1 % of bacteria can be cultured" should be not taken literally, at least, in the deep ocean.
6. The combination of results from culture-dependent and -independent studies highlighted that high proportions of the bacterial community found in the deeper layers can be retrieved by isolation. Different proportions arose when the comparisons were done against 16S rRNA amplicon TAGs, were 27.7 % of the reads were 100% identical to isolates, or to metagenomic 16S iTAGs, where 0.74 % or 6.1 % of the reads were identical to isolates at 100 % or at 97 % sequence similarity, respectively. Still, most members of the bacterial communities studied, especially for the free-living bacteria, did not match with any of our isolates regardless of the depth, thus remaining part of the uncultured diversity.
7. The zOTUs and mOTUs that matched with our isolates belonged mainly to the rare biosphere in the photic layers, but to some abundant and mid-abundant heterotrophic bacteria in the deep ocean, including members of the *Sulfitobacter* genus. These results contrast with the common knowledge that abundant bacteria in natural environments are not usually isolated.
8. Isolates retrieved in this study seem to prefer a particle-attached lifestyle, which would suggest a predominant copiotrophic behavior favoring their isolation under laboratory conditions.
9. Mercury resistant *Alteromonas* and *Marinobacter* strains, harboring *merA* and *merB* genes, were detected in samples from surface and bathypelagic waters including the NW Mediterranean Sea, the North and South Atlantic, the Southern and the Indian oceans. Among those strains, variability in the tolerances to inorganic and organic mercury compounds was found even within organisms with 99 % similarity in their 16S rRNA gene.
10. The highest tolerance to mercury compounds was detected in *Alteromonas mediterranea* strain ISS312, isolated from bathypelagic waters of the South Atlantic Ocean. This strain is a broad-spectrum resistant bacterium, not only able to degrade inorganic mercury (Hg^{2+}), but also organic

forms of this compound (MeHg), due to the presence in their genome of both, the mercuric reductase (*merA*) and the organomercurial lyase (*merB*), needed for the complete removal of mercury species from the environment.

11. *Alteromonas mediterranea* ISS312 genome is widely distributed in the global bathypelagic ocean both in the free-living and particle-attached bacterial community together with metagenomes-assembled genomes (MAGs) affiliating to *Alteromonas* and *Marinobacter* genera reconstructed from Malaspina bathypelagic metagenomic data.

12. ISS312 could be a promising candidate for future bioremediation studies given the high tolerance, the growth rates observed, its efficiency in the removal of MeHg (98.2 % of 5 µm in 24h), and its global oceanic distribution, which may favor its metabolic adaptation to different environmental conditions.

Supplementary material

SUPPLEMENTARY MATERIAL

Supplementary Information

Chapter 1: Supplementary Methods

PCR conditions

Each PCR reaction with a final volume of 25 µl contained: 2 µl of template DNA, 0.5 µl of each de-oxynucleotide triphosphate at a concentration of 10 µM, 0.75 µl of MgCl2 (1.5 mM final concentration), 0.5 µl of each primer reaching a final concentration of 0.5 µM, 0.125 µl (0.025 u/µl) of Taq DNA polymerase (Invitrogen), 2.5 µl of PCR buffer supplied by the manufacturer (Invitrogen, Paisley, UK) and Milli-Q water up to the final volume. Reactions were carried out in a Biorad thermocycler using the following program: initial denaturation at 94ºC for 5 min, followed by 30 cycles of 1 min at 94ºC, 1 min at 55ºC and 2 min at 72ºC, and a final extension step of 10 min at 72ºC. The PCR products were verified and quantified by agarose gel electrophoresis with a standard low DNA mass ladder (Invitrogen).

Phylogenetic trees

Different phylogenetic trees were included in these study: (i) phylogenetic trees with all non-redundant sequences for *Alphaproteobacteria, Gammaproteobacteria, Bacteroidetes*, and Gram-positive bacteria; (ii) phylogenetic trees to support novelty of putative novel isolates (Figure 5b in Chapter 1); and (iii) one phylogenetic tree including all isolates for alpha-diversity metrics analyses (explained in Chapter 1 Methods). The total pool of sequences to be included in the first two types of phylogenies were first aligned with the SINA web alignment tool (http://www.arb-silva.de/aligner/) (Pruesse et al., 2012) and imported into the phylogenetic software MEGA 5.2.2 (Tamura et al., 2011). The phylogenetic trees were constructed with the Neighbour Joining (NJ) algorithm using the Jukes-Cantor distance and 1000 bootstrap replicates.

In order to see phylogenetic relationships and explore the connectivity between photic-layer, mesopelagic and bathypelagic isolates, as well as to detect possible novel strains, phylogenetic trees for *Alphaproteobacteria, Gammaproteobacteria, Bacteroidetes* and Gram positive bacteria (Figures S4-7) were built. They included the non-redundant sequences dataset of each station and the taxonomic affiliation of their Closest Cultured Match (CCM) and Closest Environmental Match (CEM) obtained after BLASTn search against the RDP databases (Supplementary Tables S3 and S4).

Samples used for Illumina 16S rRNA sequencing

a) Datasets

Different datasets comprising Illumina 16S rRNA gene sequences were analyzed in order to compare them with 16S rRNA sequences from isolates obtained by traditional culture techniques. The first dataset comprised a total of 124 surface samples and 41 bathypelagic samples collected during the Malaspina 2010 Expedition and distributed across the world's oceans. Surface seawater (3m) samples were collected and filtered as described previously in Ruiz González et al. (2019). For these surface samples we focused on the 0.2–3 µm fraction, which represents mostly free□living bacteria. On the other hand, bathypelagic samples (4000 m depth) were collected and filtered as described in Salazar et al. (2016). In these case two different size fractions were analyzed representing the free-living (0.2-0.8 µm) and the particle-attached (0.8-20 µm) bacterial communities. The second dataset was formed by 80 surface samples and 39 mesopelagic samples collected during the *Tara* Oceans 2009 and *Tara* Oceans Polar Circle 2013 expeditions which covered the major oceanic provinces including the polar circles. Sampling strategy and methodology are described in Pesant et al. (2015). In this dataset we only focused in the free-living bacterial communities compressed within the 0.2 and 1.6 or 3 µm. In all three datasets, once seawater was processed, filters were flash-frozen in liquid nitrogen and stored at -80 ºC until DNA extraction.

b) DNA extraction, amplification and sequencing

The DNA from the samples of the different datasets described was extracted with a phenol-chloroform protocol, as described elsewhere (Alberti et al., 2017; Massana et al., 1997; Salazar et al., 2016). Prokaryotic barcodes for each of the datasets was generated by amplifying the V4 and V5 hyper-variable regions of the 16S rRNA gene using primers 515F-Y (5'-GTG YCA GCM GCC GCG GTA A-3') and 926R (5'-CCG YCA ATT YMT TTR AGT TT-3') described in Parada et al. (2016). Sequencing was performed in an Illumina MiSeq platform (iTAGs) using 2x250 bp paired-end approach at the Research and Testing Laboratory facility (Lubbock, TX, USA) for the Malaspina datasets and at Genoscope (Evry Cedex, France) for the *Tara* Oceans and *Tara* Polar Oceans dataset.

c) 16S rRNA Illumina sequences processing

Computing analyses were run at the MARBITS bioinformatics platform at the Institute of Marine Sciences (ICM) and at the Euler scientific compute cluster of the ETH Zürich University. The obtained amplicons were processed through the bioinformatic pipeline described in the github repository https://github.com/SushiLab/Amplicon_Recipes. Briefly, pair-end reads were merged at a minimum 90% of identity alignment, and those with ≤ 1 expected errors were selected (quality filtering). Primer matching was performed with CUTADAPT v.1.9.1. Dereplication and zOTU (zero-radius OTUs) denoising at 100% similarity (UNOISE algorithm) were performed with USEARCH v.10.0.240 (Edgar, 2010). zOTUs were taxonomically annotated against the SILVA database v132 (2017) with the LCA (lowest common ancestor) approach. Finally, zOTUs were quantified to obtain zOTU-abundance tables. Non-prokaryotic zOTUs (eukaryotes, chloroplast and mitochondria) were removed, whereas singletons (zOTUs appearing only once) were maintained.

This procedure was applied individually for the: (i) 41 Malaspina bathypelagic samples, (ii) 124 Malaspina surface samples, (iii) 119 *Tara* Oceans and *Tara* Polar Oceans surface and mesopelagic samples. Hence, 3 different zOTU abundance tables were obtained after applying the pipeline. To allow comparisons between samples, each zOTU table was randomly sampled down to lowest sampling effort using the function *rrarefy.perm* with 1000 permutations from the R package *EcolUtils* (Salazar, 2018a).

Phenotypic characterization of ISS653 and ISS1889

Detailed phenotypic characterization is described in Lucena et al. (2020b) and here we focused on the description of those test were differences were found between strains. Briefly, it included morphological, cultural, biochemical, physiological and nutritional screening and was performed by already described methods (Pujalte et al., 2018). *Mesonia algae* CECT 9441T, *Salegentibacter salegens* CECT 9443T, *Gramella echinicola* CECT 9439T and *Zunongwangia profunda* CECT 9445T were characterized in parallel for comparative purposes.

Flexirubin type pigmentation was tested according to (Bernardet et al., 2002). In addition, cellulose degradation, nitrate reduction acid production from carbohydrates in API 50CH/E, APIZYM and API 20NE profile were performed as described in (Lucena et al., 2020b). Fatty acid methyl esters were extracted from strains ISS653 and ISS1889 biomass grown in Marine Agar at 26 ºC after 72 h incubation. Extracts were prepared according to standard protocols as described for the MIDI Microbial Identification System (Sasser, 1990) at the CECT. Cellular fatty acid content was analyzed by gas chromatography with an Agilent 6850 chromatographic unit, with the MIDI Microbial Identification System using the TSBA6 method and identified using the Microbial Identification Sherlock software package (MIDI, 2008).

Supplementary Figures

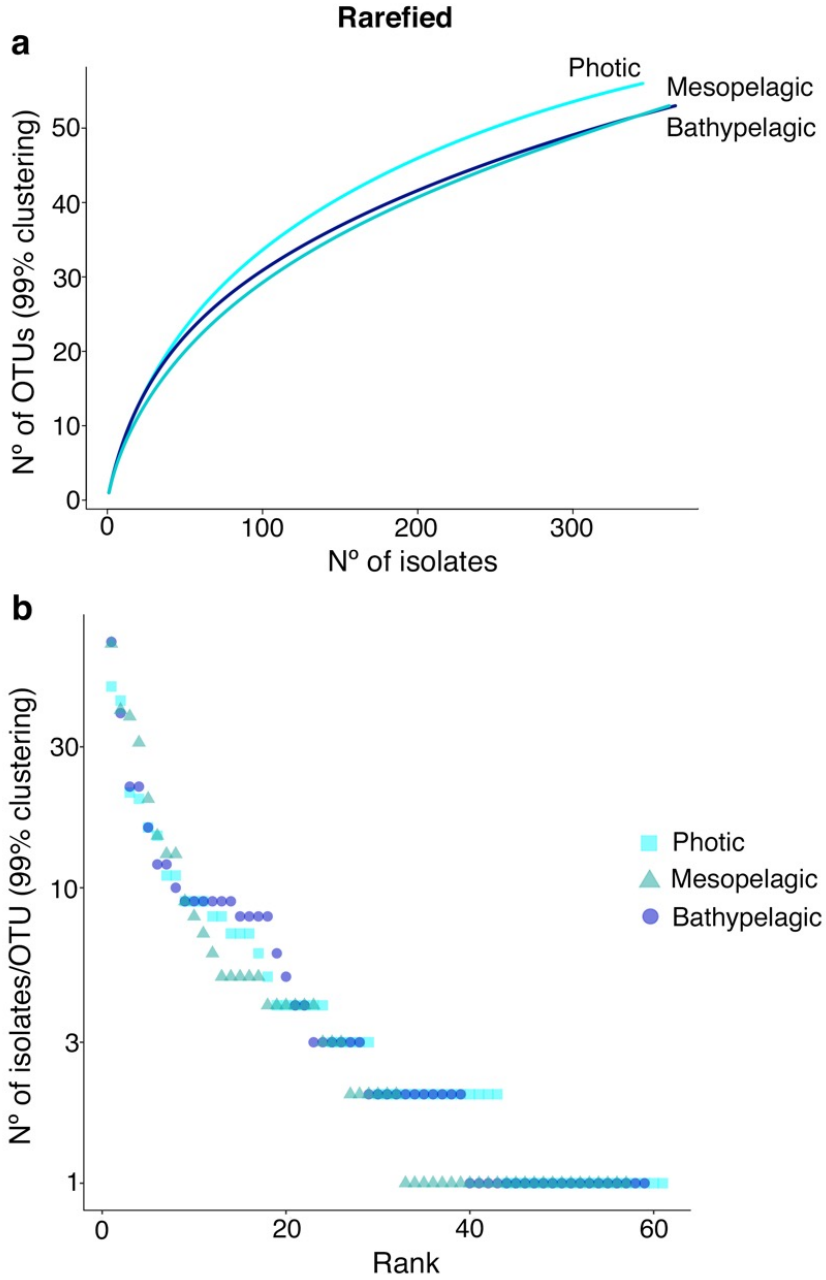

Figure S1: Rarefaction and rank abundance curves for each of the depths included in the study. **(A)** Rarefaction curves for photic, mesopelagic, and bathypelagic samples extracted from the rarefied iOTU table down to the lowest isolated layer (mesopelagic with 362 isolates). **(B)** Rank abundance plots showing the number of isolates per OTU (at 99% clustering) obtained in the three layers studied also for the rarefied iOTU table down to the lowest isolated layer. Y axis are in log10 scale.

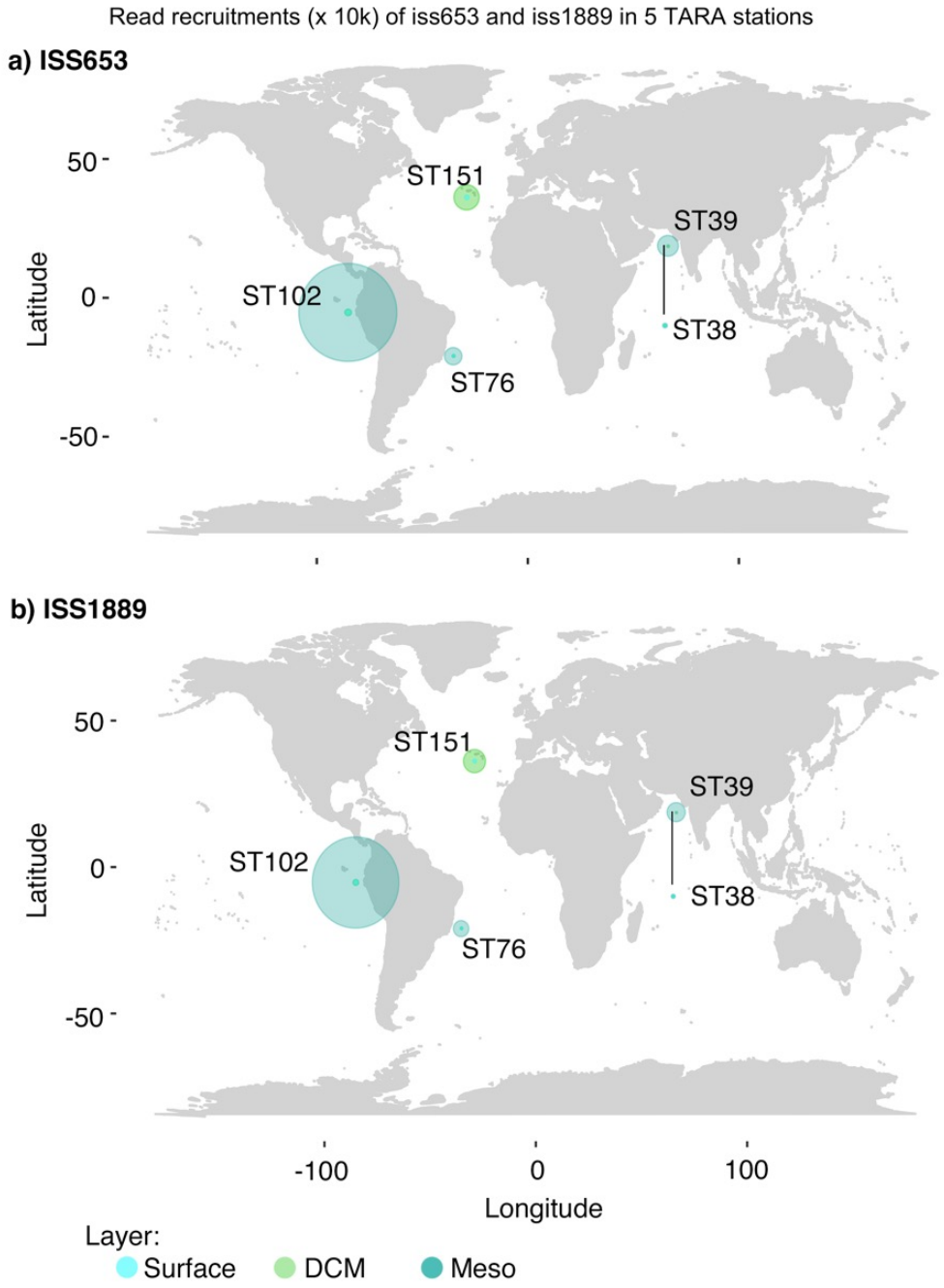

Figure S2: World map showing the read recruitment of ISS653 and ISS1889 in five *Tara* Oceans stations. They include the stations where the isolates were retrieved (ST151 and ST102) and some distant stations for the sake of comparison (ST39, ST38, ST76). ST38 is located near ST39 (Latitude 19º 2.24' N, Longitude 64º 29.24' E), but its location in the plot was slightly modified for its correct visualization. Size of the circles are the abundances of reads of each genome recruited in each station and layer (x10k). DCM, deep chlorophyll maximum; Meso, mesopelagic isolates.

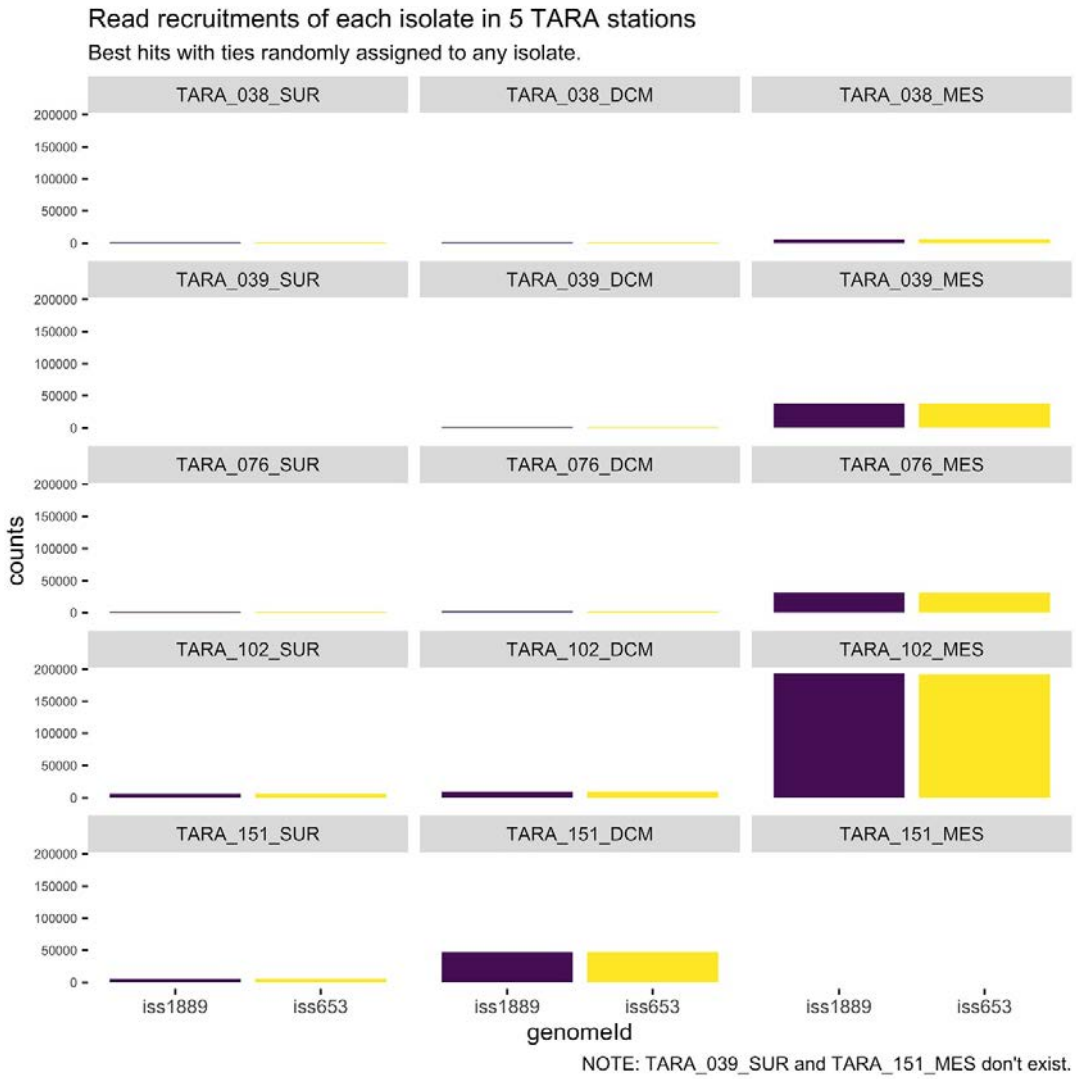

Figure S3: Read recruitment of ISS653 and ISS1889 in five *Tara* Oceans stations. The read counts recruited in each station and layer is indicated per each genome. They include the stations where the isolates were retrieved (ST151 and ST102) and some distant stations for the sake of comparison (ST39, ST38, ST76). SUR, surface isolates; DCM, deep chlorophyll maximum; Meso, mesopelagic isolates.

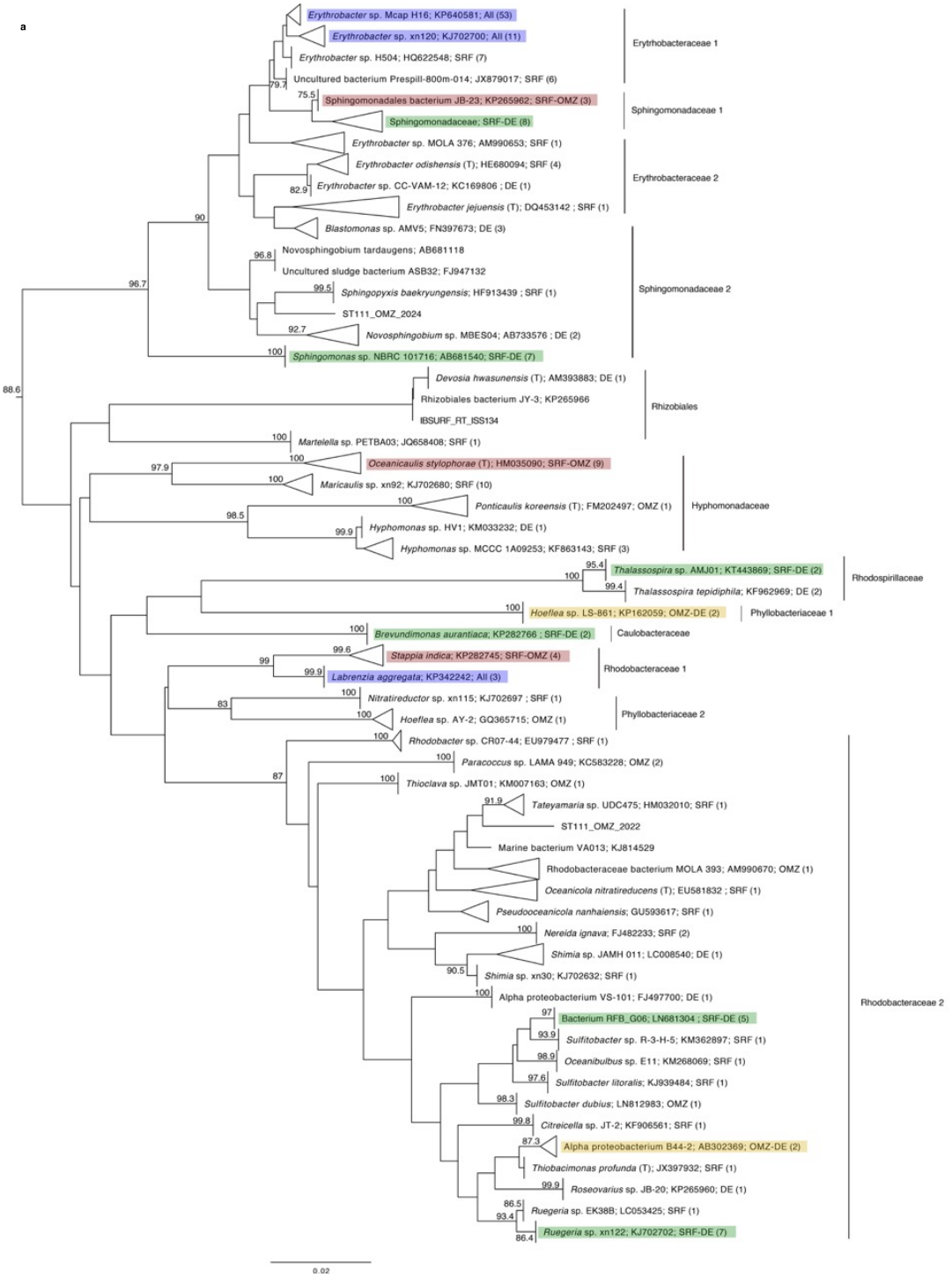

Figure S4: Phylogenetic relationships between photic-layer, mesopelagic, and bathypelagic *Alphaproteobacteria* isolates. Neighbour Joining trees of the 16S rRNA gene sequences of the reduced pool of sequences, including the non-redundant 516 isolates plus their closest cultured match (CCM) and closest uncultured or environmental match (CEM). The numbers in nodes represents bootstrap percentages > 75, calculated from 1000 replicates. The number of isolates from each specific cluster is indicated in brackets. Blue rectangles indicate group of isolates retrieved from all depths; in red; a mix between photic-layer and mesopelagic (indicated as OMZ); in green, a mix from photic-layer and bathypelagic isolates; and in yellow, a mix of mesopelagic (indicated as OMZ) and bathypelagic. DE, bathypelagic isolates; SRF, photic-layer isolates; OMZ, oxygen minimum zone isolates, which refer to the mesopelagic isolates retrieved from samples collected in areas with OMZ. The vertical lines indicate the family name or order of some groups of isolates in each tree.

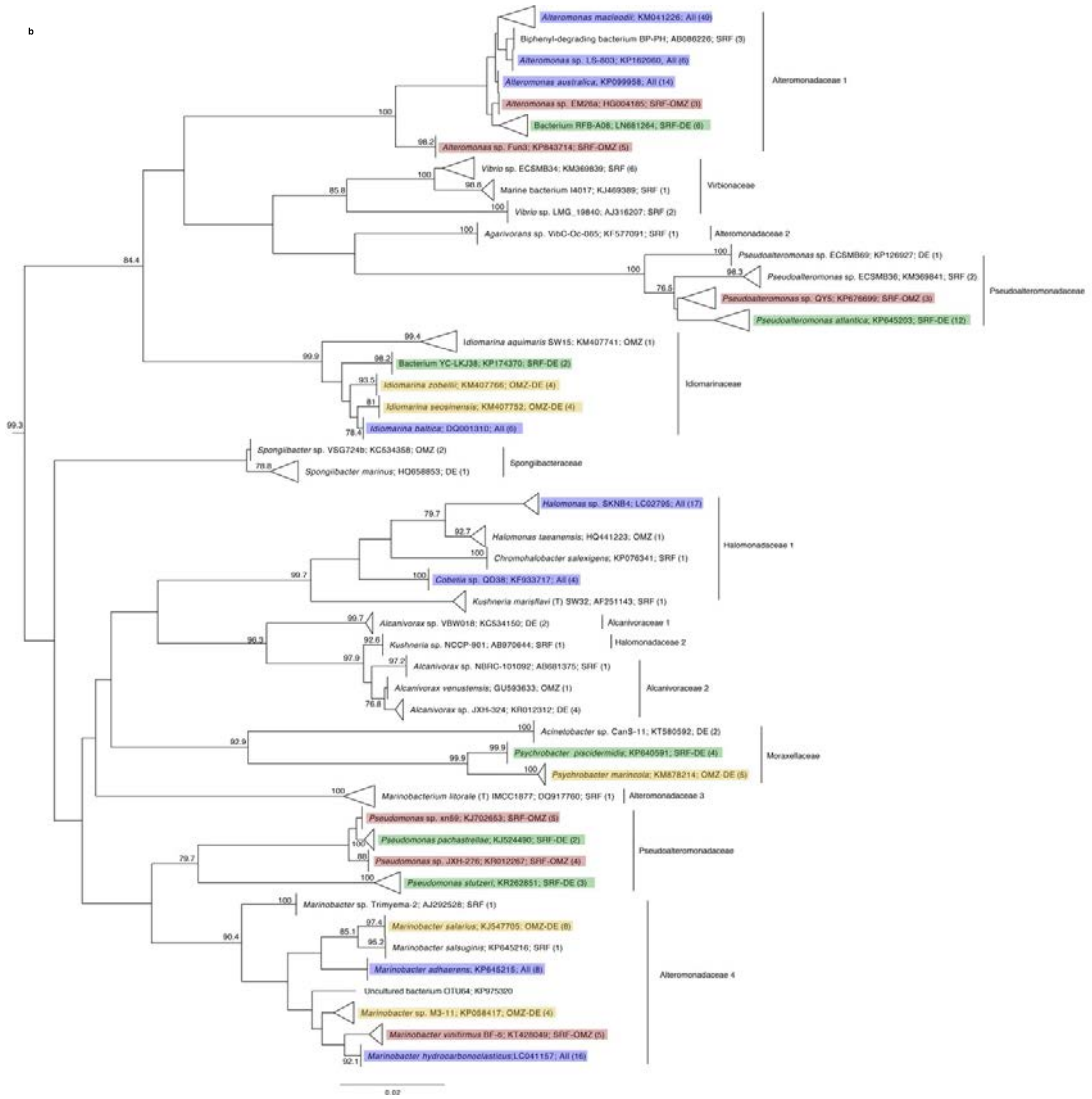

Figure S5: Phylogenetic relationships between photic-layer, mesopelagic, and bathypelagic *Gammaproteobacteria* isolates. Neighbour Joining trees of the 16S rRNA gene sequences of the reduced pool of sequences, including the non-redundant 516 isolates plus their closest cultured match (CCM) and closest uncultured or environmental match (CEM). The numbers in nodes represents bootstrap percentages > 75, calculated from 1000 replicates. The number of isolates from each specific cluster is indicated in brackets. Blue rectangles indicate group of isolates retrieved from all depths; in red; a mix between photic-layer and mesopelagic (indicated as OMZ); in green, a mix from photic-layer and bathypelagic isolates; and in yellow, a mix of mesopelagic (indicated as OMZ) and bathypelagic. DE, bathypelagic isolates; SRF, photic-layer isolates; OMZ, oxygen minimum zone isolates, which refer to the mesopelagic isolates retrieved from samples collected in areas with OMZ. The vertical lines indicate the family name or order of some groups of isolates in each tree.

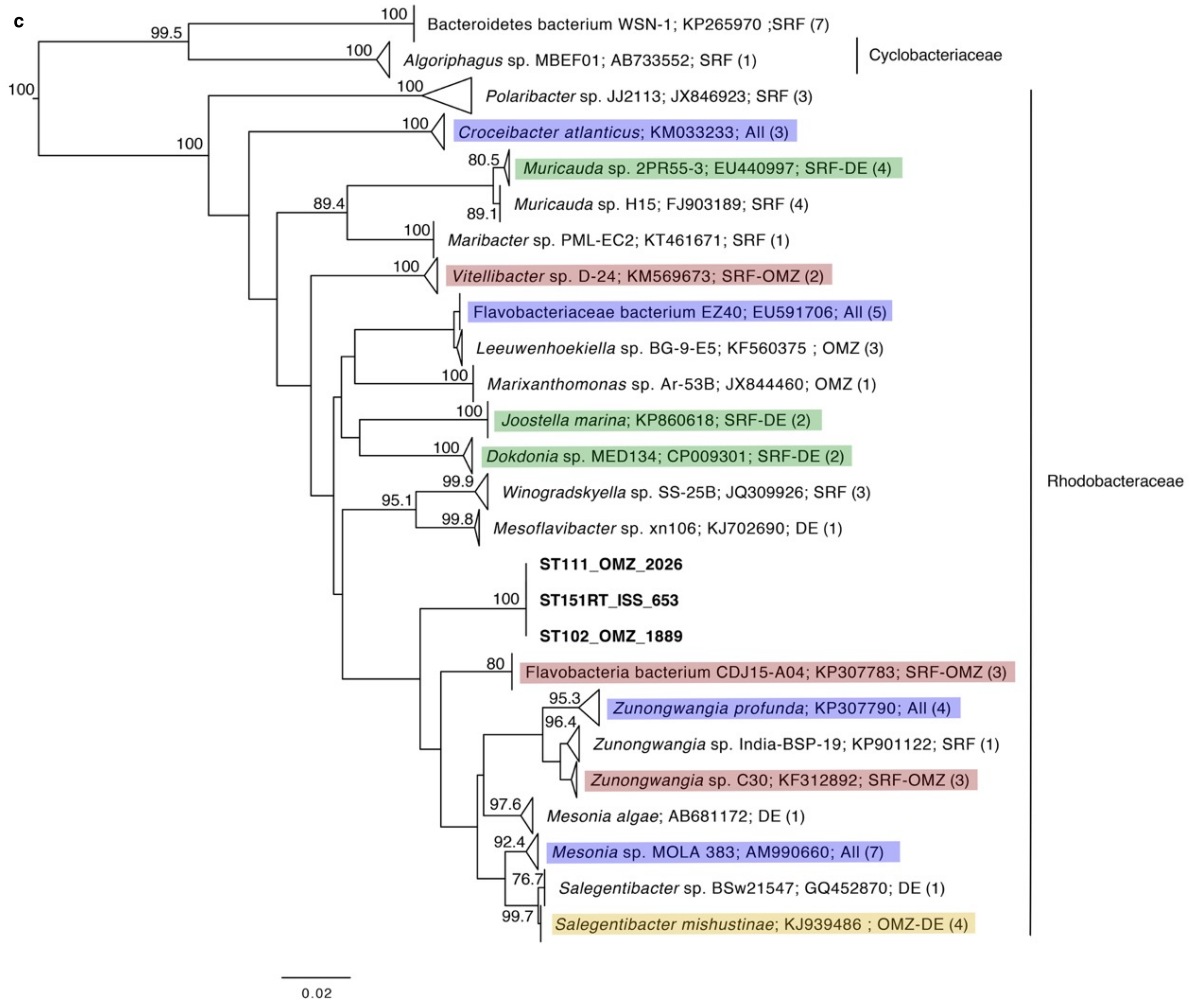

Figure S6: Phylogenetic relationships between photic-layer, mesopelagic, and bathypelagic *Bacteroidetes* isolates. Neighbour Joining trees of the 16S rRNA gene sequences of the reduced pool of sequences, including the non-redundant 516 isolates plus their closest cultured match (CCM) and closest uncultured or environmental match (CEM). The numbers in nodes represents bootstrap percentages > 75, calculated from 1000 replicates. The number of isolates from each specific cluster is indicated in brackets. Isolates in bold show the putative novel genera isolated in this study. Blue rectangles indicate group of isolates retrieved from all depths; in red; a mix between photic-layer and mesopelagic (indicated as OMZ); in green, a mix from photic-layer and bathypelagic isolates; and in yellow, a mix of mesopelagic (indicated as OMZ) and bathypelagic. DE, bathypelagic isolates; SRF, photic-layer isolates; OMZ, oxygen minimum zone isolates, which refer to the mesopelagic isolates retrieved from samples collected in areas with OMZ. The vertical lines indicate the family name or order of some groups of isolates in each tree.

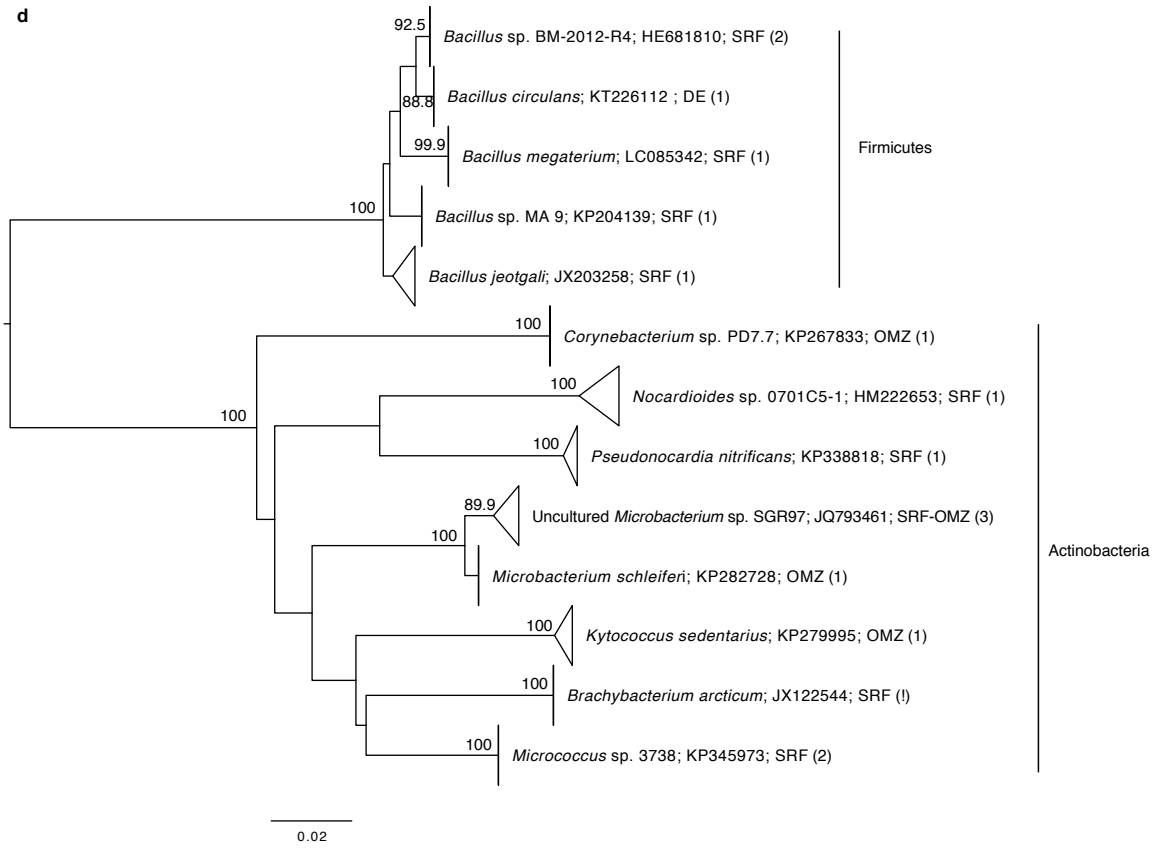

Figure S7: Phylogenetic relationships between photic-layer, mesopelagic, and bathypelagic *Gram positive bacteria* isolates. Neighbour Joining trees of the 16S rRNA gene sequences of the reduced pool of sequences, including the non-redundant 516 isolates plus their closest cultured match (CCM) and closest uncultured or environmental match (CEM). The numbers in nodes represents bootstrap percentages > 75, calculated from 1000 replicates. The number of isolates from each specific cluster is indicated in brackets. Blue rectangles indicate group of isolates retrieved from all depths; in red; a mix between photic-layer and mesopelagic (indicated as OMZ); in green, a mix from photic-layer and bathypelagic isolates; and in yellow, a mix of mesopelagic (indicated as OMZ) and bathypelagic. DE, bathypelagic isolates; SRF, photic-layer isolates; OMZ, oxygen minimum zone isolates, which refer to the mesopelagic isolates retrieved from samples collected in areas with OMZ. The vertical lines indicate the family name or order of some groups of isolates in each tree.

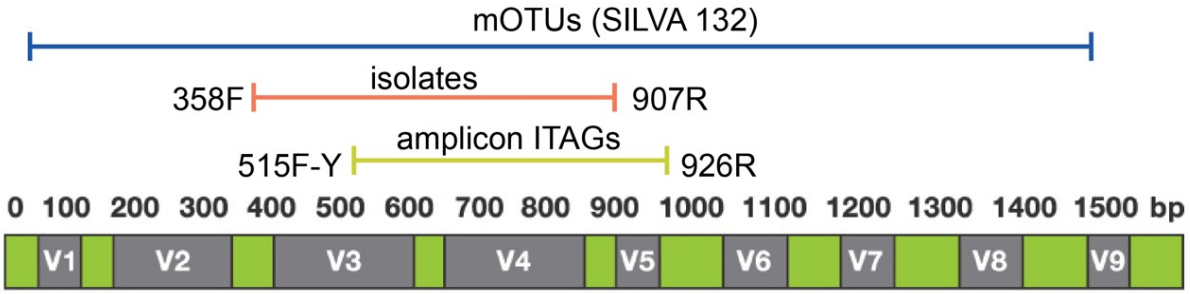

Figure S8: Graphic representation of the 16S rRNA gene and the regions covered by isolates (orange), amplicon iTAGs (yellow) and mOTUs extracted from the SILVA database v.132 (dark-blue). Forward and reverse primers used for 16S rRNA amplifications in isolates and amplicon TAGs are indicated at the left and right sides of the lines.

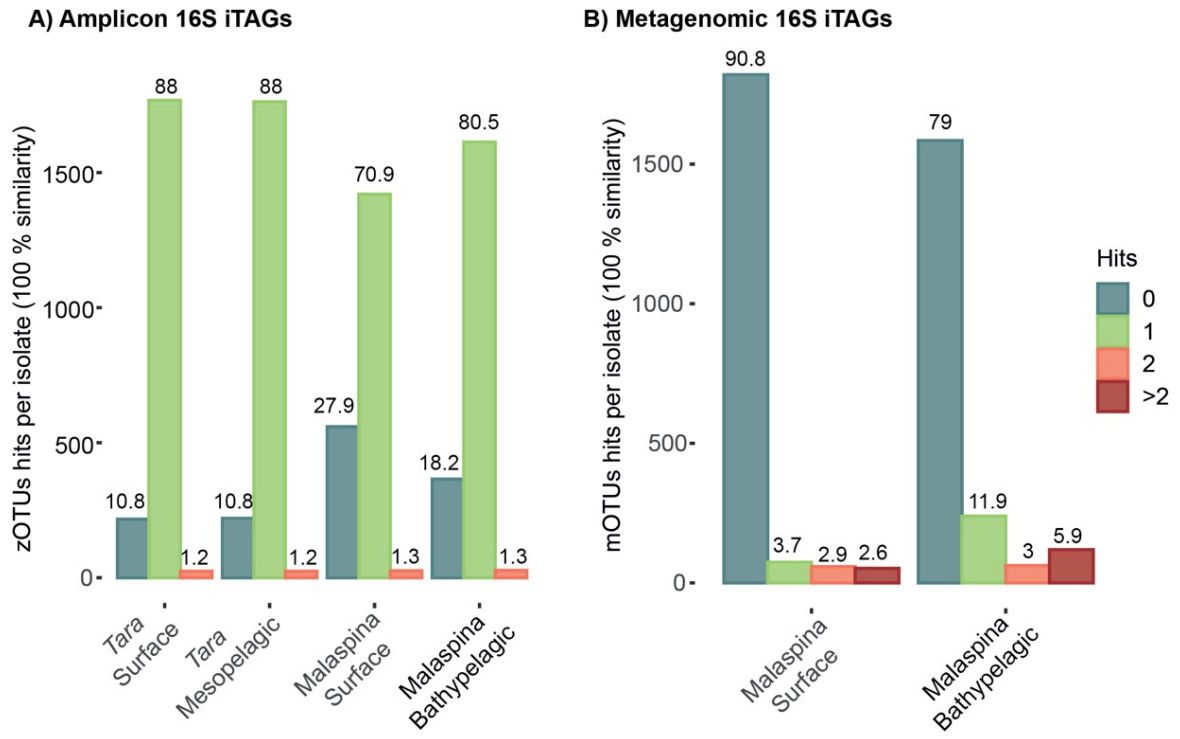

Figure S9: Number of isolates hits possible to **(A)** zOTUs and to **(B)** mOTUs at 100% sequence similarity using *usearch_global*. Percentages from total isolates in each category are indicated above each bar.

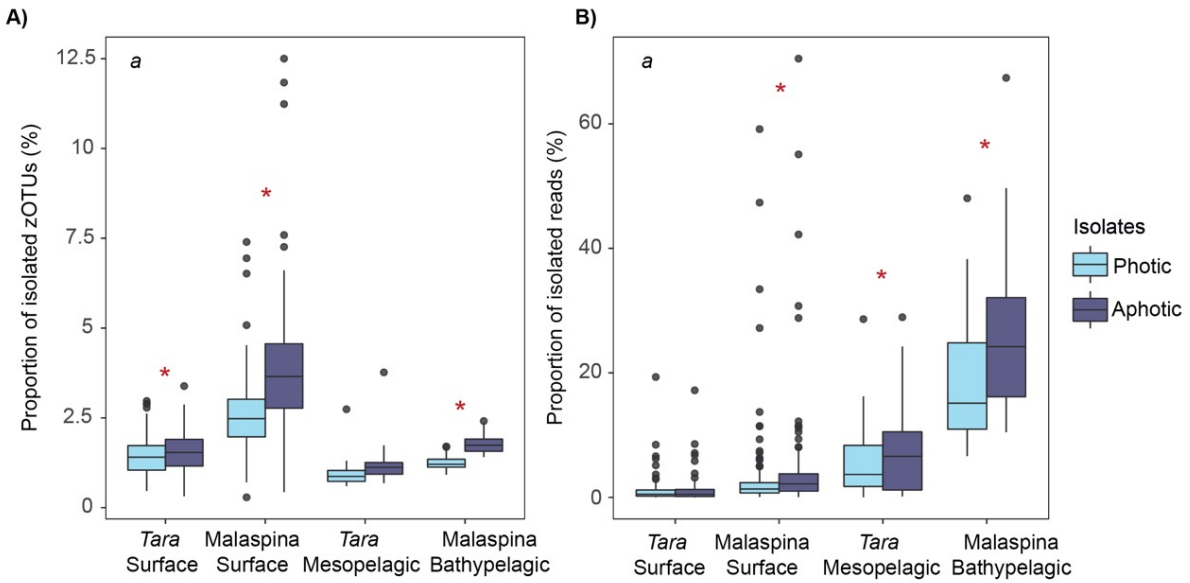

Figure S10: Differences between datasets when comparing photic and aphotic isolates separately. **(A)** Proportion of isolated zOTUs. **(B)** Proportion of reads (16S iTAGs) recruited by isolates. Comparisons done at 100% sequence similarity. Significative differences between datasets (P-value < 0.01) are indicated by an italic a in the top left corner, while significative differences (P-value < 0.01) within a dataset between photic and aphotic samples are indicated by asterisks.

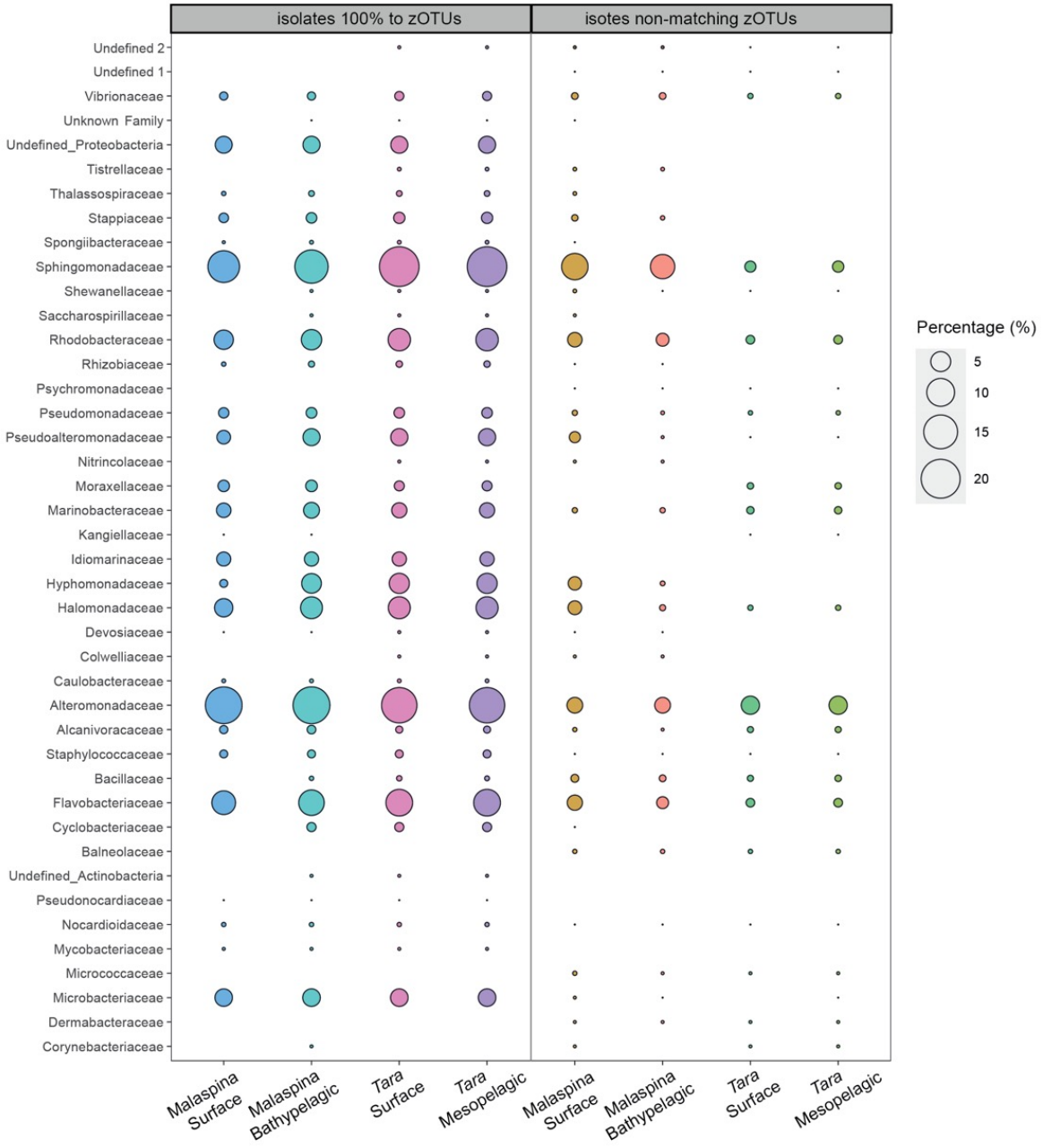

Figure S11: Taxonomic differences at the family level between those isolates that were 100% identical to zOTUs in Malaspina Surface, Malaspina Bathypelagic, *Tara* Surface and *Tara* Mesopelagic datasets. The first four columns of points represent those isolates 100% identical to zOTUs, and the last four columns of points represent those isolates that did not match with any zOTU. Size of the dots indicates the percentage of isolates in each family from the total isolates of the MARINHET_v2 collection.

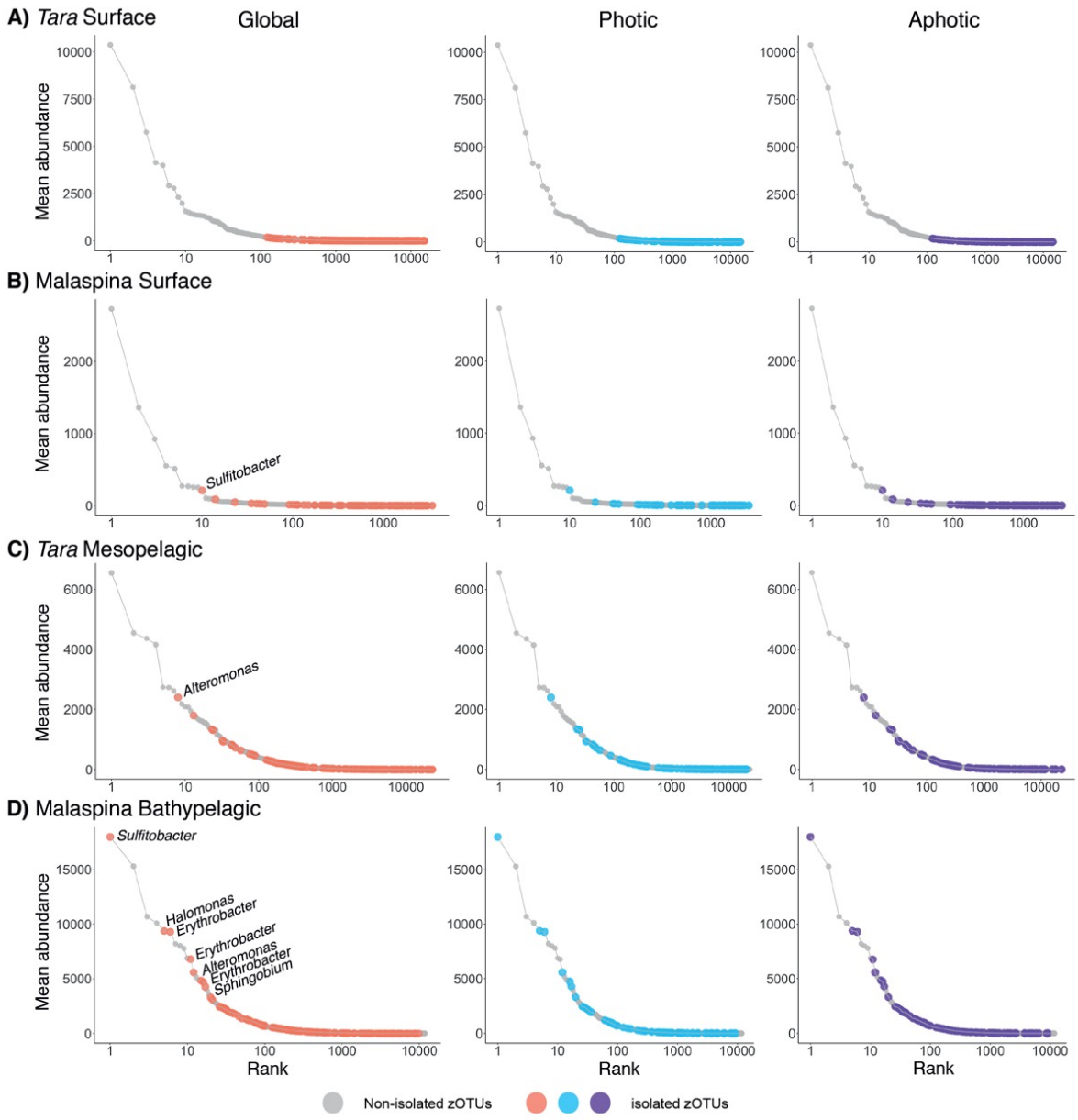

Figure S12: Rank plots showing the identified zOTUs from amplicon 16S iTAGs recruited by photic and aphotic isolates separately in the different datasets. **(A)** *Tara* Surface. **(B)** Malaspina Surface. **(C)** *Tara* Mesopelagic. **(D)** Malaspina Bathypelagic. Colour of the dots indicates which zOTUs are 100% identical to at least one isolate in the comparisons made with all photic and aphotic isolates (global), and separately with the photic and aphotic datasets: grey, non-isolated zOTUs; orange, blue or purple, isolated zOTUs. Taxonomic affiliation is indicated for the abundant isolated zOTUs (>1% abundance) in the global rank plot.

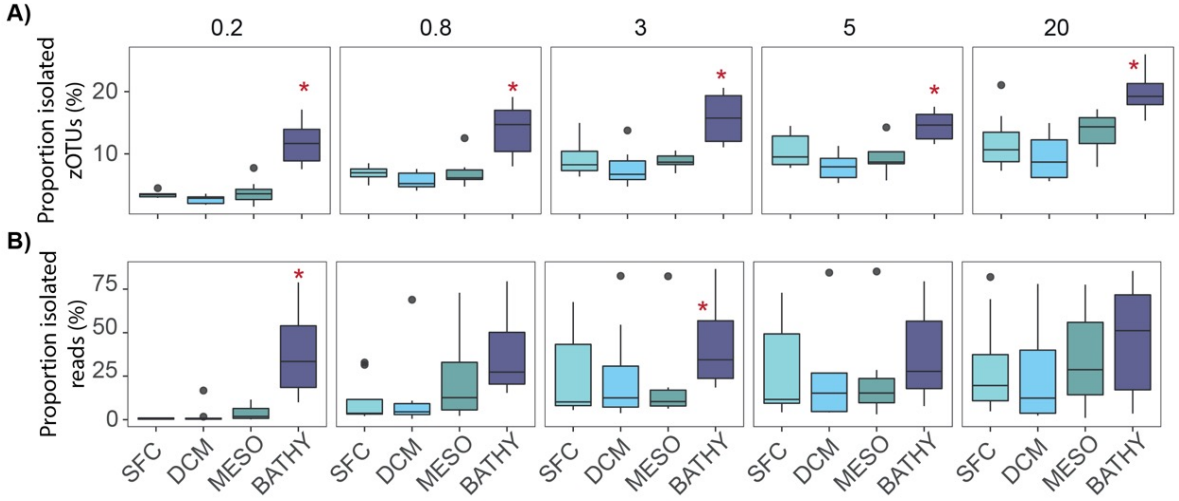

Figure S13: Boxplots from Malaspina profiles with different size fraction showing: **(A)** the percentages of isolated zOTUs that were 100% identical to at least one isolate per size fractions and depth, and **(B)** the percentages of reads that were 100% identical to at least one isolate per size fractions and depth. SFC: surface, DCM: deep chlorophyll maximum, MESO: mesopelagic, and BATHY: bathypelagic. 0.2: free-living bacteria, 0.8: bacteria attached to small particles, and 3-20: bacteria attached to big particles. Sizes of the fractions are in □m. Significative differences between layers (Kruskal-Wallis, P.value from $1.1x10^{-8}$ to $4.7x10^{-12}$) are indicated by red asterisks.

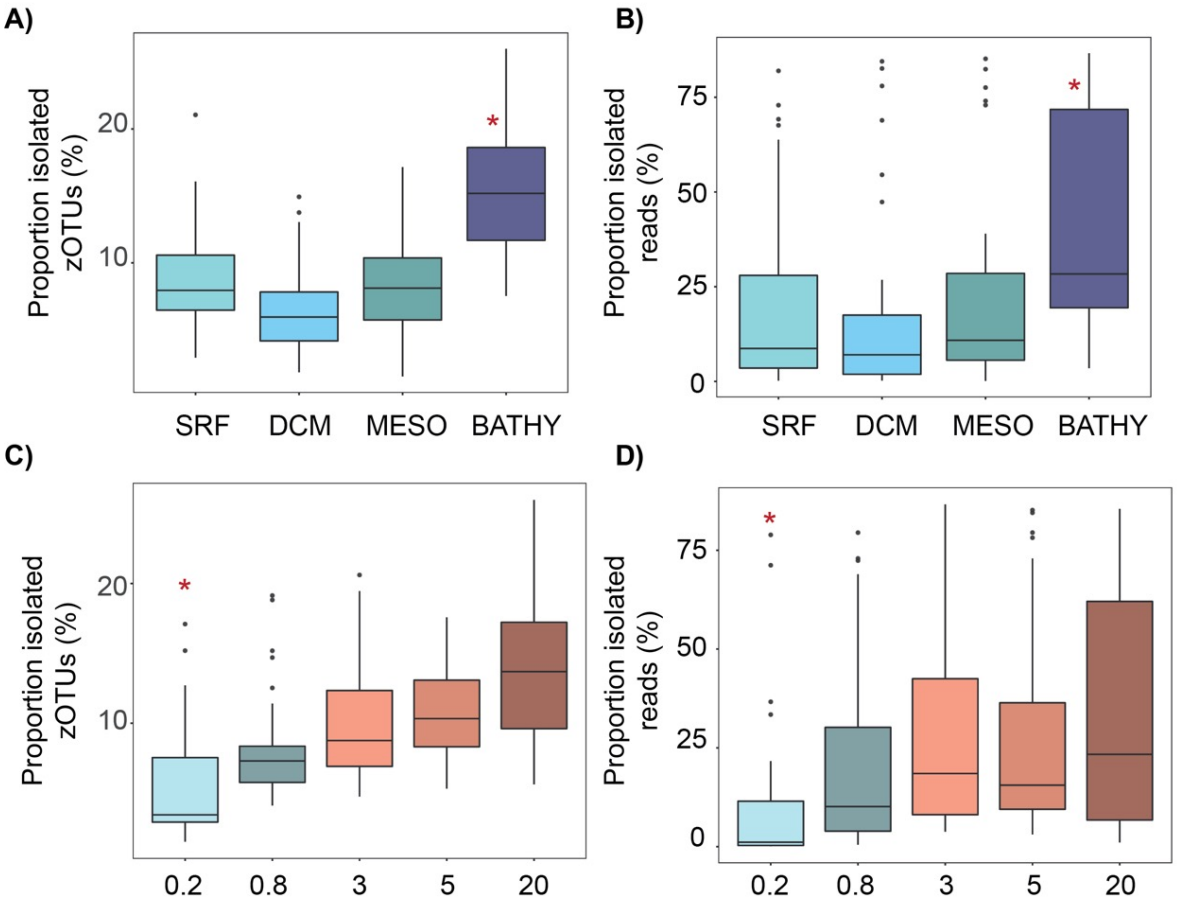

Figure S14: Boxplots from Malaspina profiles with different size fractions showing: **(A,B)** the percentages of isolated zOTUs and reads, respectively, that were 100% identical to at least one isolate per depth, and **(C,D)** the percentages of isolated zOTUs and reads, respectively, that were 100% identical to at least one isolate per size fraction. SFC: surface, DCM: deep chlorophyll maximum, MESO: mesopelagic, and BATHY: bathypelagic. 0.2: free-living bacteria, 0.8: bacteria attached to small particles, and 3-20: bacteria attached to big particles. Sizes of the fractions are in µm. Significative differences between layers or size fractions (Kruskal-Wallis, P.value <0.05) are indicated by red asterisks.

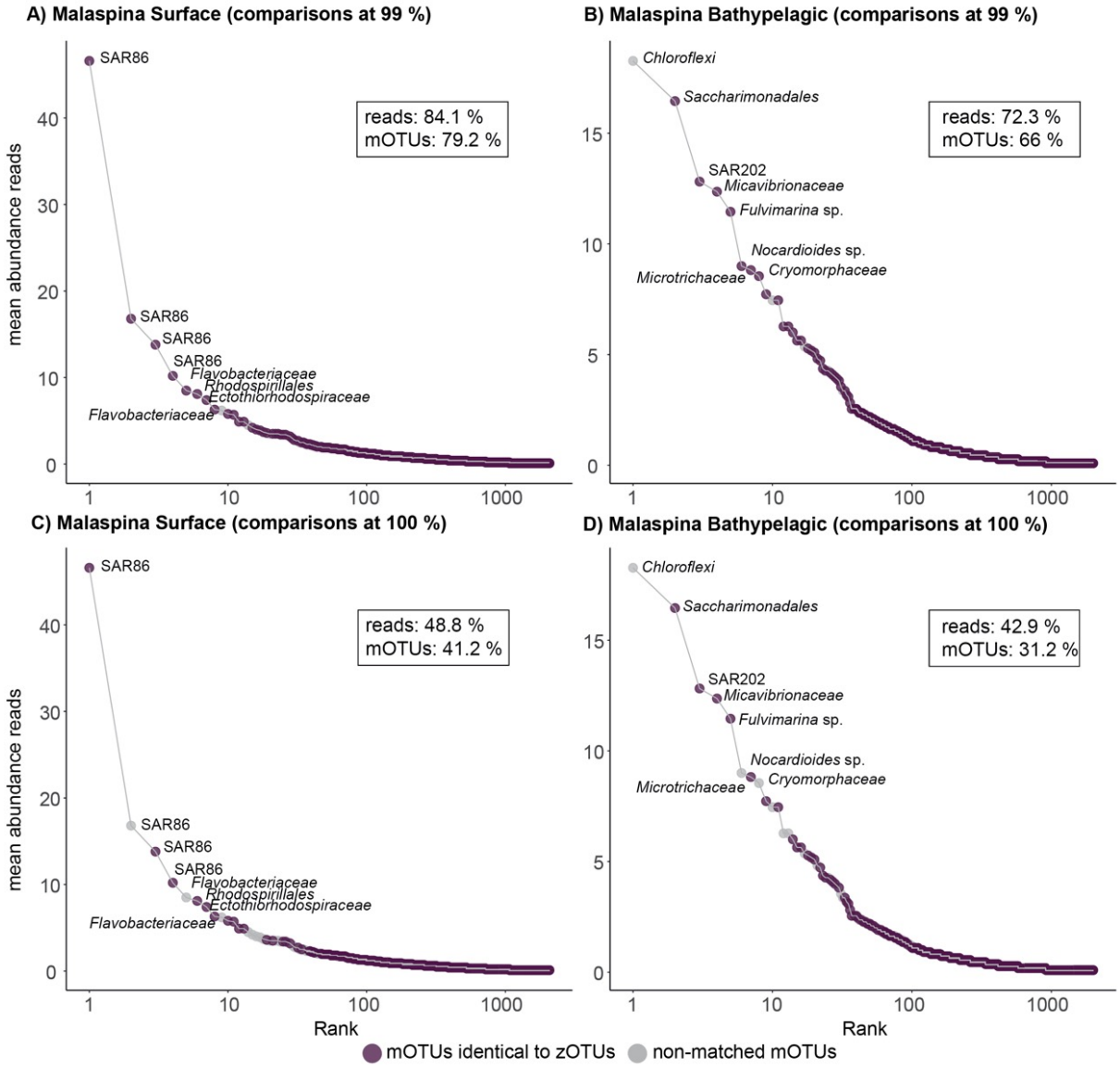

Figure S15: Rank plots showing the identified mOTUs to zOTUS in surface and bathypelagic Malaspina Expedition datasets. The identified mOTUs in Malaspina Surface and Bathypelagic samples had been ranked based on their medium abundances across samples. Color of the dots indicate if the mOTUs matched at 99% **(A,B)** or 100% similarity **(C,D)** with at least one zOTU: grey, mOTUs non identified in the zOTUs, and purple, mOTUs identified between the zOTUs. Taxonomy has been written for the top 8 most abundant mOTUs based on the mean relative abundances. Percentage of shared mOTUs identical to zOTUs and their corresponding number of reads is included in the plots per each dataset and comparison.

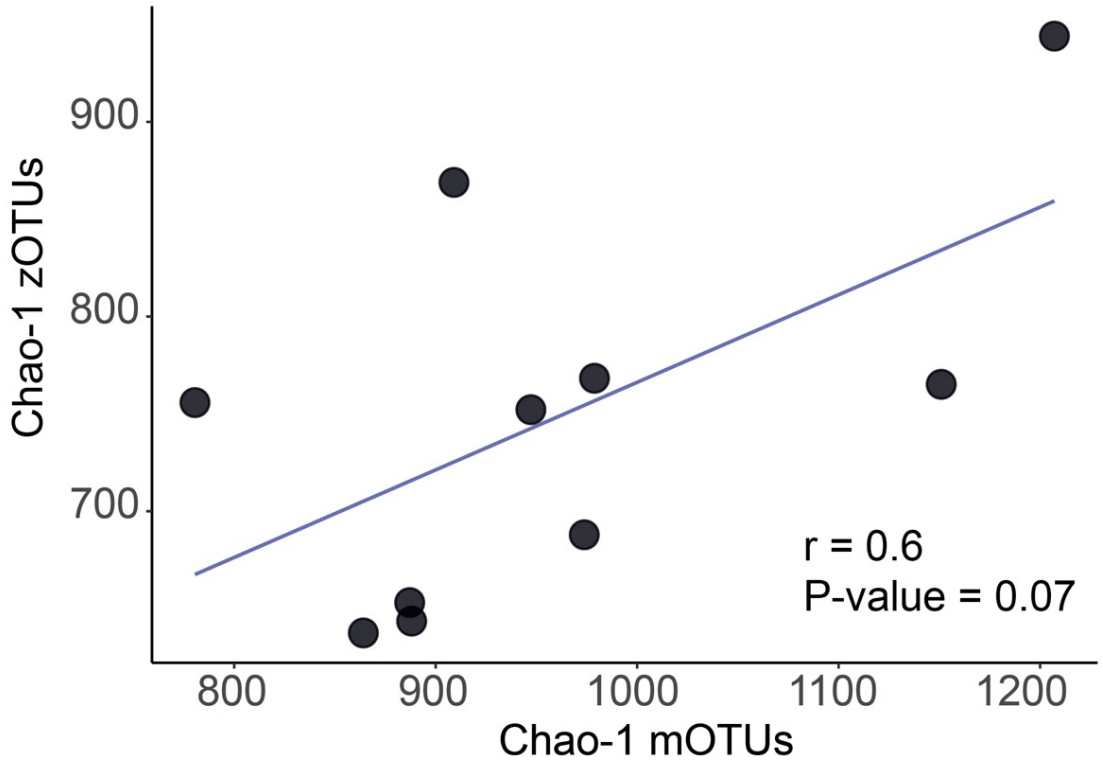

Figure S16: Correlation between Chao-1 diversity indexes calculated from the rarefied OTU-abundances tables of the mOTUs and the zOTUs from Malaspina Surface datasets. Correlation value (r) and P-value are indicated in the graph.

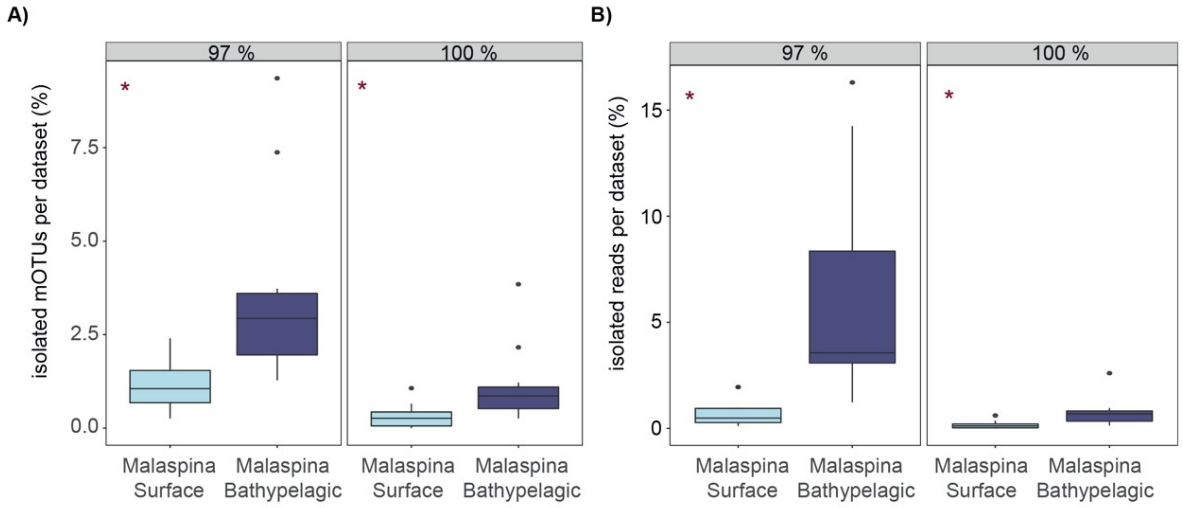

Figure S17: **(A)** Proportion of isolated mOTUs. **(B)** Proportion of isolated reads (miTAGs). Values are extracted from rarefied mOTUs-abundance tables. Outliers are indicated with grey circles. Significant differences between surface and bathypelagic Malaspina metagenomic datasets are indicated inside boxplots with red asterisks (Wilcoxon , P-value < 0.01). Notice that the scale between isolated mOTUs and the proportion of isolated reads is different.

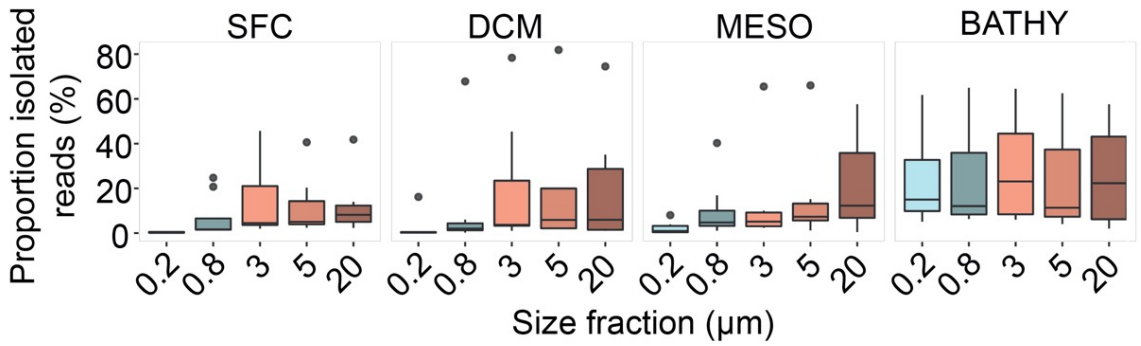

Figure S18: Distribution of bathypelagic abundant zOTUs (>1% of the reads) in all layers and size fractions. Values indicated represent the mean abundance of reads across layers and fractions. SFC: surface, DCM: deep chlorophyll maximum, MESO: mesopelagic, and BATHY: bathypelagic. 0.2: free-living bacteria, 0.8: bacteria attached to small particles, and 3-20: bacteria attached to big particles.

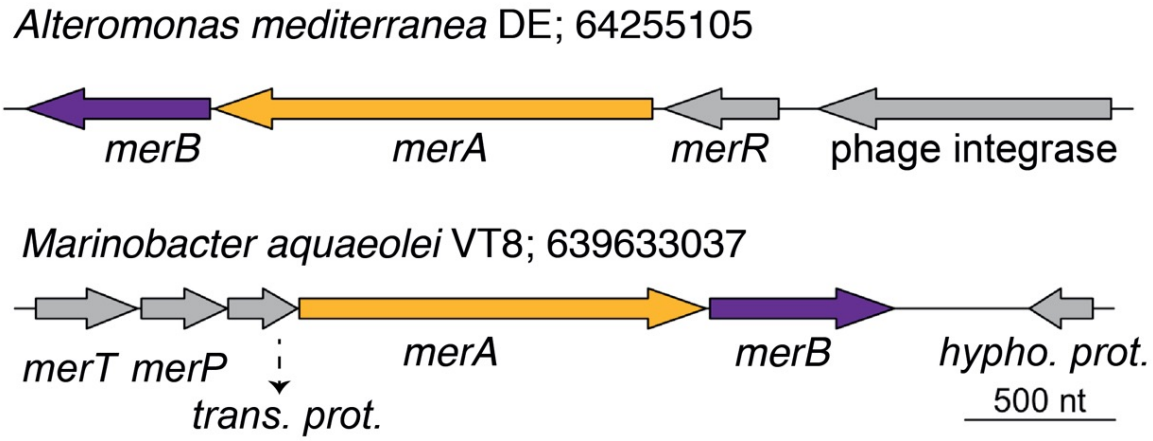

Figure S19: Co-localization of *merA* and *merB* genes in *Alteromonas mediterranea* DE and *Marinobacter aquaeolei* VT8. Coordinates extracted from the JGI/IMG database. Accession number in the IMG/JGI database indicated with the species name. Genes: *merA*, mercuric reductase; *merB*, organomercurial lyase; *merR*, *mer* operon regulator; *merT*, *merP* and trans. prot.: mercury transport proteins; hypho. prot: hypothetical protein.

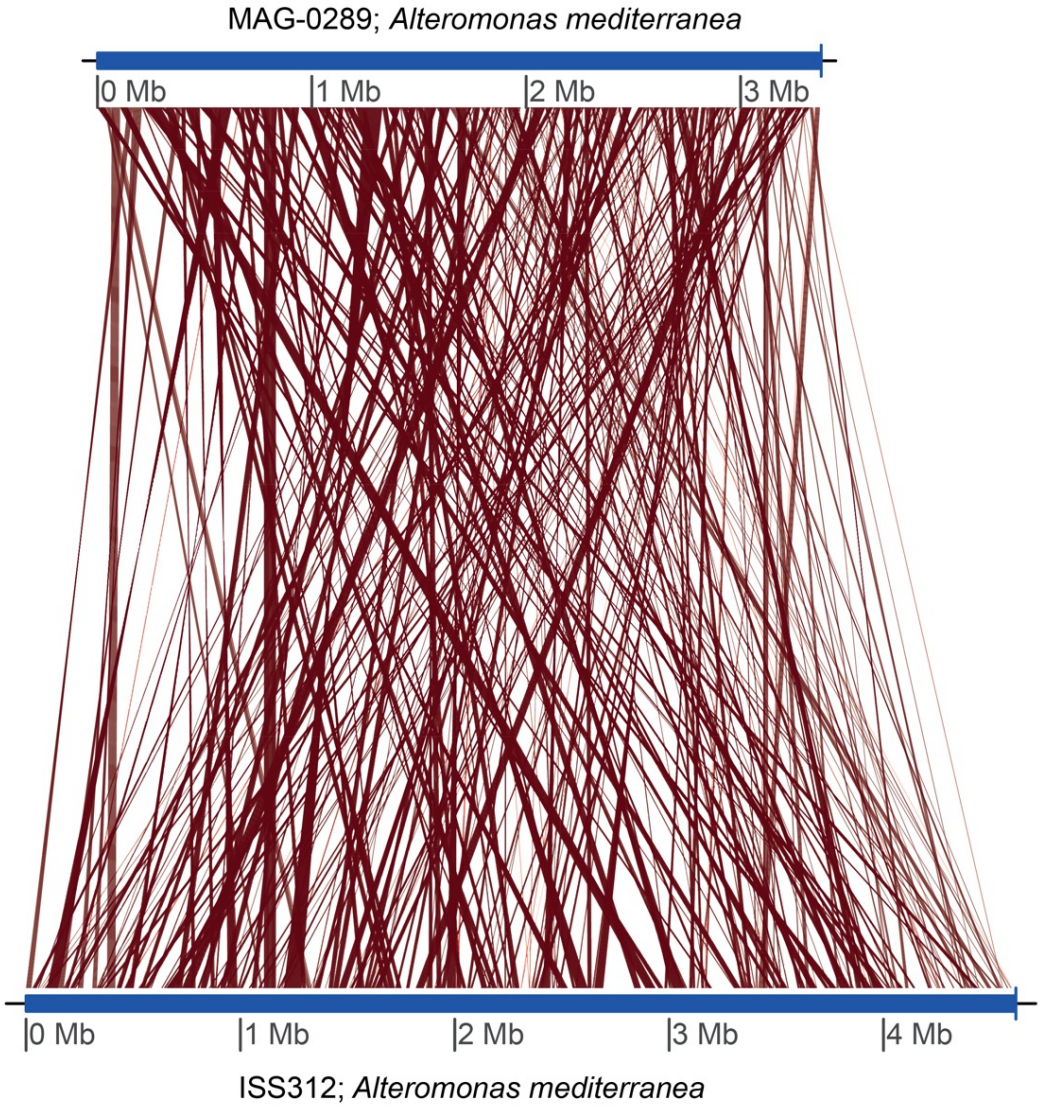

Figure S20: FastANI alignment between *Alteromonas mediterranea* genomes of the reconstructed MAG-0289 and the strain ISS312.

Supplementary Tables

Due to their size, the following tables cannot be displayed here. You can download them clicking here[14].

Table S1: Culture media and incubation conditions used for each seawater sample. Positive signs indicate which media were used. RT: room temperature.

Table S2: Non-subsampled iOTU-abundance table per depth defined at 99% sequence similarity. SRF: photic-layer; DE: bathypelagic; MES: mesopelagic.

Table S3: Metadata information of the Closest Cultured Match (CCM) obtained after BLASTn analysis of the isolates against a subset of the RDP database including only sequences from previously published cultured bacteria. SRF: photic-layer; DE: bathypelagic; MES: mesopelagic.

Table S4: Metadata information of the Closest Environmental Match (CEM) obtained after BLASTn analysis of the isolates against a subset of the RDP database including only sequences from previously published uncultured bacteria. SRF: photic-layer; DE: bathypelagic; MES: mesopelagic.

Table S5: Comparisons between the number of iOTUs (isolated OTUs) and the percentage of shared sequences between photic-layer, mesopelagic, and bathypelagic samples in non-subsampled and subsampled OTU tables.

Table S6: Comparisons of the richness and diversity indexes estimated using the isolates OTU (iOTU) tables per layer defined at 100% and 99% sequence similarity rarefied and non-rarefied to the layer with the lower number of isolates (Mesopelagic, 362 isolates). No R., stands for non-rarefied or non-subsampled OTU table, and R., stands for rarefied or subsampled OTU table.

Table S7: Subsampled iOTU-abundance table including all isolates from the photic, the mesopelagic and the bathypelagic. iOTUs obtained with clustering at 99% sequence similarity. SRF: photic-layer; DE: bathypelagic; MES: mesopelagic.

Table S8: Table indicating the number of total isolates affiliating to each genus found in the photic, mesopelagic, and bathypelagic samples. Results obtained after grouping all the iOTUs from the subsampled iOTU-abundance table (99% clustering) affiliating with the same genus. SRF: photic-layer isolates; MES: mesopelagic; DE: bathypelagic isolates.

Table S9: Table indicating the number of total isolates affiliating to each genus found in the photic, mesopelagic, and bathypelagic samples per stations. Results obtained after grouping all the isolated OTUs (iOTUs) from the non-subsampled iOTU-abundance table (99% clustering) affiliating with the same genus. SRF: photic-layer isolates; MES: mesopelagic; DE: bathypelagic isolates.

Table S10: Correspondence at 100% sequence similarity between zOTUs (zero-radious OTUs) and the top 12 iOTUs (99% clustering) including other iOTUs less abundant/rare matching with the same zOTU.

Table S11: New potential isolates best hits. Closest Cultured Match (CCM) and Closest Environmental/Uncultured Match (CEM) of the ISS653, ISS1889 and ISS2026 isolates BLASTn results against the NCBI, RDP 11 and SILVA LTP databases. Accession number and percentage of similarity are indicated together with the best hit.

[14] http://tarod.cmima.csic.es/tmp/data/marbits/acinasLab/thesis-sanz-saez/

Table S12: Annotated proteins detected in only one of the genomes of the new *Mesonia* strains. 33 proteins are codified only in ISS653, whereas 6 are unique in ISS1889.

Table S13: Characteristics of the different samples used for isolation of marine heterotrophic bacteria. Non-redundant isolates stand for the number of different isolates remaining after removing those that were 100% identical in their partial 16S rRNA gene.

Table S14: Culture media and incubation conditions used for each seawater sample. Positive signs indicate which media were used. RT: room temperature.

Table S15: Summary of the total reads, total OTUs and zOTUs, lowest number of reads and reads after rarification for each photic and aphotic *Tara* Oceans and Malaspina Expedition amplicon 16S iTAGs datasets.

Table S16: Summary of the total number of reads, lowest number of reads and number of mOTUs for Malaspina Surface and Bathypelagic metagenomic datasets.

Table S17: Summary of the proportions of isolated zOTUs (100% similarity) and OTUs (97% similarity) identical to isolates for the three levels of comparisons: global, photic and aphotic, and per stations used for isolation.

Table S18: Proportions of isolates 100% identical to zOTUs or mOTUs per dataset, and proportions of those that did not match with any zOTU or mOTU.

Table S19: Taxonomic classification of the different isolated zOTUs found in *Tara* Oceans and Malaspina Expedition amplicon 16S iTAGs datasets.

Table S20: Proportion of isolated reads at 100% similarity in the Malaspina size fraction dataset per layer and size fraction. SFC: surface, DCM: deep chlorophyll maximum, MESO: mesopelagic; BATHY: bathypelagic; sd, standard deviation; IQR, interquartile range.

Table S21: Proportion of isolated zOTUs at 100% similarity in the Malaspina size fraction dataset per layer and size fraction. SFC: surface, DCM: deep chlorophyll maximum, MESO: mesopelagic; BATHY: bathypelagic; sd, standard deviation; IQR, interquartile range.

Table S22: Similarity between metagenomic datasets (mOTUs) and amplicon 16S iTAGs datasets (zOTUs) after comparisons of mOTUs vs zOTUs at 100% and 99% sequence similarity.

Table S23: Relative mean abundances of the isolated mOTUs identified in Malaspina Expedition datasets and their classification as members of the abundant, mid-abundant or rare biosphere.

Table S24: Summary of the proportions of isolated zOTUs (100% similarity) identical to isolates removing *Cyanobacteria* from the zOTUs abundance tables in photic and aphotic *Tara* Oceans and Malaspina Expedition datasets.

Table S25: Mean abundance of the abundant isolated zOTUs (>1% reads) from the bathypelagic (BATHY) samples extracted from the Malaspina size fraction dataset and their respective mean abundances in the surface (SFC), deep chlorophyll maximum (DCM) and mesopelagic (MESO) samples per size fraction.

Table S26: Downloaded *merA* and *merB* sequences from the JGI/IMG database (2016) to design the PCR primers. In green are selected those *merA* copies that should be recognized by the primers designed.

Table S27: Candidate genera for mercury bioremediation obtained from the search of *merA* and *merB* genes in the KEGG database and from the IMG/JGI database.

Table S28: Summary of the putative isolates within the MARINHET culture collection that could harbor *merA* and *merB* genes based on the BLASTn search between the partial 16S rRNA genes sequences of the isolates and the 16S rRNA gene sequences of the candidate genera extracted from the JGI/IMG database.

Table S29: Information of the *Alteromonas* sp. and *Marinobacter* sp. strains used for PCR screening of the *merA* and *merB* genes.

Table S30: OTUs defined at 99% clustering with the *Alteromonas* and *Marinobacter* positive strains in order to select candidates for MIC analyses.

Table S31: Summary of the results for the MIC determination assays to $HgCl_2$.

Table S32: Concentrations of MeHg and inorganic mercury at different time points during the growth curves at 1 µM and 5 µM. SD: standard deviation; LOD: level of detail.

Table S33: Biotic (culture) and abiotic controls (killed and medium alone) of the MeHg degradation performed by the ISS312 *Alteromonas* strain.

Table S34: Summary of the number of positive strains per oceanographic region.

Table S35: Size of the *Alteromonas* and *Marinobacter* genomes (ISS312 and MAGs) used for biogeography analysis. Number of *merA* and *merB* copies is indicated per each genome. NA: data not available.

List of figures and tables

LIST OF FIGURES

LIST OF TABLES

Bibliography

BIBLIOGRAPHY

Acinas, S., J. Antón, and F. Rodríguez-Valera
1999. Diversity of free-living and attached bacteria in offshore Western Mediterranean waters as depicted by analysis of genes encoding 16S rRNA. *Applied and environmental microbiology*, 65(2):514–522.

Acinas, S. G., I. Ferrera, H. Sarmento, C. Díez-Vives, I. Forn, C. Ruiz-González, F. Cornejo-Castillo, G. Salazar, and J. Gasol
2015. Validation of a new catalysed reporter deposition-fluorescence in situ hybridization probe for the accurate quantification of marine Bacteroidetes populations. *Environmental Microbiology*, 17(10):3557–3569.

Acinas, S. G., V. Klepac-Ceraj, and D. E. Hunt
2003. Fine-scale phylogenetic architecture of a complex bacterial community. *Nature*, 104:551–554.

Acinas, S. G., P. Sánchez, G. Salazar, F. M. Cornejo-Castillo, M. Sebastián, R. Logares, S. Sunagawa, P. Hingamp, H. Ogata, G. Lima-Mendez, S. Roux, J. M. González, J. M. Arrieta, I. S. Alam, A. Kamau, C. Bowler, J. Raes, S. Pesant, P. Bork, S. Agustí, T. Gojobori, V. Bajic, D. Vaqué, M. B. Sullivan, C. Pedrós-Alió, R. Massana, C. M. Duarte, and J. M. Gasol
2019. Metabolic architecture of the deep ocean microbiome. *bioRxiv*, P. 635680.

Agogué, H., D. Lamy, P. R. Neal, M. L. Sogin, and G. J. Herndl
2011. Water mass-specificity of bacterial communities in the North Atlantic revealed by massively parallel sequencing. *Molecular Ecology*, 20(2):258–274.

Alava, J. J., W. W. L. Cheung, P. S. Ross, and U. R. Sumaila
2017. Climate change-contaminant interactions in marine food webs: Toward a conceptual framework. *Global Change Biology*, 23(10):3984–4001.

Alava, J. J., A. M. Cisneros-Montemayor, U. R. Sumaila, and W. W. L. Cheung
2018. Projected amplification of food web bioaccumulation of MeHg and PCBs under climate change in the Northeastern Pacific. *Scientific Reports*, 8(1):13460.

Alberti, A., J. Poulain, S. Engelen, K. Labadie, S. Romac, I. Ferrera, G. Albini, J.-M. Aury, C. Belser, A. Bertrand, C. Cruaud, C. Da Silva, C. Dossat, F. Gavory, S. Gas, J. Guy, M. Haquelle, E. Jacoby, O. Jaillon, A. Lemainque, E. Pelletier, G. Samson, M. Wessner, G. T. Genoscope Technical Team, S. G. Acinas, M. Royo-Llonch, F. M. Cornejo-Castillo, R. Logares, B. Fernández-Gómez, C. Bowler, G. Cochrane, C. Amid, P. T. Hoopen, C. De Vargas, N. Grimsley, E. Desgranges, S. Kandels-Lewis, H. Ogata, N. Poulton, M. E. Sieracki, R. Stepanauskas, M. B. Sullivan, J. R. Brum, M. B. Duhaime, B. T. Poulos, B. L. Hurwitz, T. O. C. Tara Oceans Consortium Coordinators, S. Pesant, E. Karsenti, and P. Wincker
2017. Viral to metazoan marine plankton nucleotide sequences from the Tara Oceans expedition. *Scientific data*, 4:170093.

Alonso-Sáez, L. and J. M. Gasol
2007. Seasonal variations in the contributions of different bacterial groups to the uptake of low-molecular-weight compounds in northwestern Mediterranean coastal waters. *Applied and environmental microbiology*, 73(11):3528–33525.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman
1990. Basic local alignment search tool. *Journal of molecular biology*, 215(3):403–410.

Amann, R., W. Ludwig, and K. Schleifer
1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Reviews*, 59(1):143–169.

Amaral-Zettler, L., L. F. Artigas, J. Baross, L. Bharathi, A. Boetius, D. Chandramohan, G. Herndl, K. Kogure, P. Neal, C. Pedrós -Alió, A. Ramette, S. Schouten, L. Stal, A. Thessen, J. de Leeuw, and M. Sogin
2010. A global census of marine microbes. In *Life in the World's Oceans*, Pp. 221–245.

Amos, H. M., D. J. Jacob, D. G. Streets, and E. M. Sunderland
2013. Legacy impacts of all-time anthropogenic emissions on the global mercury cycle. *Global Biogeochemical Cycles*, 27(2):410–421.

Andrady, A. L.
2011. Microplastics in the marine environment. *Marine Pollution Bulletin*, 62(8):1596–1605.

Arístegui, J., C. M. Duarte, S. Agustí, M. Doval, X. A. Álvarez-Salgado, and D. A. Hansell
2002. Dissolved organic carbon support of respiration in the dark ocean. *Science*, 371:298.

Arístegui, J., J. M. Gasol, C. M. Duarte, and G. J. Herndl
2009. Microbial oceanography of the dark oceans pelagic realm. *Limnology and Oceanography*, 54(5):1501–1529.

Arrieta, J. M., S. G. Acinas, R. Massana, and D. Vaqué
2012. Concentración de biomasa de picoeucariotas, bacterias y virus marinos para la extracción de ácidos nucleicos. In *Libro blanco de métodos y técnicas de trabajo oceanográfico*, E. Moreno□Ostos, ed., Pp. 521–532. CSIC.

Auta, H., C. Emenike, and S. Fauziah
2017. Distribution and importance of microplastics in the marine environment: A review of the sources, fate, effects, and potential solutions. *Environment International*, 102:165–176.

Azam, F.
1998. Microbial control of oceanic carbon flux: The plot thickens. *Science*, 280(5364):694–696.

Bakir, F., S. Damluji, L. Amin-Zaki, M. Murtadha, A. Khalidi, N. Al-Rawi, S. Tikriti, H. Dhahir, T. Clarkson, J. Smith, and R. Doherty
1973. Methylmercury poisoning in Iraq. Technical Report 4096.

Baltar, F., J. Arístegui, J. M. Gasol, I. Lekunberri, and G. J. Herndl
2010a. Mesoscale eddies: hotspots of prokaryotic activity and differential community structure in the ocean. *The ISME Journal*, 4(8):975–988.

Baltar, F., J. Arístegui, E. Sintes, J. M. Gasol, T. Reinthaler, and G. J. Herndl
2010b. Significance of non-sinking particulate organic carbon and dark CO 2 fixation to heterotrophic carbon demand in the mesopelagic northeast Atlantic. *Geophysical Research Letters*, 37(9):1–6.

Bar-On, Y. M., R. Phillips, and R. Milo
2018. The biomass distribution on Earth. *Proceedings of the National Academy of Sciences of the United States of America*, 115(25):6506–6511.

Barkay, T.
1987. Adaptation of aquatic microbial communities to Hg stress. *Applied and environmental microbiology*, 53(12):2725–2732.

Barkay, T., K. Kritee, E. Boyd, and G. Geesey
2010. A thermophilic bacterial origin and subsequent constraints by redox, light and salinity on the evolution of the microbial mercuric reductase. *Environmental Microbiology*, 12(11):2904–2917.

Barkay, T., S. M. Miller, and A. O. Summers
2003. Bacterial mercury resistance from atoms to ecosystems. *FEMS Microbiology Reviews*, 27(2-3):355–384.

Bartlett, D. H.
1999. Microbial adaptations to the psychrosphere/piezosphere. Technical Report 1.

Bass, L., C. Liebert, M. Lee, A. Summers, D. White, S. Thayer, and J. Maurer
1999. Incidence and characterization of integrons, genetic elements mediating multiple-drug resistance, in avian Escherichia coli. *Antimicrobial agents and chemotherapy*, 43(12):2925–2929.

Baumann, L., P. Baumann, M. Mandel, and R. D. Allen
1972. Taxonomy of aerobic marine eubacteria. *Journal of bacteriology*, 110(1):402–429.

Ben-Dov, E., E. Kramarsky-Winter, and A. Kushmaro
2009. An in situ method for cultivating microorganisms using a double encapsulation technique. *FEMS Microbiology Ecology*, 68(3):363–371.

Benotti, M. J. and B. J. Brownawell
2009. Microbial degradation of pharmaceuticals in estuarine and coastal seawater. *Environmental Pollution*, 157(3):994–1002.

Bernardet, J.-F., Y. Nakagawa, B. Holmes, S. o. t. t. o. F. of Prokaryotes, and C. like bacteria of the International Committee on Systematics
2002. Proposed minimal standards for describing new taxa of the family Flavobacteriaceae and emended description of the family. *International Journal of Systematic and Evolutionary Microbiology*, 52(3):1049–1070.

Berzas Nevado, J., R. Rodríguez Martín-Doimeadios, E. Krupp, F. Guzmán Bernardo, N. Rodríguez Fariñas, M. Jiménez Moreno, D. Wallace, and M. Patiño Ropero
2011. Comparison of gas chromatographic hyphenated techniques for mercury speciation analysis. *Journal of Chromatography A*, 1218(28):4545–4551.

Blum, J. D., B. N. Popp, J. C. Drazen, C. Anela Choy, and M. W. Johnson
2013. Methylmercury production below the mixed layer in the North Pacific Ocean. *Nature Geoscience*, 6(10):879–884.

Bochdansky, A. B., M. A. Clouse, and G. J. Herndl
2016. Dragon kings of the deep sea: marine particles deviate markedly from the common number-size spectrum. *Scientific Reports*, 6(1):22633.

Boitard, L., D. Cottinet, N. Bremond, J. Baudry, and J. Bibette
2015. Growing microbes in millifluidic droplets. *Engineering in Life Sciences*, 15(3):318–326.

Bolhuis, H. and M. Cretoiu
2016. What is so special about marine microorganisms? Introduction to the marine microbiome—From diversity to Biotechnological Potential. In *The Marine Microbiome*, L. Stal and M. Cretoiu, eds., Pp. 3–20. Springer.

Bopp, L., C. Bowler, L. Guidi, É. Karsenti, and C. de Vargas
2015. The Ocean: A carbon pump. In *Ocean climate*, Pp. 12–17.

Boss, E., A. Waite, H. Y. Frank Muller-Karger, R. Wanninkhof, J. Uitz, S. Thomalla, H. Sosik, B. Sloyan, G. G. Anthony Richardson, Patricia Miloslavich, Johannes Karstensen, K. Fennel, H. Claustre, M. Cornejo, I. Berman-Frank, S. Batten, and S. Acinas
2018. The time is right to expand biological measurements in ocean observing programs. *Limnology and Oceanography Bulletin*, (august):1–2.

Bowman, J. P.
2006. The marine clade of the family Flavobacteriaceae: The genera Aequorivita, Arenibacter, Cellulophaga, Croceibacter, Formosa, Gelidibacter, Gillisia, Maribacter, Mesonia, Muricauda, Polaribacter, Psychroflexus, Psychroserpens, Robiginitalea, Salegentibacter. In *The Prokaryotes*, M. Dworkin, S. Falkow, E. Rosenberg, K. Schleifer, and E. Stackebrandt, eds., Pp. 677–694. New York, NY: Springer New York.

Bowman, K. L., R. E. Collins, A. M. Agather, C. H. Lamborg, C. R. Hammerschmidt, D. Kaul, C. L. Dupont, G. A. Christensen, and D. A. Elias
2020. Distribution of mercury□cycling genes in the Arctic and equatorial Pacific Oceans and their relationship to mercury speciation. *Limnology and Oceanography*, 65(S1):S310–S320.

Boyd, E. S. and T. Barkay
2012. The mercury resistance operon: From an origin in a geothermal environment to an efficient detoxification machine. *Frontiers in Microbiology*, 3(October):349.

Bridou, R., M. Monperrus, P. R. Gonzalez, R. Guyoneaud, and D. Amouroux
2011. Simultaneous determination of mercury methylation and demethylation capacities of various sulfate-reducing bacteria using species-specific isotopic tracers. *Environmental Toxicology and Chemistry*, 30(2):337–344.

Bristow, L. A., W. Mohr, S. Ahmerkamp, and M. M. Kuypers
2017. Nutrients that limit growth in the ocean. *Current Biology*, 27:R474–R478.

Browne, H. P., S. C. Forster, B. O. Anonye, N. Kumar, B. A. Neville, M. D. Stares, D. Goulding, and T. D. Lawley
2016. Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature*, 533(7604):543–546.

Buerger, S., A. Spoering, E. Gavrish, C. Leslin, L. Ling, and S. Epstein
2012. Microbial scout hypothesis and microbial discovery. *Applied and Environmental Microbiology*, 78(9):3221–3228.

Castro Da Silva, M. A., A. Cavalett, A. Spinner, D. C. Rosa, R. B. Jasper, M. C. Quecine, M. L. Bonatelli, A. Pizzirani-Kleiner, G. Corção, A. Oliveira, and S. Lima
2013. Phylogenetic identification of marine bacteria isolated from deep-sea sediments of the eastern South Atlantic Ocean. *SpringerPlus*, 2:127.

Cavicchioli, R., M. Ostrowski, F. Fegatella, A. Goodchild, and N. Guixa-Boixereu
2003. Life under nutrient limitation in oligotrophic marine environments: an eco/physiological perspective of Sphingopyxis alaskensis (formerly Sphingomonas alaskensis). *Microbial ecology*, 45(3):203–217.

Chen, P., L. Zhang, X. Guo, X. Dai, L. Liu, L. Xi, J. Wang, L. Song, Y. Wang, Y. Zhu, L. Huang, and Y. Huang
2016. Diversity, biogeography, and biodegradation potential of Actinobacteria in the deep-sea dediments along the Southwest Indian ridge. *Frontiers in Microbiology*, 7:1340.

Chiu, H.-H., W. Y. Shieh, S. Y. Lin, C.-M. Tseng, P.-W. Chiang, and I. Wagner-Dö Bler
2007. Alteromonas tagae sp. nov. and Alteromonas simiduii sp. nov., mercury-resistant bacteria isolated from a Taiwanese estuary. *International Journal of Systematic and Evolutionay Microbiology*, 57:1209–1216.

Cho, J.-C. and S. J. Giovannoni
2004. Cultivation and growth characteristics of a diverse group of oligotrophic marine Gammaproteobacteria. *Applied and environmental microbiology*, 70(1):432–440.

Ciok, A., K. Budzik, M. K. Zdanowski, J. Gawor, J. Grzesiak, P. Decewicz, R. Gromadka, D. Bartosik, and L. Dziewit
2018. Plasmids of psychrotolerant polaromonas spp. isolated From Arctic and Antarctic glaciers - diversity and role in adaptation to polar environments. *Frontiers in microbiology*, 9:1285.

Clarkson, T.
1993. Mercury: major issues in environmental health. *Environmental health perspectives*, 100:31–38.

Clarkson, T. W. and L. Magos
2006. The toxicology of mercury and its chemical compounds. *Critical reviews in toxicology*, 36(8):609–662.

Colwell, R. K. and J. A. Coddington
1994. Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 345(1311):101–118.

Connon, S. A. and S. J. Giovannoni
2002. High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. *Applied and Environmental Microbiology*, 68(8):3878–3885.

core team, R.
2017. R core team. A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria https://www.R-project. org/.

Cossa, D., B. Averty, and N. Pirrone
2009. The origin of methylmercury in open Mediterranean waters. *Limnology and Oceanography*, 54(3):837–844.

Cossa, D., L.-E. Heimbürger, D. Lannuzel, S. R. Rintoul, E. C. Butler, A. R. Bowie, B. Averty, R. J. Watson, and T. Remenyi
2011. Mercury in the Southern Ocean. *Geochimica et Cosmochimica Acta*, 75(14):4037–4052.

Cossa, D., J.-M. Martin, K. Takayanagi, and J. Sanjuan
1997. The distribution and cycling of mercury species in the western Mediterranean. *Deep Sea Research Part II*, 44(3-4):721–740.

Cottrell, M. and D. Kirchman
2000. Community composition of marine bacterioplankton determined by 16S rRNA gene clone libraries and fluorescence In situ hybridization. *Applied and environmental microbiology*, 66(12):5116–5122.

Cottrell, M. T. and D. L. Kirchman
2016. Transcriptional control in marine copiotrophic and oligotrophic bacteria with streamlined genomes. *Applied and environmental microbiology*, 82(19):6010–6018.

Covelli, S., J. Faganeli, M. Horvat, and A. Brambati
2001. Mercury contamination of coastal sediments as the result of long-term cinnabar mining activity (Gulf of Trieste, Northern Adriatic sea). *Applied Geochemistry*, 16(5):541–558.

Crespo, B. G., T. Pommier, B. Fernández-Gómez, and C. Pedrós-Alió
2013. Taxonomic composition of the particle-attached and free-living bacterial assemblages in the Northwest Mediterranean Sea analyzed by pyrosequencing of the 16S rRNA. *MicrobiologyOpen*, 2(4):541–552.

Crespo, B. G., P. J. Wallhead, R. Logares, and C. Pedrós-alió
2016. Probing the rare biosphere of the North-West Mediterranean Sea: an experiment with high sequencing effort. *PloS one*, 11(7):e0159195.

Crutzen, P. J.
2002. Geology of mankind. *Nature*, 415(6867):23.

Curtis, T. P., W. T. Sloan, and J. W. Scannell
2002. Estimating prokaryotic diversity and its limits. *Proceedings of the National Academy of Sciences of the United States of America*, 99(16):10494–10499.

Dash, H. R., N. Mangwani, J. Chakraborty, S. Kumari, and S. Das
2013. Marine bacteria: Potential candidates for enhanced bioremediation. *Applied Microbiology and Biotechnology*, 97:561–571.

De, J. and N. Ramaiah
2007. Characterization of marine bacteria highly resistant to mercury exhibiting multiple resistances to toxic chemicals. *Ecological Indicators*, 7(3):511–520.

De, J., N. Ramaiah, A. Mesquita, and X. N. Verlekar
2003. Tolerance to various toxicants by marine bacteria highly resistant to mercury. *Marine Biotechnology*, 5(2):185–193.

del Giorgio, P. A. and C. M. Duarte
2002. Respiration in the open ocean. *Nature*, 420(6914):379–384.

DeLong, E. F.
1992. Archaea in coastal marine environments. *Proceedings of the National Academy of Sciences of the United States of America*, 89(12):5685–5689.

Deming, J. W. and J. A. Baross
2000. Survival, dormancy, and nonculturable cells in extreme deep-sea environment. In *Nonculturable microorganisms in the environments*, R. R. Colwell and D. J. Grimes, eds., Pp. 147–197. ASM Press, Washington DC.

Deng, X. and P. Wang
2012. Isolation of marine bacteria highly resistant to mercury and their bioaccumulation process. *Bioresource Technology*, 121:342–347.

Deppe, U., H.-H. Richnow, W. Michaelis, and G. Antranikian
2005. Degradation of crude oil by an arctic microbial consortium. *Extremophiles*, 9(6):461–470.

Díez-Vives, C., S. Nielsen, P. Sánchez, O. Palenzuela, I. Ferrera, M. Sebastián, C. Pedrós-Alió, J. M. Gasol, and S. G. Acinas
2019. Delineation of ecologically distinct units of marine Bacteroidetes in the Northwestern Mediterranean Sea. *Molecular Ecology*, 28(11):2846–2859.

Duarte, C. M.
2015. Seafaring in the 21St Century: The Malaspina 2010 Circumnavigation Expedition. *Limnology and Oceanography Bulletin*, 24(1):11–14.

Duarte, C. M., D. K. Ngugi, I. Alam, J. Pearman, A. Kamau, V. M. Eguiluz, T. Gojobori, S. G. Acinas, J. M. Gasol, V. Bajic, and X. Irigoien
2020. Sequencing effort dictates gene discovery in marine microbial metagenomes. *Environmental Microbiology*, Pp. 1462–2920.15182.

Ducklow, H., D. Steinberg, and K. Buesseler
2001. Upper ocean carbon export and the biological pump. *Oceanography*, 14(4):50–58.

Edgar, R. C.
2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26(19):2460–2461.

Eguchi, M. and Y. Ishida
1990. Oligotrophic properties of heterotrophic bacteria and in situ heterotrophic activity in pelagic seawaters. *FEMS Microbiology Ecology*, 73:23–30.

Eguíluz, V. M., G. Salazar, J. Fernández-Gracia, J. K. Pearman, J. M. Gasol, S. G. Acinas, S. Sunagawa, X. Irigoien, and C. M. Duarte
2019. Scaling of species distribution explains the vast potential marine prokaryote diversity. *Scientific Reports*, 9(1):18710.

Eilers, H., J. Pernthaler, F. O. Glöckner, and R. Amann
2000. Culturability and in situ abundance of pelagic Bacteria from the North Sea. *Applied and Environmental Microbiology*, 66(7):3044–3051.

Eloe, E. A., F. Malfatti, J. Gutierrez, K. Hardy, W. E. Schmidt, K. Pogliano, J. Pogliano, F. Azam, and D. H. Bartlett
2011a. Isolation and characterization of a psychropiezophilic alphaproteobacterium. *Applied and environmental microbiology*, 77(22):8145–8153.

Eloe, E. A., C. N. Shulse, D. W. Fadrosh, S. J. Williamson, E. E. Allen, and D. H. Bartlett
2011b. Compositional differences in particle-associated and free-living microbial assemblages from an extreme deep-ocean environment. *Environmental Microbiology Reports*, 3(4):449–458.

Enrico, M., G. L. Roux, N. Marusczak, L.-E. Heimbürger, A. Claustres, X. Fu, R. Sun, and J. E. Sonke
2016. Atmospheric mercury transfer to peat bogs dominated by gaseous elemental mercury dry deposition. *Environmental Science & Technology*, 50(5):2405–2412.

Eriksen, H. H. and F. X. Perrez
2014. The Minamata convention: A comprehensive response to a global problem. *Review of European, Comparative and International Environmental Law*, 23(2):195–210.

Faith, D. P.
1992. Conservation evaluation and phylogenetic diversity. *Biological Conservation*, 61(1):1–10.

Fang, J., L. Zhang, and D. A. Bazylinski
2010. Deep-sea piezosphere and piezophiles: geomicrobiology and biogeochemistry. *Trends in microbiology*, 18:413–422.

Fashola, M. O., V. M. Ngole-Jeme, and O. O. Babalola
2016. Heavy metal pollution from gold mines: Environmental effects and bacterial strategies for resistance. *International journal of environmental research and public health*, 13(11).

Felske, A. D., W. Fehr, B. V. Pauling, H. von Canstein, and I. Wagner-Döbler
2003. Functional profiling of mercuric reductase (mer A) genes in biofilm communities of a technical scale biocatalyzer. *BMC microbiology*, 3:22.

Fenchel, T.
2008. The microbial loop – 25 years later. *Journal of Experimental Marine Biology and Ecology*, 366(1-2):99–103.

Ferrari, B. C., S. J. Binnerup, and M. Gillings
2005. Microcolony cultivation on a soil substrate membrane system selects for previously uncultured soil bacteria. *Applied and Environmental Microbiology*, 71(12):8714–8720.

Ferrera, I., A. B. Banta, and A.-L. Reysenbach
2014. Spatial patterns of Aquificales in deep-sea vents along the Eastern Lau Spreading Center (SW Pacific). *Systematic and Applied Microbiology*, 37(6):442–448.

Ferrera, I. and O. Sánchez
2016. Insights into microbial diversity in wastewater treatment systems: How far have we come? *Biotechnology Advances*, 34(5):790–802.

Field, C. B., M. J. Behrenfeld, J. T. Randerson, and P. Falkowski
1998. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science (New York, N.Y.)*, 281(5374):237–240.

Finster, K. W. and K. U. Kjeldsen
2010. Desulfovibrio oceani subsp. oceani sp. nov., subsp. nov. and Desulfovibrio oceani subsp. galateae subsp. nov., novel sulfate-reducing bacteria isolated from the oxygen minimum zone off the coast of Peru. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, 97(3):221–229.

Fisher, M. and E. Triplett
1999. Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. *Applied and environmental microbiology*, 65(10):4630–4636.

Floyd, M. M., J. Tang, M. Kane, and D. Emerson
2005. Captured diversity in a culture collection: case study of the geographic and habitat distributions of environmental isolates held at the american type culture collection. *Applied and environmental microbiology*, 71(6):2813–2823.

Fong, J. C., B. E. De Guzman, C. H. Lamborg, and M. P. Sison-Mangus
2019. The mercury-tolerant microbiota of the zooplankton Daphnia aids in host survival and maintains fecundity under mercury Stress. *Environmental science & technology*, 53:14688–14699.

Fontanez, K. M., J. M. Eppley, T. J. Samo, D. M. Karl, and E. F. DeLong
2015. Microbial community structure and function on sinking particles in the North Pacific Subtropical Gyre. *Frontiers in microbiology*, 6:469.

Francisco, D., R. Mah, and A. Rabin
1973. Acridine orange-epifluorescence technique for counting bacteria in natural waters. *Trans Am Microsc Soc*, 92(3):416–421.

Fuhrman, J., K. McCallum, and A. Davis
1992. Novel major archaebacterial group from marine plankton. *Nature*, 356(6365):148–149.

Fuhrman, J. A. and F. Azam
1980. Bacterioplankton secondary production estimates for coastal waters of British Columbia, Antarctica, and California. *Applied and environmental microbiology*, 39(6):1085–1095.

Fuhrman, J. A., J. A. Cram, and D. M. Needham
2015. Marine microbial community dynamics and their ecological interpretation. *Nature Reviews Microbiology*, 13(3):133–146.

Fuhrman, J. A. and A. A. Davis
1997. Widespread Archaea and novel Bacteria from the deep sea as shown by 16S rRNA gene sequences. *Marine Ecology Progress Series*, 150(1-3):275–285.

Galand, P. E., M. Potvin, E. O. Casamayor, and C. Lovejoy
2010. Hydrography shapes bacterial biogeography of the deep Arctic Ocean. *The ISME Journal*, 4(4):564–576.

Galata, V., T. Fehlmann, C. Backes, and A. Keller
2019. PLSDB: a resource of complete bacterial plasmids. *Nucleic acids research*, 47(D1):D195–D202.

Gao, W., D. Navarroli, J. Naimark, W. Zhang, S.-h. Chao, and D. R. Meldrum
2013. Microbe observation and cultivation array (MOCA) for cultivating and analyzing environmental microbiota. *Microbiome*, 1(1):4.

García-Martínez, J., S. G. Acinas, R. Massana, and F. Rodríguez-Valera
2002. Prevalence and microdiversity of Alteromonas macleodii-like microorganisms in different oceanic regions. *Environmental microbiology*, 4(1):42–50.

Gärtner, A., M. Blümel, J. Wiese, and J. F. Imhoff
2011. Isolation and characterisation of bacteria from the Eastern Mediterranean deep sea. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, 100:421–435.

Gasol, J. M. and P. A. del Giorgio
2000. Using flow cytometry for counting natural planktonic bacteria and understanding the structure of planktonic bacterial communities. *Scientia Marina*, 64(2):197–224.

Gasol, J. M. and D. L. Kirchman
2018. Introduction: The evolution of microbial ecology of the ocean. In *Microbial ecology of the oceans*, J. M. Gasol and D. L. Kirchman, eds., Pp. 1–46. Wiley-Blackwell.

GESAMP
1969. Report of the 1st Session. Technical report, GESAMP, London.

Gibbons, S. M., J. G. Caporaso, M. Pirrung, D. Field, R. Knight, and J. A. Gilbert
2013. Evidence for a persistent microbial seed bank throughout the global ocean. *Proceedings of the National Academy of Sciences of the United States of America*, 110(12):4651–4655.

Gionfriddo, C. M., M. T. Tate, R. R. Wick, M. B. Schultz, A. Zemla, M. P. Thelen, R. Schofield, D. P. Krabbenhoft, K. E. Holt, and J. W. Moreau
2016. Microbial mercury methylation in Antarctic sea ice. *Nature microbiology*, 1(10):16127.

Giovannoni, S., T. Britschgi, C. Moyer, and K. Field
1990. Genetic diversity in Sargasso Sea bacterioplankton. *Nature*, 345(6270):60–63.

Giovannoni, S. and U. Stingl
2007. The importance of culturing bacterioplankton in the ' omics ' age. *Nature reviews. Microbiology*, 5:820–826.

Giovannoni, S. J.
2017. SAR11 bacteria: The most abundant plankton in the oceans. *Annual Review of Marine Science*, 9(1):231–255.

Giovannoni, S. J., L. Bibbs, J.-C. Cho, M. D. Stapels, R. Desiderio, K. L. Vergin, M. S. Rappé, S. Laney, L. J. Wilhelm, H. J. Tripp, E. J. Mathur, and D. F. Barofsky
2005. Proteorhodopsin in the ubiquitous marine bacterium SAR11. *Nature*, 438(7064):82–85.

Giovannoni, S. J., T. D. Mullins, and K. G. Field
1995. Microbial diversity in oceanic systems: rRNA approaches to the study of unculturable microbes. In *Molecular Ecology of Aquatic Microbes*, I. Joint, ed., Pp. 217–248. Berlin, Heidelberg: Springer Berlin Heidelberg.

González, J., B. Fernández-Gómez, A. Fernàndez-Guerra, L. Gómez-Consarnau, O. Sánchez, M. Coll-Lladó, J. Del Campo, L. Escudero, R. Rodríguez-Martínez, L. Alonso-Sáez, M. Latasa, I. Paulsen, O. Nedashkovskaya, I. Lekunberri, J. Pinhassi, and C. Pedrós-Alió
2008. Genome analysis of the proteorhodopsin-containing marine bacterium Polaribacter sp. MED152 (Flavobacteria). *Proceedings of the National Academy of Sciences of the United States of America*, 105(25):8724–8729.

Griffin, H., T. Foster, S. Silver, and T. Misra
1987. Cloning and DNA sequence of the mercuric- and organomercurial-resistance determinants of plasmid pDU1358. *Proceedings of the National Academy of Sciences of the United States of America*, 84(10):3112–3116.

Grosche, A., H. Sekaran, I. Pérez-Rodríguez, V. Starovoytov, and C. Vetriani
2015. Cetia pacifica gen. nov., sp. nov., a chemolithoautotrophic, thermophilic, nitrate-ammonifying bacterium from a deep-sea hydrothermal vent. *International journal of systematic and evolutionary microbiology*, 65(Pt 4):1144–1150.

Gruber, N.
2016. Elusive marine nitrogen fixation. *Proceedings of the National Academy of Sciences of the Unitated States of America*, 113(16):4246–4248.

Gurevich, A., V. Saveliev, N. Vyahhi, and G. Tesler
2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics (Oxford, England)*, 29(8):1072–1075.

Gutierrez, T., J. F. Biddle, A. Teske, and M. D. Aitken
2015. Cultivation-dependent and cultivation-independent characterization of hydrocarbon-degrading bacteria in Guaymas Basin sediments. *Frontiers in Microbiology*, 6(JUL):1–12.

Gutleben, J., M. C. De Mares, J. Dirk Van Elsas, H. Smidt, J. Overmann, and D. Sipkema
2017. The multi-omics promise in context: from sequence to microbial isolate. *Critical Reviews in Microbiology*, 44(2):212–219.

Gworek, B., O. Bemowska-Kałabun, M. Kijeńska, and J. Wrzosek-Jakubowska
2016. Mercury in Marine and Oceanic Waters—a Review. *Water, Air, & Soil Pollution*, 227(10):371.

Hall, B. G., H. Acar, A. Nandipati, and M. Barlow
2014. Growth rates made easy. *Mol. Biol. Evol*, (31):232–238.

Hammerschmidt, C. R. and K. L. Bowman
2012. Vertical methylmercury distribution in the subtropical North Pacific Ocean. *Marine Chemistry*, 132-133:77–82.

Hanan, A. E., M. A. E. Gehan, and A. E. S. Nermeen
2011. Cadmium resisting bacteria in Alexandria Eastern Harbor (Egypt) and optimization of cadmium bioaccumulation by Vibrio harveyi. *African Journal of Biotechnology*, 10(17):3412–3423.

Handley, K. M. and J. R. Lloyd
2013. Biogeochemical implications of the ubiquitous colonization of marine habitats and redox gradients by Marinobacter species. *Frontiers in Microbiology*, 4:136.

Harada, M.
1995. Minamata disease: methylmercury poisoning in Japan caused by environmental pollution. *Critical reviews in toxicology*, 25(1):1–24.

Harada, M., K. Ito, N. Nakajima, S. Yamamura, M. Tomita, H. Suzuki, and S. Amachi
2019. Genomic analysis of Pseudomonas sp. strain SCT, an iodate-reducing bacterium isolated from marine sediment, reveals a possible use for bioremediation. *G3*, 9(5):1321–1329.

Harding, G., J. Dalziel, and P. Vass
2018. Bioaccumulation of methylmercury within the marine food web of the outer Bay of Fundy, Gulf of Maine. *PloS one*, 13(7):e0197220.

Harmsen, H., D. Prieur, and C. Jeanthon
1997. Distribution of microorganisms in deep-sea hydrothermal vent chimneys investigated by whole-cell hybridization and enrichment culture of thermophilic subpopulations. *Applied and Environmental Microbiology*, 63(7):2876–2883.

Hassanshahian, M., G. Emtiazi, G. Caruso, and S. Cappello
2014. Bioremediation (bioaugmentation/biostimulation) trials of oil polluted seawater: a mesocosm simulation study. *Marine environmental research*, 95:28–38.

Heimbürger, L.-E., D. Cossa, J.-C. Marty, C. Migon, B. Averty, A. Dufour, and J. Ras
2010. Methyl mercury distributions in relation to the presence of nano- and picophytoplankton in an oceanic water column (Ligurian Sea, North-western Mediterranean). *Geochimica et Cosmochimica Acta*, 74(19):5549–5559.

Herndl, G. and T. Reinthaler
2013. Microbial control of the dark end of the biological pump. *Nature Geoscience*, 6(9):718–724.

Hesselman, M. C., D. I. Odoni, B. M. Ryback, S. de Groot, R. G. A. van Heck, J. Keijsers, P. Kolkman, D. Nieuwenhuijse, Y. M. van Nuland, E. Sebus, R. Spee, H. de Vries, M. T. Wapenaar, C. J. Ingham, K. Schroën, V. A. P. Martins dos Santos, S. K. Spaans, F. Hugenholtz, and M. W. J. van Passel
2012. A multi-platform flow device for microbial (co-) cultivation and microscopic analysis. *PLoS ONE*, 7(5):e36982.

Heywood, J. L., M. V. Zubkov, G. A. Tarran, B. M. Fuchs, and P. M. Holligan
2006. Prokaryoplankton standing stocks in oligotrophic gyre and equatorial provinces of the Atlantic Ocean: Evaluation of inter-annual variability. *Deep-Sea Research Part II*, 53(14-16):1530–1547.

Hirayama, H., M. Sunamura, K. Takai, T. Nunoura, T. Noguchi, H. Oida, Y. Furushima, H. Yamamoto, T. Oomori, and K. Horikoshi
2007. Culture-dependent and -independent characterization of microbial communities associated with a shallow submarine hydrothermal system occurring within a coral reef off Taketomi Island, Japan. *Applied and environmental microbiology*, 73(23):7642–7656.

Hobbie, J., R. Daley, and S. Jasper
1977. Use of nucleopore filters for counting bacteria by fluorescence microscopu. *Applied and environmental microbiologyl*, 33(5):1225–1228.

Huber, J. A., D. A. Butterfield, and J. A. Baross
2002. Temporal changes in archaeal diversity and chemistry in a mid-ocean ridge subseafloor habitat. *Appl. Environ. Microbiol.*, 68(4):1585–1594.

Hunter-Cevera, J. C., D. Karl, and M. Buckley
2005. Marine microbial diversity: the key to Earth's habitability (A report from the American Academy of Microbiology). Technical report, American Society of Microbiology, Washington DC, American Society for Microbiology.

Hwang, C. Y., I. Lee, Y. Cho, Y. M. Lee, Y.-J. Jung, K. Baek, S.-I. Nam, and H. K. Lee
2015. Sediminicola arcticus sp. nov., a psychrophilic bacterium isolated from deep-sea sediment, and emended description of the genus Sediminicola. *INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY*, 65(Pt 5):1567–1571.

Ibarbalz, F. M., N. Henry, M. C. Brandão, S. Martini, G. Busseni, H. Byme, L. P. Coelho, H. Endo, J. Gasol, A. C. Gregiory, F. Mahé, J. Rigonato, M. Royo-Llonch, G. Salazar, I. Sanz-Sáez, E. Scalco, D. Soviadan, A. A. Zayed, A. Zingone, K. Labadie, J. Ferland, C. Marec, S. Kandels, M. Picheral, D. Céline, J. Poulain, S. Pisarev, M. Carmichael, S. Pesant, T. O. Coordinators, M. Babin, E. Boss, D. Iudiocone, O. Jaillon, S. Acinas, H. Ogata, E. Pelletier, L. Stemmann, M. B. Sullivan, S. Sunagawa, L. Bopp, C. De Vargas, L. Karp-Boss, P. Wincker, F. Lombard, C. Bowler, and L. Zinger
2019. Global trends in marine plankton diversity across kingdoms of life. *Cell*, 179:1084–1097.

Ingham, C. J., A. Sprenkels, J. Bomer, D. Molenaar, A. van den Berg, J. E. T. van Hylckama Vlieg, and W. M. de Vos
2007. The micro-Petri dish, a million-well growth chip for the culture and high-throughput screening of microorganisms. *Proceedings of the National Academy of Sciences*, 104(46):18217–18222.

Ivars-Martinez, E., A.-B. Martin-Cuadrado, G. D'Auria, A. Mira, S. Ferriera, J. Johnson, R. Friedman, and F. Rodriguez-Valera
2008. Comparative genomics of two ecotypes of the marine planktonic copiotroph Alteromonas macleodii suggests alternative lifestyles associated with different kinds of particulate organic matter. *The ISME journal*, 2(12):1194–1212.

Jannasch, H.
1958. Studies on planktonic bacteria by means of a direct membrane filter method. Technical report.

Jayaprakashvel, M., S. Vijay, C. Karthigeyan, and A. J. Hussain
2015. Isolation and characterization of mercury resistant bacteria from the coastal are of Chennai, India. *International Journal of Advanced Research*, 4(8):64–76.

Jensen, P. R., C. A. Kauffman, and W. Fenical
1996. High recovery of culturable bacteria from the surfaces of marine algae. *Marine Biology*, 126(1):1–7.

Joint, I., M. Mühling, and J. Querellou
2010. Culturing marine bacteria - an essential prerequisite for biodiscovery. *Microbial biotechnology*, 3(5):564–575.

Jones, S. E. and J. T. Lennon
2010. Dormancy contributes to the maintenance of microbial diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 107(13):5881–5886.

Kaeberlein, T., K. Lewis, and S. S. Epstein
2002. Isolating 'Uncultivable' microorganisms in pure culture in a simulated natural environment. *Science*, 296(5570):1127–1129.

Kai, W., Y. Peisheng, M. Rui, J. Wenwen, and S. Zongze
2017. Diversity of culturable bacteria in deep-sea water from the South Atlantic Ocean. *Bioengineered*, 8(5):572–584.

Kane, A. L., B. Al-Shayeb, P. V. Holec, S. Rajan, N. E. Le Mieux, S. C. Heinsch, S. Psarska, K. G. Aukema, C. A. Sarkar, E. A. Nater, and J. A. Gralnick
2016. Toward bioremediation of methylmercury using silica encapsulated Escherichia coli harboring the mer operon. *PLoS ONE*, 11(1):e0147036.

Kanehisa, M. and S. Goto
2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 28(1):27–30.

Karagas, M. R., A. L. Choi, E. Oken, M. Horvat, R. Schoeny, E. Kamai, W. Cowell, P. Grandjean, and S. Korrick
2012. Evidence on the human health effects of low-level methylmercury exposure. *Environmental Health Perspectives*, 120(6):799–806.

Karsenti, E., S. G. Acinas, P. Bork, C. Bowler, C. De Vargas, J. Raes, M. Sullivan, D. Arendt, F. Benzoni, J.-M. Claverie, M. Follows, G. Gorsky, P. Hingamp, D. Iudicone, O. Jaillon, S. Kandels-Lewis, U. Krzic, F. Not, H. Ogata, S. Pesant, E. G. Reynaud, C. Sardet, M. E. Sieracki, S. Speich, D. Velayoudon, J. Weissenbach, P. Wincker, and T. O. Consortium
2011. A holistic approach to marine eco-systems biology. *PLoS Biology*, 9:7–11.

Kaye, J. Z. and J. A. Baross
2000. High incidence of halotolerant bacteria in Pacific hydrothermal-vent and pelagic environments. *FEMS microbiology ecology*, 32(3):249–260.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, S. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meinties, and A. Drummond
2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12):1647–1649.

Kembel, S., P. Cowan, M. Helmus, W. Cornwell, H. Morlon, D. Ackerly, S. Blomberg, and C. Webb
2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, 26:1463–1464.

Kim, M. K. and K. D. Zoh
2012. Fate and transport of mercury in environmental media and human exposure. *Journal of Preventive Medicine and Public Health*, 45(6):335–343.

Kirchman, D. L.
2002. The ecology of Cytophaga-Flavobacteria in aquatic environments. *FEMS Microbiology Ecology*, 39:91–100.

Köck-Schulmeyer, M., C. Postigo, M. Farré, D. Barceló, and M. López de Alda
2019. Medium to highly polar pesticides in seawater: Analysis and fate in coastal areas of Catalonia (NE Spain). *Chemosphere*, 215:515–523.

Kogure, K., U. Simidu, and N. Taga
1979. A tentative direct microscopic method for counting living marine bacteria. *Canadian journal of microbiology*, 25(3):415–420.

Kozlov, A. M., D. Darriba, T. Flouri, B. Morel, and A. Stamatakis
2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, 35(21):4453–4455.

Kultima, J. R., S. Sunagawa, J. Li, W. Chen, H. Chen, D. R. Mende, M. Arumugam, Q. Pan, B. Liu, J. Qin, J. Wang, and P. Bork
2012. MOCAT: A Metagenomics Assembly and Gene Prediction Toolkit. *PLoS ONE*, 7(10):e47656.

Lagier, J., F. Armougom, M. Million, P. Hugon, I. Pagnier, C. Robert, F. Bittar, G. Fournous, G. Gimenez, M. Maraninchi, J.-F. Trape, E. Koonin, B. La Scola, and D. Raoult
2012. Microbial culturomics: paradigm shift in the human gut microbiome study. *Clinical microbiology and infection*, 18(12):1185–1193.

Lamborg, C., K. Bowman, C. Hammerschmidt, C. Gilmour, K. Munson, N. Selin, and C. M. Tseng
2014a. Mercury in the anthropocene ocean. *Oceanography*, 27(1):76–87.

Lamborg, C. H., W. F. Fitzgerald, A. W. H. Damman, J. M. Benoit, P. H. Balcom, and D. R. Engstrom
2002. Modern and historic atmospheric mercury fluxes in both hemispheres: Global and regional mercury cycling implications. *Global Biogeochemical Cycles*, 16(4):1104.

Lamborg, C. H., C. R. Hammerschmidt, and K. L. Bowman
2016. An examination of the role of particles in oceanic mercury cycling. *Philosophical Transactions of the Royal Society A*, 374(2081):20150297.

Lamborg, C. H., C. R. Hammerschmidt, K. L. Bowman, G. J. Swarr, K. M. Munson, D. C. Ohnemus, P. J. Lam, L.-E. Heimbürger, M. J. A. Rijkenberg, and M. A. Saito
2014b. A global ocean inventory of anthropogenic mercury based on water column measurements. *Nature*, 512(7512):65–68.

Lambrechts, S., A. Willems, and G. Tahon
2019. Uncovering the uncultivated majority in Antarctic soils: Toward a synergistic approach. *Frontiers in Microbiology*, 10:242.

Lara, E., J. M. Arrieta, I. Garcia-Zarandona, J. A. Boras, C. M. Duarte, S. Agustí, P. F. Wassmann, and D. Vaqué
2013. Experimental evaluation of the warming effect on viral, bacterial and protistan communities in two contrasting Arctic systems. *Aquatic Microbial Ecology*, 70(1):17–32.

Laurier, F., R. Mason, G. Gill, and L. Whalin
2004. Mercury distributions in the North Pacific Ocean—20 years of observations. *Marine Chemistry*, 90(1-4):3–19.

Lauro, F. M., D. McDougald, T. Thomas, T. J. Williams, S. Egan, S. Rice, M. Z. DeMaere, L. Ting, H. Ertan, J. Johnson, S. Ferriera, A. Lapidus, I. Anderson, N. Kyrpides, A. C. Munk, C. Detter, C. S. Han, M. V. Brown, F. T. Robb, S. Kjelleberg, and R. Cavicchioli
2009. The genomic basis of trophic strategy in marine bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 106(37):15527–15533.

Lee, C., J. Y. Kim, W. I. Lee, K. L. Nelson, J. Yoon, and D. L. Sedlak
2008. Bactericidal effect of zero-valent iron nanoparticles on Escherichia coli. *Environmental Science & Technology*, 42(13):4927–4933.

Lehnherr, I., V. L. St. Louis, H. Hintelmann, and J. L. Kirk
2011. Methylation of inorganic mercury in polar marine waters. *Nature Geoscience*, 4(5):298–302.

Lekunberri, I., J. Gasol, S. Acinas, L. Gómez-Consarnau, B. Crespo, E. Casamayor, R. Massana, C. Pedrós-Alió, and J. Pinhassi
2014. The phylogenetic and ecological context of cultured and whole genome-sequenced planktonic bacteria from the coastal NW Mediterranean Sea. *Systematic and Applied Microbiology*, 37(3):216–228.

Letunic, I. and P. Bork
2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research*, 47(W1):W256–W259.

Li, P., X. Feng, G. Qiu, L. Shang, and Z. Li
2009. Mercury pollution in Asia: a review of the contaminated sites. *Journal of Hazardous Materials*, 168(2-3):591–601.

Lima de Silva, A. A., M. A. R. de Carvalho, S. A. L. de Souza, P. M. T. Dias, R. G. da Silva Filho, C. S. de Meirelles Saramago, C. A. de Melo Bento, and E. Hofer
2012. Heavy metal tolerance (Cr, Ag AND Hg) in bacteria isolated from sewage. *Brazilian journal of microbiology*, 43(4):1620–1631.

Lima-Mendez, G., K. Faust, N. Henry, J. Decelle, S. Colin, F. Carcillo, S. Chaffron, J. C. Ignacio-Espinosa, S. Roux, F. Vincent, L. Bittner, Y. Darzi, J. Wang, S. Audic, L. Berline, G. Bontempi, A. M. Cabello, L. Coppola, F. M. Cornejo-Castillo, L. De Meester, I. Ferrera, M.-J. Garet-Delmas, L. Guidi, E. Lara, S. Pesant, M. Royo-Llonch, G. Salazar, P. Sánchez, M. Sebastian, C. Souffreau, C. Dimier, M. Picheral, M. B. Sullivan, E. Karsenti, C. Bowler, C. de Vargas, and J. Raes
2015. Determinants of community structure in the global plankton interactome. *Science*, 348(6237):1262073.

Liu, Q., J. Fang, J. Li, L. Zhang, B.-B. Xie, X.-L. Chen, and Y.-Z. Zhang
2018. Depth-resolved variations of cultivable bacteria and their extracellular enzymes in the water column of the New Britain trench. *Frontiers in Microbiology*, 9:135.

Liu, W., T. Marsh, H. Cheng, and L. Forney
1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Applied and environmental microbiology*, 63(11):4516–4522.

Lloyd, K. G., A. D. Steen, J. Ladau, J. Yin, and L. Crosby
2018. Phylogenetically novel uncultured microbial cells dominate Earth microbiomes. *mSystems*, 3(5):e00055–18.

Locey, K. J. and J. T. Lennon
2016. Scaling laws predict global microbial diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 113(21):5970–5975.

Logares, R., S. Sunagawa, G. Salazar, F. M. Cornejo-Castillo, I. Ferrera, H. Sarmento, P. Hingamp, H. Ogata, C. de Vargas, G. Lima-Mendez, J. Raes, J. Poulain, O. Jaillon, P. Wincker, S. Kandels-Lewis, E. Karsenti, P. Bork, and S. G. Acinas
2014. Metagenomic 16S rDNA Illumina tags are a powerful alternative to amplicon sequencing to explore diversity and structure of microbial communities. *Environmental Microbiology*, 16(9):2659–2671.

Lohmann, R. and J. Dachs
2019. Polychlorinated biphenyls in the global ocean. In *World seas: an environmental evaluation*, C. Sheppard, ed., Pp. 269–282. Academic Press.

Longhurst, A. R.
2007. *Ecological geography of the sea*. Academic Press.

López-Pacheco, I. Y., A. Silva-Núñez, C. Salinas-Salazar, A. Arévalo-Gallegos, L. A. Lizarazo-Holguin, D. Barceló, H. M. Iqbal, and R. Parra-Saldívar
2019. Anthropogenic contaminants of high concern: Existence in water resources and their adverse effects. *Science of The Total Environment*, 690:1068–1088.

López-Pérez, M., A. Gonzaga, A.-B. Martin-Cuadrado, O. Onyshchenko, A. Ghavidel, R. Ghai, and F. Rodriguez-Valera
2012. Genomes of surface isolates of Alteromonas macleodii: the life of a widespread marine opportunistic copiotroph. *Scientific reports*, 2:696.

Lucena, T., D. Arahal, I. Sanz-Sáez, S. Acinas, O. Sánchez, R. Aznar, C. Pedrós-Alió, and M. Pujalte
2020a. Thalassocella blandensis gen. nov., sp. nov., a novel member of the family Cellvibrionaceae. *International journal of systematic and evolutionary microbiology*, 70(2):1231–1239.

Lucena, T., I. Sanz-Sáez, D. R. Arahal, S. G. Acinas, O. Sánchez, C. Pedrós-Alió, R. Aznar, and M. J. Pujalte
2020b. Mesonia oceanica sp. nov., isolated from oceans during the Tara oceans expedition, with a preference for mesopelagic waters. *International journal of systematic and evolutionary microbiology*, 70(7):4329–4338.

Luna, G. M.
2015. Biotechnological Potential of Marine Microbes. In *Springer Handbook of Marine Biotechnology*, S. Kim, ed., Pp. 651–661. Springer Berlin Heidelberg.

Ma, L., J. Kim, R. Hatzenpichler, M. A. Karymov, N. Hubert, I. M. Hanan, E. B. Chang, and R. F. Ismagilov
2014. Gene-targeted microfluidic cultivation validated by isolation of a gut bacterium listed in Human Microbiome Project's Most Wanted taxa. *Proceedings of the National Academy of Sciences of the United States of America*, 111(27):9768.

Malcolm, E. G., J. K. Schaefer, E. B. Ekstrom, C. B. Tuit, A. Jayakumar, H. Park, B. B. Ward, and F. M. Morel
2010. Mercury methylation in oxygen deficient zones of the oceans: No evidence for the predominance of anaerobes. *Marine Chemistry*, 122(1-4):11–19.

Martin-Platero, A. M., B. Cleary, K. Kauffman, S. P. Preheim, D. J. McGillicuddy, E. J. Alm, and M. F. Polz
2018. High resolution time series reveals cohesive but short-lived communities in coastal plankton. *Nature Communications*, 9(1):266.

Martiny, A. C.
2019. High proportions of bacteria are culturable across major biomes. *ISME Journal*, 13:2125–2128.

Mason, R., K. Rolfhus, and W. Fitzgerald
1998. Mercury in the North Atlantic. *Marine Chemistry*, 61(1-2):37–53.

Mason, R. P., A. L. Choi, W. F. Fitzgerald, C. R. Hammerschmidt, C. H. Lamborg, A. L. Soerensen, and E. M. Sunderland
2012. Mercury biogeochemical cycling in the ocean and policy implications. *Environmental research*, 119:101–17.

Mason, R. P. and G.-R. Sheu
2002. Role of the ocean in the global mercury cycle. *Global Biogeochemical Cycles*, 16(4):1093.

Massana, R., E. DeLong, and C. Pedrós-Alió
2000. A few cosmopolitan phylotypes dominate planktonic archaeal assemblages in widely different oceanic provinces. *Applied and environmental microbiology*, 66(5):1777–1787.

Massana, R., A. E. Murray, C. M. Preston, and E. F. Delong
1997. Vertical Distribution and Phylogenetic Characterization of Marine Planktonic Archaea in the Santa Barbara Channel. Technical Report 1.

Math, R. K., H. M. Jin, J. M. Kim, Y. Hahn, W. Park, E. L. Madsen, and C. O. Jeon
2012. Comparative genomics reveals adaptation by Alteromonas sp. SN2 to marine tidal-flat conditions: Cold tolerance and aromatic hydrocarbon metabolism. *PLoS ONE*, 7(4):e35784.

Mathema, B. V., B. C. Thakuri, and M. Sillanpää
2011. Bacterial mer operon-mediated detoxification of mercurial compounds: a short review. *Archives of Microbiology*, 193:837–844.

McBride, M. J.
2014. The family Flavobacteriaceae. In *The Prokaryotes*, E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson, eds., Pp. 643–676. Berlin, Heidelberg: Springer Berlin Heidelberg.

Menezes, L. D., G. L. Fernandes, A. B. Mulla, R. M. Meena, and S. R. Damare
2018. Diversity of culturable Sulphur-oxidising bacteria in the oxygen minimum zones of the northern Indian Ocean. *Journal of Marine Systems*, 209(May 2017):103085.

Mergler, D., H. A. Anderson, L. H. M. Chan, K. R. Mahaffey, M. Murray, M. Sakamoto, and A. H. Stern
2007. Methylmercury exposure and health effects in humans: A worldwide concern. *https://doi.org/10.1579/0044-7447(2007)36[3:MEAHEI]2.0.CO;2*, 36(1):3–11.

Mestre, M., C. Ruiz-González, R. Logares, C. M. Duarte, J. M. Gasol, and M. M. Sala
2018. Sinking particles promote vertical connectivity in the ocean microbiome. *Proceedings of the National Academy of Sciences of the United States of America*, 115(29):E6799–E6807.

MIDI
2008. Sherlock Microbial Identification System Operating Manual, version 6.1. Technical report, MIDI Inc., Newark.

Miller, M. and T. Clarkson
1973. *Mercury, mercurial, and mercaptans.* Thomas, Springfield, ILL.

Mindlin, S., G. Kholodii, Z. Gorlenko, S. Minakhina, L. Minakhin, E. Kalyaeva, A. Kopteva, M. Petrova, O. Yurieva, and V. Nikiforov
2001. Mercury resistance transposons of Gram-negative environmental bacteria and their classification. *Research in Microbiology*, 152(9):811–822.

Moalic, Y., D. Desbruyères, C. M. Duarte, A. F. Rozenfeld, C. Bachraty, and S. Arnaud-Haond
2012. Biogeography revisited with network theory: retracing the history of hydrothermal vent communities. *Systematic Biology*, 61(1):127–137.

Mojiri, A., J. L. Zhou, A. Ohashi, N. Ozaki, and T. Kindaichi
2019. Comprehensive review of polycyclic aromatic hydrocarbons in water sources, their effects and treatments. *Science of The Total Environment*, 696:133971.

Møller, A. K., T. Barkay, M. A. Hansen, A. Norman, L. H. Hansen, S. J. Sørensen, E. S. Boyd, and N. Kroer
2014. Mercuric reductase genes (merA) and mercury resistance plasmids in High Arctic snow, freshwater and sea-ice brine. *FEMS Microbiology Ecology*, 87(1):52–63.

Monperrus, M., E. Tessier, D. Amouroux, A. Leynaert, P. Huonnic, and O. Donard
2007. Mercury methylation, demethylation and reduction rates in coastal and marine surface waters of the Mediterranean Sea. *Marine Chemistry*, 107(1):49–63.

Mora, C., D. P. Tittensor, S. Adl, A. G. B. Simpson, and B. Worm
2011. How many species are there on Earth and in the Ocean? *PLoS Biology*, 9(8):e1001127.

Morishita, K., K. Nakamura, K. Tuchiya, K. Nishimura, M. Iwahara, and O. Yagi
2006. Removal of methylmercury from a fish broth by Alteromonas macledii isolated from Minamata Bay. *Japanese Journal of Water Treatment Biology*, 42(2):45–51.

Morris, R. M., M. S. Rappé, S. A. Connon, K. L. Vergin, W. A. Siebold, C. A. Carlson, and S. J. Giovannoni
2002. SAR11 clade dominates ocean surface bacterioplankton communities. *Nature*, 420(6917):806–810.

Mulla, A., G. Fernandes, L. Menezes, R. M. Meena, H. Naik, M. Gauns, and S. Damare
2018. Diversity of culturable nitrate-reducing bacteria from the Arabian Sea oxygen minimum zone. *Deep-Sea Research Part II: Topical Studies in Oceanography*, 156(December 2017):27–33.

Muller, E. E., E. Glaab, P. May, N. Vlassis, and P. Wilmes
2013. Condensing the omics fog of microbial communities. *Trends in microbiology*, 21(7):325–333.

Mullins, T. D., T. B. Britschgi, R. L. Krest, and S. J. Giovannoni
1995. Genetic comparisons reveal the same unknown bacterial lineages in Atlantic and Pacific bacterioplankton communities. *Limnology and Oceanography*, 40(1):148–158.

Munson, K. M., C. H. Lamborg, R. M. Boiteau, and M. A. Saito
2018. Dynamic mercury methylation and demethylation in oligotrophic marine water. *Biogeosciences*, 15:6451–6460.

Muyzer, G., E. de Waal, and A. Uitterlinden
1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and environmental microbiology*, 59(3):695–700.

Nakagawa, S. and K. Takai
2008. Deep-sea vent chemoautotrophs: diversity, biochemistry and ecological significance. *FEMS Microbiology Ecology*, 65(1):1–14.

Nakamura, K., J. Aoki, and M. Yamamoto
2000. Mercury volatilization by the most mercury-resistant bacteria from the seawater of Minamata Bay in various physiological conditions. *Clean Products and Processes*, 2:174–178.

Nakamura, K., M. Iwahara, and K. Furukawa
2001. Screening of organomercurial-volatilizing bacteria in the mercury-polluted sediments and seawater of Minamata Bay in Japan. *Clean Products and Processes*, 3(2):104–107.

Nedashkovskaya, O. I., S. B. Kim, N. V. Zhukova, J. Kwak, V. V. Mikhailov, and K. S. Bae
2006. Mesonia mobilis sp. nov., isolated from seawater, and emended description of the genus Mesonia. *INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY*, 56(10):2433–2436.

Nichols, D., N. Cahoon, E. M. Trakhtenberg, L. Pham, A. Mehta, A. Belanger, T. Kanigan, K. Lewis, and S. S. Epstein
2010. Use of ichip for high-throughput in situ cultivation of 'uncultivable' microbial species. *Applied and environmental microbiology*, 76(8):2445–2450.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. Henry, H. Stevens, E. Szoecs, and H. Wagner
2018. Vegan: community ecology package. R package version 2.5-3. https://CRAN. R-project.org/package=vegan.

Olson, B., T. Barkay, and R. Colwell
1979. Role of plasmids in mercury transformation by bacteria isolated from the aquatic environment. *Applied and environmental microbiology*, 38(3):478–485.

Orani, A. M., E. Vassileva, S. Azemard, and C. Alonso-Hernandez
2020. Trace elements contamination assessment in marine sediments from different regions of the Caribbean Sea. *Journal of Hazardous Materials*, 399:122934.

Oregaard, G. and S. J. Sørensen
2007. High diversity of bacterial mercuric reductase genes from surface and sub-surface floodplain soil (Oak Ridge, USA). *The ISME Journal*, 1(5):453–467.

Oremland, R., C. Culbertson, and M. Winfrey
1991. Methylmercury decomposition in sediments and bacterial cultures: involvement of methanogens and sulfate reducers in oxidative demethylation. *Applied and environmental microbiology*, 57(1):130–137.

Osborn, A. M., K. D. Bruce, P. Strike, and D. A. Ritchie
1997. Distribution, diversity and evolution of the bacterial mercury resistance (mer) operon. *FEMS Microbiology Reviews*, 19(4):239–262.

Overmann, J., B. Abt, and J. Sikorski
2017. Present and future of culturing bacteria. *Annu. Rev. Microbiol*, 71:711–730.

Overmann, J. and C. Lepleux
2016. Marine Bacteria and Archaea: Diversity, adaptations, and culturability. In *The Marine Microbiome*, L. Stal and M. Cretoiu, eds., Pp. 21–56. Springer.

Oyetibo, G. O., S. T. Ishola, W. Ikeda-Ohtsubo, K. Miyauchi, M. O. Ilori, and G. Endo
2015. Mercury bioremoval by Yarrowia strains isolated from sediments of mercury-polluted estuarine water. *Applied Microbiology and Biotechnology*, 99(8):3651–3657.

Pace, N. R., D. A. Stahl, D. J. Lane, and G. J. Olsen
1986. The analysis of natural microbial populations by ribosomal RNA sequences. In *Advances in microbial ecology*, K. Marshall, ed., Pp. 1–55. Springer, Boston, MA.

Page, K. A., S. A. Connon, and S. J. Giovannoni
2004. Representative freshwater bacterioplankton isolated from Crater Lake, Oregon. *Applied and environmental microbiology*, 70(11):6542–6550.

Palma Esposito, F., C. J. Ingham, R. Hurtado-Ortiz, C. Bizet, D. Tasdemir, and D. de Pascale
2018. Isolation by miniaturized culture chip of an Antarctic bacterium Aequorivita sp. with antimicrobial and anthelmintic activity. *Biotechnology reports*, 20:e00281.

Pande, S. and C. Kost
2017. Bacterial unculturability and the formation of intercellular metabolic networks. *Trends in Microbiology*, 25(5):349–361.

Panwichian, S., D. Kantachote, B. Wittayaweerasak, and M. Mallavarapu
2011. Removal of heavy metals by exopolymeric substances produced by resistant purple nonsulfur bacteria isolated from contaminated shrimp ponds. *Electronic Journal of Biotechnology*, 14:ISSN: 0717–3458.

Parada, A. E., D. M. Needham, and J. A. Fuhrman
2016. Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology*, 18(5):1403–1414.

Paradis, E., J. Claude, and K. Strimmer
2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20:289–290.

Park, J., A. Kerner, M. A. Burns, and X. N. Lin
2011. Microdroplet-enabled highly parallel co-cultivation of microbial communities. *PLoS ONE*, 6(2):e17019.

Parker, C. T., B. J. Tindall, and G. M. Garrity
2019. International code of nomenclature of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*, 69(1A):S1–S111.

Parks, D. H., M. Chuvochina, D. W. Waite, C. Rinke, A. Skarshewski, P.-A. Chaumeil, and P. Hugenholtz
2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature Biotechnology*, 36(10):996–1004.

Parks, D. H., M. Imelfort, C. T. Skennerton, P. Hugenholtz, and G. W. Tyson
2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome research*, 25(7):1043–1055.

Parks, D. H., C. Rinke, M. Chuvochina, P.-A. Chaumeil, B. J. Woodcroft, P. N. Evans, P. Hugenholtz, and G. W. Tyson
2017. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nature Microbiology*, 2:1533–1542.

Parte, A. C., J. Sardà Carbasse, J. P. Meier-Kolthoff, L. C. Reimer, and M. Göker
2020. List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. *International Journal of Systematic and Evolutionary Microbiology*, P. ijsem004332.

Pedler, B. E., L. I. Aluwihare, and F. Azam
2014. Single bacterial strain capable of significant contribution to carbon cycling in the surface ocean. *Proceedings of the National Academy of Sciences of the United States of America*, 111(20):7202–7207.

Pedrós-Alió, C.
2012. The rare bacterial biosphere. *Annual Review of Marine Science*, 4(1):449–466.

Pedrós-Alió, C., S. Acinas, R. Logares, and R. Massana
2018. Marine microbial diversity as seen by high-throughput sequencing. In *Microbial Ecology of the Oceans*, J. M. Gasol and D. L. Kirchman, eds., Pp. 47–98. Wiley-Blackwell.

Pesant, S., F. Not, M. Picheral, S. Kandels-Lewis, N. Le Bescot, G. Gorsky, D. Iudicone, E. Karsenti, S. Speich, R. Troublé, C. Dimier, S. Searson, T. O. C. Coordinators, S. G. Acinas, P. Bork, E. Boss, C. Bowler, C. D. Vargas, M. Follows, G. Gorsky, N. Grimsley, P. Hingamp, D. Iudicone, O. Jaillon, S. Kandels-Lewis, L. Karp-Boss, E. Karsenti, U. Krzic, F. Not, H. Ogata, S. Pesant, J. Raes, E. G. Reynaud, C. Sardet, M. Sieracki, S. Speich, L. Stemmann, M. B. Sullivan, S. Sunagawa, D. Velayoudon, J. Weissenbach, and P. Wincker
2015. Open science resources for the discovery and analysis of Tara Oceans data. *Scientific Data*, 2(1):150023.

Pinhassi, J., U. Zweifel, and A. Hagström
1997. Dominant marine bacterioplankton species found among colony-forming bacteria. *Applied and environmental microbiology*, 63(9):3359–3366.

Podar, M., C. C. Gilmour, C. C. Brandt, A. Soren, S. D. Brown, B. R. Crable, A. V. Palumbo, A. C. Somenahally, and D. A. Elias
2015. Global prevalence and distribution of genes and microorganisms involved in mercury methylation. *Science advances*, 1(9):e1500675.

Pomeroy, L., P. Williams, F. Azam, and J. Hobbie
2007. The microbial loop. *Oceanography*, 20(2):28–33.

Pommier, T., P. R. Neal, J. M. Gasol, M. Coll, S. G. Acinas, and C. Pedrós-Alió
2010. Spatial patterns of bacterial richness and evenness in the NW Mediterranean Sea explored by pyrosequencing of the 16S rRNA. *Aquatic Microbial Ecology*, 61(3):221–233.

Porter, K. G. and Y. S. Feig
1980. The use of DAPI for identifying and counting aquatic microflora1. *Limnology and Oceanography*, 25(5):943–948.

Pruesse, E., J. Peplies, and F. O. Glöckner
2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics (Oxford, England)*, 28(14):1823–1829.

Puig, P., A. Palanques, J. A. Sanchez-Cabeza, and P. Masqué
1999. Heavy metals in particulate matter and sediments in the southern Barcelona sedimentation system (North-western Mediterranean). *Marine Chemistry*, 63(3-4):311–329.

Pujalte, M., T. Lucena, L. Rodrigo-Torres, and D. Arahal
2018. Comparative genomics of Thalassobius including the description of Thalassobius activus sp. nov., and Thalassobius autumnalis sp. nov. *Frontiers in microbiology*, 8:2645.

Rabiner, L. and B. Juang
1986. An introduction to hidden Markov models. *IEEE ASSP Magazine*, 3(1):4–16.

Rahmstorf, S.
2003. Thermohaline circulation: The current climate. *Nature*, 421(6924):699–699.

Ramadan, E., M. Maged, A. E. Hosseiny, F. S. Chambergo, J. C. Setubal, and H. E. Dorry
2019. Molecular adaptations of bacterial mercuric reductase to the hypersaline kebrit deep in the Red Sea. *Appl. Environ. Microbiol.*, 85(4):e01431–18.

Rappé, M. S., S. A. Connon, K. L. Vergin, and S. J. Giovannoni
2002. Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature*, 418(6898):630–633.

Rasmussen, L., C. Zawadsky, S. Binnerup, G. Oregaard, S. Sørensen, and N. Kroer
2008. Cultivation of hard-to-culture subsurface mercury-resistant bacteria and discovery of new merA gene sequences. *Applied and environmental microbiology*, 74(12):3795–3803.

Raven, J. and P. G. Falkowski
1999. Oceanic sinks for atmospheric CO2. *Plant, Cell and Environment*, 22(6):741–755.

Razumov, A.
1932. The direct method of calculation of bacteria in water: comparison with the Koch method. *Mikrobiologija*, 1:131–146.

Robinson, C., D. K. Steinberg, T. R. Anderson, J. Arístegui, C. A. Carlson, J. R. Frost, J.-F. Ghiglione, S. Hernández-León, G. A. Jackson, R. Koppelmann, B. Quéguiner, O. Ragueneau, F. Rassoulzadegan, B. H. Robison, C. Tamburini, T. Tanaka, K. F. Wishner, and J. Zhang
2010. Mesopelagic zone ecology and biogeochemistry – a synthesis. *Deep Sea Research Part II: Topical Studies in Oceanography*, 57(16):1504–1518.

Robinson, J. and O. Tuovinen
1984. Mechanisms of microbial resistance and detoxification of mercury and organomercury compounds: physiological, biochemical, and genetic analyses. *Microbiological reviews*, 48(2):95–124.

Rodríguez Martín-Doimeadios, R. C., M. Monperrus, E. Krupp, D. Amouroux, and O. F. X. Donard
2003. Using speciated isotope dilution with GC–inductively coupled plasma MS to determine and unravel the artificial formation of monomethylmercury in certified reference sediments. *Analitycal Chemistry*, 75(13):3202–3211.

Ruiz González, C., R. Logares, M. Sebastián, M. Mestre, R. Rodríguez□Martínez, M. Galí, M. M. Sala, S. G. Acinas, C. M. Duarte, and J. M. Gasol
2019. Higher contribution of globally rare bacterial taxa reflects environmental transitions across the surface ocean. *Molecular Ecology*, 28(8):1930–1945.

Ruiz-González, C., M. Mestre, M. Estrada, M. Sebastián, G. Salazar, S. Agustí, E. Moreno-Ostos, I. Reche, X. A. Álvarez-Salgado, X. A. G. Morán, C. M. Duarte, M. M. Sala, and J. M. Gasol
2020. Major imprint of surface plankton on deep ocean prokaryotic structure and activity. *Molecular Ecology*, 29(10):1820–1838.

Sahm, K., C. Knoblauch, and R. Amann
1999. Phylogenetic affiliation and quantification of psychrophilic sulfate-reducing isolates in marine Arctic sediments. *Applied and environmental microbiology*, 65(9):3976–3981.

Saiz-Lopez, A., S. P. Sitkiewicz, D. Roca-Sanjuán, J. M. Oliva-Enrich, J. Z. Dávalos, R. Notario, M. Jiskra, Y. Xu, F. Wang, C. P. Thackray, E. M. Sunderland, D. J. Jacob, O. Travnikov, C. A. Cuevas, A. U. Acuña, D. Rivero, J. M. C. Plane, D. E. Kinnison, and J. E. Sonke
2018. Photoreduction of gaseous oxidized mercury changes global atmospheric mercury speciation, transport and deposition. *Nature Communications*, 9(1):4796.

Salazar, G.
2018a. EcolUtils: Utilities for community ecology analysis. R package version 0.1. https://github.com/GuillemSalazar/EcolUtils.

Salazar, G.
2018b. *Large-scale biogeography of marine pelagic Bacteria and Archaea. PhD Thesis.* PhD thesis, Universitat Politècnica de Barcelona and Institut de Ciències del Mar.

Salazar, G., F. Cornejo-Castillo, V. Benítez-Barrios, E. Fraile-Nuez, X. Álvarez-Salgado, C. Duarte, J. Gasol, and S. Acinas
2016. Global diversity and biogeography of deep-sea pelagic prokaryotes. *The ISME Journal*, 10:596–608.

Salazar, G., F. Cornejo-Castillo, E. Borrull, C. Díez-Vives, E. Lara, D. Vaqué, J. Arrieta, C. Duarte, J. Gasol, and S. Acinas
2015. Particle-association lifestyle is a phylogenetically conserved trait in bathypelagic prokaryotes. *Molecular Ecology*, 24:5692–5706.

Sambo, F., F. Finotello, E. Lavezzo, G. Baruzzo, G. Masi, E. Peta, M. Falda, S. Toppo, L. Barzon, and B. D. Camillo
2018. Optimizing PCR primers targeting the bacterial 16S ribosomal RNA gene. *BMC Bioinformatics*, 19:343.

Sánchez, O., J. M. Gasol, R. Massana, J. Mas, and C. Pedrós-Alió
2007. Comparison of different denaturing gradient gel electrophoresis primer sets for the study of marine bacterioplankton communities. *Applied and environmental microbiology*, 73(18):5962–5967.

Sanz-Sáez, I., G. Salazar, P. Sánchez, E. Lara, M. Royo-Llonch, E. L. Sà, T. Lucena, M. J. Pujalte, D. Vaqué, C. M. Duarte, J. M. Gasol, C. Pedrós-Alió, O. Sánchez, and S. G. Acinas
2020. Diversity and distribution of marine heterotrophic bacteria from a large culture collection. *BMC Microbiology*, 20(1):207.

Saranya, K., A. Sundaramanickam, S. Shekhar, S. Swaminathan, and T. Balasubramanian
2017. Bioremediation of mercury by Vibrio fluvialis screened from industrial effluents. *BioMed research international*, 2017:6509648.

Sass, A., H. Sass, M. Coolen, H. Cypionka, and J. Overmann
2001. Microbial communities in the chemocline of a hypersaline deep-sea basin (Urania basin, Mediterranean Sea). *Applied and environmental microbiology*, 67(12):5392–5402.

Sasser, M.
1990. Identification of bacteria by gas chromatography of cellular fatty acids, MIDI Technical Note 101. Technical report, MIDI Inc., Newark.

Scanlan, D., M. Ostrowski, S. Mazard, A. Dufresne, L. Garczarek, W. Hess, A. Post, M. Hagemann, I. Paulsen, and F. Partensky
2009. Ecological genomics of marine Picocyanobacteria. *Microbiology and molecular biology reviews*, 73(2):249–299.

Schaefer, J. K., J. Yagi, J. R. Reinfelder, T. Cardona, K. M. Ellickson, S. Tel-Or, and T. Barkay
2004. Role of the bacterial organomercury lyase (MerB) in controlling methylmercury accumulation in mercury-contaminated natural waters. *Environmental Science and Technology*, 38(16):4304–4311.

Schauer, R., C. Bienhold, A. Ramette, and J. Harder
2010. Bacterial diversity and biogeography in deep-sea surface sediments of the South Atlantic Ocean. *The ISME journal*, 4(2):159–170.

Schut, F., E. J. De Vries, J. C. Gotitschal, B. R. Robertson, W. Harder, R. A. Prins, and D. K. Button
1993. Isolation of typical marine bacteria by dilution culture: Growth, maintenance, and characteristics of isolates under laboratory conditions. *Applied and environmental microbiology*, 59(7):2150–2160.

Schut, F., R. A. Prins, and J. C. Gottscha
1997. Oligotrophy and pelagic marine bacteria: facts and fiction. *Aquatic Microbial Ecology*, 2:177–202.

Sebastián, M., J.-C. Auguet, C. X. Restrepo-Ortiz, M. M. Sala, C. Marrasé, and J. M. Gasol
2017. Deep ocean prokaryotic communities are remarkably malleable when facing long-term starvation. *Environmental Microbiology*, 20(2):713–723.

Seemann, T.
2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics (Oxford, England)*, 30(14):2068–2069.

Sekiguchi, T., T. Sato, M. Enoki, H. Kanehiro, K. Umematsu, and C. Kato
2010. Isolation and characterization of biodegradable plastic degrading bacteria. *JAMSTEC Report of Research and Development*, 11(September):33–41.

Selje, N., T. Brinkhoff, and M. Simon
2005. Detection of abundant bacteria in the Weser estuary using culture-dependent and culture-independent approaches. *Aquatic Microbial Ecology*, 39(1):17–34.

Shade, A., C. S. Hogan, A. K. Klimowicz, M. Linske, P. S. McManus, and J. Handelsman
2012. Culturing captures members of the soil rare biosphere. *Environmental Microbiology*, 14(9):2247–2252.

Shannon, C. E.
1948. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423.

Sharma, S. and S. Chatterjee
2017. Microplastic pollution, a threat to marine ecosystem and human health: a short review. *Environmental Science and Pollution Research*, 24(27):21530–21547.

Simbahan, J., E. Kurth, J. Schelert, A. Dillman, E. Moriyama, S. Jovanovich, and P. Blum
2005. Community analysis of a mercury hot spring supports occurrence of domain-specific forms of mercuric reductase. *Applied and environmental microbiology*, 71(12):8836–8845.

Singer, E., E. A. Webb, W. C. Nelson, J. F. Heidelberg, N. Ivanova, A. Pati, and K. J. Edwards
2011. Genomic potential of Marinobacter aquaeolei, a biogeochemical 'opportunitroph'. *Applied and Environmental Microbiology*, 77(8):2763–2771.

Smith, D. C., M. Simon, A. L. Alldredge, and F. Azam
1992. Intense hydrolytic enzyme activity on marine aggregates and implications for rapid particle dissolution. *Nature*, 359(6391):139–142.

Soerensen, A. L., E. M. Sunderland, C. D. Holmes, D. J. Jacob, R. M. Yantosca, H. Skov, J. H. Christensen, S. A. Strode, and R. P. Mason
2010. An improved global model for air-sea exchange of mercury: High concentrations over the North Atlantic. *Environmental Science & Technology*, 44(22):8574–8580.

Sogin, M. L., H. G. Morrison, J. A. Huber, D. M. Welch, S. M. Huse, P. R. Neal, J. M. Arrieta, and G. J. Herndl
2006. Microbial diversity in the deep sea and the underexplored 'rare biosphere'. *Proceedings of the National Academy of Sciences of the United States of America*, 103(32):12115–12120.

Sotero-Martins, A., M. Silva De Jesus, M. Lacerda, J. C. Moreira, A. Luzia, L. Filgueiras, P. Rubens, G. Barrocas, F. Oswaldo, and C. Escola
2008. A conservative region of the mercuric reductase gene (merA) as a molecular marker of bacterial mercury resistance. *Brazilian Journal of Microbiology*, 39:307–310.

Sprouffske, K. and A. Wagner
2016. Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinformatics*, 17(1):172.

Stackebrandt, E. and J. Ebers
2006. Taxonomic parameters revisited: tarnished gold standards. *Microbiology Today*, 33:152–155.

Staley, J. and A. Konopka
1985. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annual Review of Microbiology*, 39(1):321–346.

Steen, A. D., A. Crits-Christoph, P. Carini, K. M. Deangelis, N. Fierer, K. G. Lloyd, and J. C. Thrash
2019. High proportions of bacteria and archaea across most biomes remain uncultured. *The ISME Journal*, 13:3126–3130.

Stefani, F. O. P., T. H. Bell, C. Marchand, I. E. de la Providencia, A. El Yassimi, M. St-Arnaud, and M. Hijri
2015. Culture-dependent and -independent methods capture different microbial community fractions in hydrocarbon-contaminated soils. *PLOS ONE*, 10(6):e0128272.

Steffen, W., P. J. Crutzen, and J. R. McNeill Ambio
2007. The anthropocene: Are humans now overwhelming the great forces of nature? *Ambio*, 36(8):614–621.

Stepanauskas, R. and M. E. Sieracki
2007. Matching phylogeny and metabolism in the uncultured marine bacteria, one cell at a time. *Proceedings of the National Academy of Sciences of the United States of America*, 104(21):9052–9057.

Sul, W. J., T. A. Oliver, H. W. Ducklow, L. A. Amaral-Zettler, and M. L. Sogin
2013. Marine bacteria exhibit a bipolar distribution. *Proceedings of the National Academy of Sciences of the United States of America*, 110(6):2342–2347.

Sunagawa, S., S. G. Acinas, P. Bork, C. Bowler, S. G. Acinas, M. Babin, P. Bork, E. Boss, C. Bowler, G. Cochrane, C. de Vargas, M. Follows, G. Gorsky, N. Grimsley, L. Guidi, P. Hingamp, D. Iudicone, O. Jaillon, S. Kandels, L. Karp-Boss, E. Karsenti, M. Lescot, F. Not, H. Ogata, S. Pesant, N. Poulton, J. Raes, C. Sardet, M. Sieracki, S. Speich, L. Stemmann, M. B. Sullivan, S. Sunagawa, P. Wincker, D. Eveillard, G. Gorsky, L. Guidi, D. Iudicone, E. Karsenti, F. Lombard, H. Ogata, S. Pesant, M. B. Sullivan, P. Wincker, and C. de Vargas
2020. Tara Oceans: towards global ocean ecosystems biology. *Nature Reviews Microbiology*, 18(8):428–445.

Sunagawa, S., L. P. Coelho, S. Chaffron, J. R. Kultima, K. Labadie, G. Salazar, B. Djahanschiri, G. Zeller, D. R. Mende, A. Alberti, F. M. Cornejo-Castillo, P. I. Costea, C. Cruaud, F. D'Ovidio, S. Engelen, I. Ferrera, J. M. Gasol, L. Guidi, F. Hildebrand, F. Kokoszka, C. Lepoivre, G. Lima-Mendez, J. Poulain, B. T. Poulos, M. Royo-Llonch, H. Sarmento, S. Vieira-Silva, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis, E. Boss, M. Follows, L. Karp-Boss, U. Krzic, E. G. Reynaud, C. Sardet, M. Sieracki, D. Velayoudon, C. Bowler, C. De Vargas, G. Gorsky, N. Grimsley, P. Hingamp, D. Iudicone, O. Jaillon, F. Not, H. Ogata, S. Pesant, S. Speich, L. Stemmann, M. B. Sullivan, J. Weissenbach, P. Wincker, E. Karsenti, J. Raes, S. G. Acinas, and P. Bork
2015. Structure and function of the global ocean microbiome. *Science*, 348(6237):1261359.

Sunderland, E. M., D. P. Krabbenhoft, J. W. Moreau, S. A. Strode, and W. M. Landing
2009. Mercury sources, distribution, and bioavailability in the North Pacific Ocean: Insights from data and models. *Global Biogeochemical Cycles*, 23(2):GB2010.

Sunderland, E. M. and R. P. Mason
2007. Human impacts on open ocean mercury concentrations. *Global Biogeochemical Cycles*, 21(4):GB4022.

Sutton, T. T., M. R. Clark, D. C. Dunn, P. N. Halpin, A. D. Rogers, J. Guinotte, S. J. Bograd, M. V. Angel, J. A. A. Perez, K. Wishner, R. L. Haedrich, D. J. Lindsay, J. C. Drazen, A. Vereshchaka, U. Piatkowski, T. Morato, K. Błachowiak-Samołyk, B. H. Robison, K. M. Gjerde, A. Pierrot-Bults, P. Bernal, G. Reygondeau, and M. Heino
2017. A global biogeographic classification of the mesopelagic zone. *Deep Sea Research Part I: Oceanographic Research Papers*, 126:85–102.

Suzuki, M. T., M. S. Rappé, Z. W. Haimberger, H. Winfield, N. Adair, J. Ströbel, and S. J. Giovannoni
1997. Bacterial diversity among small-subunit rRNA gene clones and cellular isolates from the same seawater sample. *Applied and environmental microbiology*, 63(3):983–989.

Swan, B. K., B. Tupper, A. Sczyrba, F. M. Lauro, M. Martinez-Garcia, J. M. González, H. Luo, J. J. Wright, Z. C. Landry, N. W. Hanson, B. P. Thompson, N. J. Poulton, P. Schwientek, S. G. Acinas, S. J. Giovannoni, M. A. Moran, S. J. Hallam, R. Cavicchioli, T. Woyke, and R. Stepanauskas
2013. Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. *Proceedings of the National Academy of Sciences of the United States of America*, 110(28):11463–11468.

Tabor, P. S., K. Ohwada, and R. R. Colwell
1981. Filterable marine bacteria found in the deep sea: Distribution, taxonomy, and response to starvation. *Microbial Ecology*, 7(1):67–83.

Tait, R. and F. Dipper
1998a. The Oceans. In *Elements of Marine Ecology*, R. Tait and F. Dipper, eds., Pp. 1–24. Butterworth-Heinemann.

Tait, R. and F. Dipper
1998b. The seawater habitat—physical and chemical conditions. In *Elements of marine ecologycology*, R. Tait and F. Dipper, eds., Pp. 96–165. Butterworth-Heinemann.

Talley, L. D., G. L. Pickard, W. J. Emery, and J. H. Swift
2011. Typical distributions of water characteristics. In *Descriptive Physical Oceanography*, L. D. Talley, G. L. Pickard, W. J. Emery, and J. H. Swift, eds., Pp. 67–110. Academic Press.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar
2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution*, 28(10):2731–9.

Tanaka, T., K. Kawasaki, S. Daimon, W. Kitagawa, K. Yamamoto, H. Tamaki, M. Tanaka, C. H. Nakatsu, and Y. Kamagata
2014. A hidden pitfall in the preparation of agar media undermines microorganism cultivability. *Applied and Environmental Microbiology*, 80(24):7659–7666.

Teira, E., P. Lebaron, H. van Aken, and G. J. Herndl
2006. Distribution and activity of Bacteria and Archaea in the deep water masses of the North Atlantic. *Limnology and Oceanography*, 51(5):2131–2144.

Terahara, T., K. Yamada, J. Nakayama, Y. Igarashi, T. Kobayashi, and C. Imada
2016. Bacterial community structures of deep-sea water investigated by molecular biological techniques. *Gene*, 576(2):696–700.

Thrash, J. C.
2020. Towards culturing the microbe of your choice. *Environmental Microbiology Reports*, Pp. 1758–2229.12898.

Tonon, L. A. C., A. P. B. Moreira, and F. Thompson
2014. The family Erythrobacteraceae. In *The Prokaryotes*, E. Rosenberg, D. Ef, S. Lory, E. Stackebrandt, and F. Thompson, eds., Pp. 213–235. Berlin, Heidelberg: Springer Berlin Heidelberg.

Torsvik, V. and L. Øvreås
2002. Microbial diversity and function in soil: from genes to ecosystems. *Current Opinion in Microbiology*, 5(3):240–245.

Tringe, S. G., C. von Mering, A. Kobayashi, A. A. Salamov, K. Chen, H. W. Chang, M. Podar, J. M. Short, E. J. Mathur, J. C. Detter, P. Bork, P. Hugenholtz, and E. M. Rubin
2005. Comparative metagenomics of microbial communities. *Science*, 308(5721):554–557.

Tsygankov, V. Y.
2019. Organochlorine pesticides in marine ecosystems of the Far Eastern Seas of Russia (2000–2017). *Water Research*, 161:43–53.

Tyson, G. W., I. Lo, B. J. Baker, E. E. Allen, P. Hugenholtz, and J. F. Banfield
2005. Genome-directed isolation of the key nitrogen fixer Leptospirillum ferrodiazotrophum sp. nov. from an acidophilic microbial community. *Applied and environmental microbiology*, 71(10):6319–6324.

Uddin, S., S. W. Fowler, M. Behbehani, A. N. Al-Ghadban, P. W. Swarzenski, and N. Al-Awadhi
2020. A review of radioactivity in the Gulf region. *Marine Pollution Bulletin*, 159(July):111481.

UN Environment
2019. Global mercury assesment 2018. Technical report, UN Environment Programme, Chemicals and Health Brancg, UN Environment Programme, Chemicals and Health Branch Geneva, Switzerland.

Venter, J. C., K. Remington, J. F. Heidelberg, A. L. Halpern, D. Rusch, J. A. Eisen, D. Wu, I. Paulsen, K. E. Nelson, W. Nelson, D. E. Fouts, S. Levy, A. H. Knap, M. W. Lomas, K. Nealson, O. White, J. Peterson, J. Hoffman, R. Parsons, H. Baden-Tillson, C. Pfannkoch, Y.-H. Rogers, and H. O. Smith
2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science (New York, N.Y.)*, 304(5667):66–74.

Vetriani, C., Y. S. Chew, S. M. Miller, J. Yagi, J. Coombs, R. A. Lutz, and T. Barkay
2005. Mercury adaptation among bacteria from a deep-sea hydrothermal vent. *Applied and environmental microbiology*, 71(1):220–226.

Vichi, M., J. I. Allen, S. Masina, and N. J. Hardman-Mountford
2011. The emergence of ocean biogeochemical provinces: A quantitative assessment and a diagnostic for model evaluation. *Global Biogeochemical Cycles*, 25(2):GB2005.

Villar, E., G. K. Farrant, M. Follows, L. Garczarek, S. Speich, S. Audic, L. Bittner, B. Blanke, J. R. Brum, C. Brunet, R. Casotti, A. Chase, J. R. Dolan, F. D'Ortenzio, J.-P. Gattuso, N. Grima, L. Guidi, C. N. Hill, O. Jahn, J.-L. Jamet, H. Le Goff, C. Lepoivre, S. Malviya, E. Pelletier, J.-B. Romagnan, S. Roux, S. Santini, E. Scalco, S. M. Schwenck, A. Tanaka, P. Testor, T. Vannier, F. Vincent, A. Zingone, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis, S. G. Acinas, P. Bork, E. Boss, C. de Vargas, G. Gorsky, H. Ogata, S. Pesant, M. B. Sullivan, S. Sunagawa, P. Wincker, E. Karsenti, C. Bowler, F. Not, P. Hingamp, D. Iudicone, and D. Iudicone
2015. Environmental characteristics of Agulhas rings affect interocean plankton transport. *Science*, 348(6237):1261447–1261447.

Voss, M., H. W. Bange, J. W. Dippner, J. J. Middelburg, J. P. Montoya, and B. Ward
2013. The marine nitrogen cycle: recent discoveries, uncertainties and the potential relevance of climate change. *Philosophical transactions of the Royal Society of London.*, 368(1621):20130121.

Wang, F., R. W. Macdonald, D. A. Armstrong, and G. A. Stern
2012. Total and methylated mercury in the Beaufort Sea: The role of local and recent organic remineralization. *Environmetal Science and Technology*, 46:11821–11828.

Wang, H., J. Tomasch, M. Jarek, and I. Wagner-Döbler
2014. A dual-species co-cultivation system to study the interactions between Roseobacters and dinoflagellates. *Frontiers in Microbiology*, 5:311.

Ward, D., R. Weller, and M. Bateson
1990. 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature*, 345(6270):63–65.

Webb, C. O., D. D. Ackerly, M. A. McPeek, and M. J. Donoghue
2002. Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, 33(1):475–505.

Webb, P.
2020. *Introduction to cceanography*. Creative Commons Attribution 4.0 Unported License also.

Wells, N.
2015. OCEANOGRAPHIC TOPICS | General Processes. In *Encyclopedia of Atmospheric Sciences*, G. R. North, J. Pyle, and F. Zhang, eds., Pp. 290–300. Academic Press.

Wenning, R. J. and L. Martello
2014. POPs in marine and freshwater environments. In *Environmental forensics for persistent organic pollutants*, G. O'Sullivan and C. Sandau, eds., Pp. 357–390. Elsevier.

Whitman, W., D. Coleman, and W. Wiebe
1998. Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences of the United States of America*, 95(12):6578–6583.

Wickham, H.
2019. Welcome to the tidyverse. *Journal of Open Source Software*, 4(43):1686.

Wiegand, I., K. Hilpert, and R. E. Hancock
2008. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nature Protocols*, 3(2):163–175.

Wiener, J., D. Krabbenhoft, G. Heinz, and A. Scheuhammer
2002. Ecotoxicology of mercury. In *Handbook of Ecotoxicology, Second Edition*, D. Hoffman, B. Rattner, G. Burton, and J. Cairns, eds., Pp. 409–464. Lewis Publishers.

Wilhelmsson, D., R. C. Thompson, K. Holmström, O. Lindén, and H. Eriksson-Hägg
2013. Marine Pollution. In *Managing ocean environments in a changing climate*, chapter Chapter 6, Pp. 127–169. Elsevier.

Woese, C. and G. Fox
1977. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences of the United States of America*, 74(11):5088–5090.

Woese, C., O. Kandler, and M. Wheelis
1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences of the United States of America*, 87(12):4576–4579.

Wurch, L., R. J. Giannone, B. S. Belisle, C. Swift, S. Utturkar, R. L. Hettich, A.-L. Reysenbach, and M. Podar
2016. Genomics-informed isolation and characterization of a symbiotic Nanoarchaeota system from a terrestrial geothermal environment. *Nature communications*, 7:12115.

Yang, S.-J., I. Kang, and J.-C. Cho
2016. Expansion of cultured bacterial diversity by large-scale dilution-to-extinction culturing from a single seawater sample. *Microbial Ecology*, 71(1):29–43.

Yarza, P., P. Yilmaz, E. Pruesse, F. O. Glöckner, W. Ludwig, K.-H. Schleifer, W. B. Whitman, J. Euzéby, R. Amann, and R. Rosselló-Móra
2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature reviews. Microbiology*, 12(9):635–645.

Yuan, J., Q. Lai, F. Sun, T. Zheng, and Z. Shao
2015. The diversity of PAH-degrading bacteria in a deep-sea water column above the Southwest Indian ridge. *Frontiers in Microbiology*, 6(AUG):853.

Yılmaz, A. B., A. Yanar, and E. N. Alkan
2017. Review of heavy metal accumulation on aquatic environment in Northern East Mediterrenean Sea part I: some essential metals. *Rev Environ Health*, 32:119–163.

Zeder, M., A. Ellrott, and R. Amann
2011. Automated sample area definition for high-throughput microscopy. *Cytometry Part A*, 79A(4):306–310.

Zeder, M. and J. Pernthaler
2009. Multispot live-image autofocusing for high-throughput microscopy of fluorescently stained bacteria. *Cytometry Part A*, 75A(9):781–788.

Zehr, J. P., S. R. Bench, B. J. Carter, I. Hewson, F. Niazi, T. Shi, H. J. Tripp, and J. P. Affourtit
2008. Globally distributed uncultivated oceanic N2-fixing cyanobacteria lack oxygenic photosystem II. *Science*, 322(5904):1110–1112.

Zehr, J. P. and B. B. Ward
2002. Nitrogen cycling in the ocean: new perspectives on processes and paradigms. *Applied and environmental microbiology*, 68(3):1015–1024.

Zeng, Y., Y. Zou, J. M. Grebmeier, J. He, and T. Zheng
2012. Culture-independent and -dependent methods to investigate the diversity of planktonic bacteria in the northern Bering Sea. *Polar Biology*, 35(1):117–129.

Zengler, K., G. Toledo, M. Rappe, J. Elkins, E. J. Mathur, J. M. Short, and M. Keller
2002. Cultivating the uncultured. *Proceedings of the National Academy of Sciences of the United States of America*, 99(24):15681–15696.

Zengler, K., M. Walcher, G. Clark, I. Haller, G. Toledo, T. Holland, E. J. Mathur, G. Woodnutt, J. M. Short, and M. Keller
2005. High□throughput cultivation of microorganisms using microcapsules. *Methods in Enzymology*, 397:124–130.

Zeyaullah, M., B. Islam, and A. Ali
2009. Isolation, identification and PCR amplification of merA gene from highly mercury polluted Yamuna river. *African Journal of Biotechnology*, 9(24):3510–3514.

Zhang, W., L. Chen, and D. Liu
2012. Characterization of a marine-isolated mercury-resistant Pseudomonas putida strain SP1 and its potential application in marine mercury reduction. *Applied Microbiology and Biotechnology*, 93(3):1305–1314.

Zheng, R., S. Wu, N. Ma, and C. Sun
2018. Genetic and physiological adaptations of marine bacterium Pseudomonas stutzeri 273 to mercury stress. *Frontiers in microbiology*, 9:682.

Zinger, L., L. A. Amaral-Zettler, J. A. Fuhrman, M. C. Horner-Devine, S. M. Huse, D. B. M. Welch, J. B. Martiny, M. Sogin, A. Boetius, and A. Ramette
2011. Global patterns of bacterial beta-diversity in seafloor and seawater ecosystems. *PloS one*, 6(9):e24570.

Zinger, L., A. Boetius, and A. Ramette
2014. Bacterial taxa-area and distance-decay relationships in marine environments. *Molecular ecology*, 23(4):954–964.

Zobell, C. E.
1941. Studies on marine bacteria. I. The requirements of heterotrophic aerobes. *Journal of Marine Research*, 4:42–75.

Zobell, C. E. and R. Y. Morita
1957. Barophilic bacteria in some deep sea sediments. *Journal of bacteriology*, 73(4):563–568.