Marine Genetic Resources in Areas Beyond National Jurisdiction: Promoting Marine Scientific Research and Enabling Equitable Benefit Sharing: Supplementary Material

S1.0 Examples of marine genetic resources

Genetic Resources for Pharmaceuticals

Natural products play a dominant role in the discovery and development of new drugs, especially antibiotic and cytotoxic agents (Newman & Cragg, 2016). To date, 13 pharmaceutical drugs derived from marine organisms have been clinically approved, including treatments for cancer, neuropathic pain, an antiviral (against *Herpes simplex*), and for treatment of hypertriglyceridemia (Blasiak et al., 2020a; Midwestern University, 2020). A further 24 marine-derived products are in clinical trials and a further 250 in preclinical investigations (Blasiak et al., 2020a; Midwestern University, 2020).

Cytarabine, also known as ara-C, provides an early example of a drug derived from a marine organism (Blasiak et al., 2020a). This drug was identified in a Caribbean marine sponge, *Tectitethya crypta* (Rogers, 2019). Werner Bergman, a scientist studying steroid-like compounds in marine organisms found boiling the sponge yielded a large amount of a crystalline substance. The physical properties of this substance resembled those of thymidine, a component of DNA, and he named it spongothymidine. He found another substance, which he called spongouridine, with similar properties to uridine, a component of RNA. These substances are nucleosides and inhibit the replication of DNA and RNA, a form of chemical defence. An artificial form of the chemical, arabinosyl cytosine, was synthesized and found to inhibit cancer cells in rats. This substance was licensed as a drug for cancer treatment, Cytosar-U® in 1969 (Blasiak et al., 2020a). It is now used as a cell replication inhibitor in the treatment of several leukaemias and non-Hodgkin's lymphoma (Rogers, 2019; Blasiak et al., 2020a).

Figure S1. (a) *Tectitethya crypta* (Zea et al., 2014) and (b) the drug Cytarabine (Pfizer, Netherlands; https://www.pfizer.nl/product/cytarabine-hospira).

Genetic Resources for Research Tools

Marine organisms have been the source of important chemicals used in research. For example, green fluorescent protein (GFP) was discovered by Osamu Shimomura who was studying bioluminescence in the jellyfish *Aequorea victoria.* He and his colleagues collected over 10,000 specimens of the jellyfish from Puget Sound (USA) and collected the proteins

responsible for the bioluminescence from the edge of the umbrellas. The bioluminescent protein was called aequorin and glowed blue, but a second protein fluoresced green under blue or UV light. After thirty years of study, researchers discovered the DNA sequence for GFP, a small protein produced by a single gene. This discovery allowed artificial insertion of the gene next to other genes in an organism so that the GFP gene expressed whenever and wherever the adjacent gene was expressed, fluorescing green when exposed to blue or ultraviolet light. The uses of GFP in tracking where and when a gene was actively transcribed represented an enormous scientific advance and won Shimomura the Nobel Prize for Chemistry in 2008. GFPs have been used, for example, to trace the circuitry of the brain, the entry of viruses or bacteria into tissues and cells and the regeneration of organs, such as the kidney (Rogers, 2019). They have also helped to track the processes involved in growth and development of organisms (Rogers, 2019). Other examples of natural product chemicals enabling the advancement of research derived from marine organisms include high-fidelity Taq-polymerases isolated from deep-sea microorganisms, used in a wide range of applications for molecular biology to replicate DNA in the polymerase chain reaction (Takagi et al., 1997; Terpe 2013; Hikida et al., 2017). Many marine organisms also produce chemicals, usually toxins, which selectively block a variety of receptors and voltage-gated channels across cell membranes and therefore are used as research tools (e.g. conotoxins; Bjørn-Yoshimoto et al., 2020). It is interesting to note that a significant proportion of products developed from marine species have been used for molecular and functional biology applications.

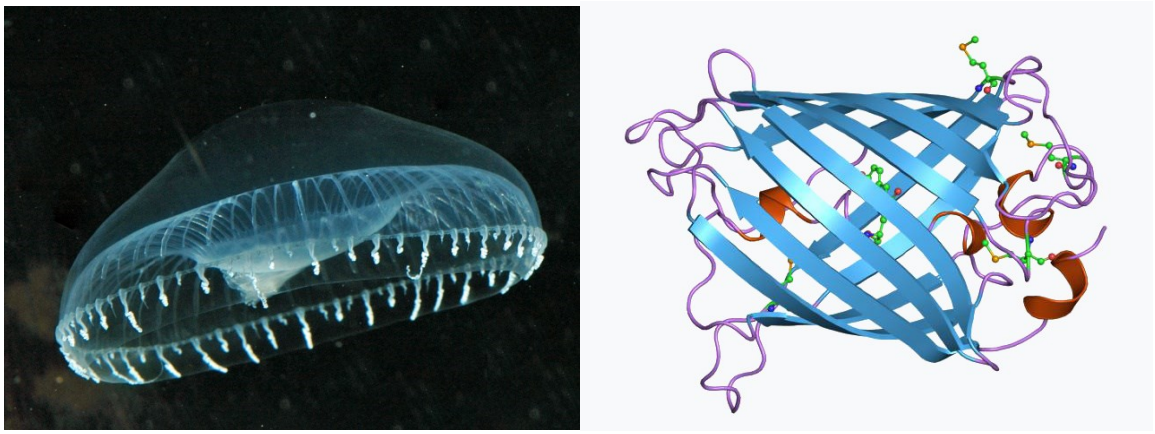

Figure S2. (a) The jellyfish *Aequorea victoria* (Sierra Blakely c/o Wikimedia Commons) and (b) a diagram of the GFP molecule (Jawahar Swaminathan and MSD Staff at http://www.ebi.ac.uk c/o Wikimedia Commons).

Genetic Resources for Industrial Processes

Many industrial processes currently utilize enzymes because they can achieve complex chemical reactions under benign conditions, using water as a solvent, thus reducing chemical and energy wastage. The global market for industrial enzymes was expected to reach US$ 7 Bn by 2018, with a compound annual growth rate of 8.2% between 2013 and 2018 (Sarmiento et al., 2015). Cold adapted enzymes – those that function at low temperature, thus reducing cellular energy needed by lowering the temperature from perhaps 30 ºC to that of cold water (8-15 ºC) - have attracted particular interest. Another favourable property of these enzymes is easy inactivation of further activity when desired by raising the temperature. Scientists discovered these enzymes by investigating extremophiles, those organisms, particularly bacteria and archaea, that thrive in low temperature environments such as earth's polar regions (Bruno et al., 2019). The lowest temperature at which an extremophile has been observed to grow and reproduce is -15 ºC,

but numerous enzymes derived from polar organisms are currently or may potentially be used in industrial processes (see Bruno et al., 2019). Industries can also utilize high temperature enzymes obtained from organisms inhabiting deep-sea hydrothermal vents (e.g., hydrolases such as amylases, glycosidases and proteases; Dalmaso et al., 2015).

Figure S3. Extreme environments in the ocean: (a) Arctic sea ice (rapidly disappearing with climate change; AD Rogers) and (b) deep-sea high temperature hydrothermal vents from the East Scotia Ridge, Southern Ocean (NERC CHESSO project).

Management and Conservation

A variety of management and conservation purposes require better understanding of genetic variation within and among species. It is estimated, for example that up to 90% of marine species are undescribed (e.g. Mora et al., 2011) and genetic tools have now become critical in identification of marine species and understanding how they fit within the Tree of Life (e.g. for hydroids, Moura et al., 2018). For many groups of marine organisms, especially including macroalgae and many invertebrates, DNA sequencing is routinely used for identifying and classifying novel species (e.g. Abdelkrim et al., 2018; Schneider et al., 2019). For microorganisms, including microbial eukaryotes, bacteria and Archaea, identification, classification and quantification are almost solely achieved through molecular genetic approaches (see Gasol and Kirchman, 2018; Pedrós-Alió et al., 2018). Understanding of the functional ecology of microbial taxa is also undertaken through genomic approaches (e.g. identification of functional genes from environmental samples and their patterns of expression; Gasol and Kirchman, 2018; Morris, 2018). Environmental DNA (eDNA) approaches are now being used to establish the species richness of communities (see Ruppert et al., 2019 for review) or for the identification of specific taxa (e.g. identifying harmful algal bloom [HAB] species, Liu et al., 2020; or species of conservation concern such as manatees, Hunter et al., 2018, and European eels, Cardás et al., 2020). They have also been used for studying the ecology of species, for example, in analysis of diet (Ruppert et al., 2019). Important aspects of species biology, such as mating patterns and reproductive systems can also be assessed using genetic markers (e.g.sexual versus asexual reproduction; paternity studies; Frankham et al., 2010). For many commercial species and also species of conservation concern it is important to identify discrete populations as these are often adapted to distinct environments and therefore have different patterns of growth, maturation and other biological attributes. Such populations have to be managed as separate entities because they respond to exploitation or other forms of mortality differently (Carvalho and Hauser, 1995). For species of conservation concern DNA sequencing and use of genetic markers are important in identifying issues with inbreeding and loss of genetic diversity, accumulation of deleterious mutations, population fragmentation, hybridization and for forensics (e.g. detection of illegal hunting of threatened species; Frankham et al., 2010).

They can also be used to study evolutionary processes in invasive species (Frankham et al., 2010; Ruppert et al., 2019). Understanding genetic connectivity is also critical in the design of networks of marine protected areas to ensure the persistence and recovery of marine populations, communities and ecosystems through dispersal of eggs/spores, larvae, juveniles and adult organisms (e.g. Carr et al., 2017; Balbar and Metaxas, 2019).

An example of the use of genetic data in the management and conservation of marine species is in resolving the misidentification of the European common skate, *Dipturis batis*, which was listed as critically endangered by the IUCN in 2006. This species once occurred commonly in the Northern Atlantic Ocean and Mediterranean Sea, especially in bottom-trawl catches. World landings decreased to 502 tonnes in 2005, but a study in 2006/2007 noted inconsistencies in species identification. A closer study of the species morphology revealed that the common skate may indeed be two distinct species, provisionally named the blue skate (*Dipturis* cf. *flossada*) and the flapper skate (*Dipturis* cf. *intermedia*) which have markedly different maximum sizes (143.2 and 228.8 cm respectively), growth rates, and age of sexual maturity. Besides morphological analysis, this study also extracted DNA from muscle samples, amplified and sequenced parts of the mitochondrial DNA, and built a phylogenetic tree showing the relationship between the species analysed (Figure 4). As is standard practice for species descriptions, the researchers deposited these data in GenBank (NCBI) and placed the samples from which they derived the digital sequence information in the Museum of Natural History Paris (MNHN) for reference. The latter (deposition of voucher specimens from which genetic data are derived) is not done commonly but is something we recommend as best practice (see Pleijel et al., 2008). The phylogenetic tree clearly shows that the two species confused for a single species are not even closely related. Once correctly assigned, the 2005 catch contained 8,300 adults of the smaller *Dipturis* cf. *flossada* and only 140 adult *Dipturis* cf. *intermedia.* The rarity of the latter, larger species is consistent with the lower rates of reproduction observed for larger-bodied species, which is therefore less able to sustain exploitation. This example illustrates the critical need for clear taxonomic identification and genetic determination or assessment of genetic diversity; and the need for species descriptions based both on morphology and molecular data.

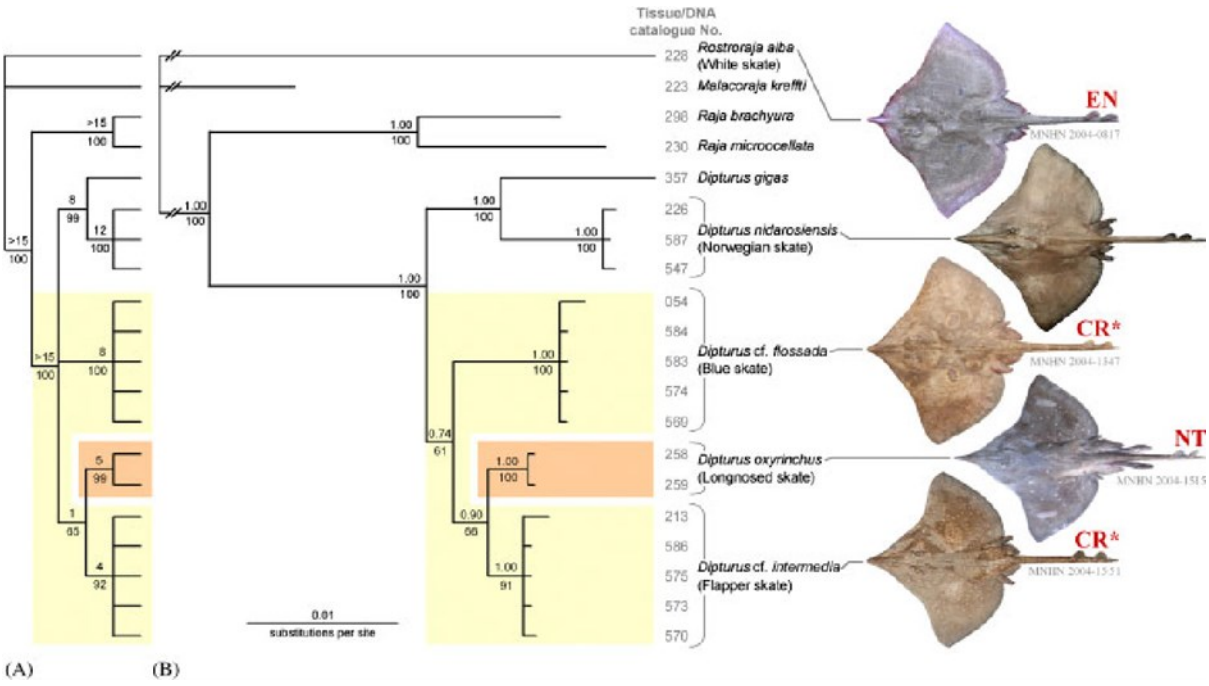

Figure S4. Phylogenetic tree showing relationships between different species of skate (Figure reproduced with permission from Iglésias et al., 2010).

S2.0 Monetary and non-monetary benefits

Discussions around the BBNJ agreement have identified the potential for both monetary and non-monetary benefits to arise from research on MGR. 'Monetary' benefits might come from commercial products derived from MGR from ABNJ and could be used for various purposes, although they have been linked to funding of conservation and sustainable use of biodiversity. 'Non-monetary' benefits include a range of capacity-building activities such as training in marine science, science co-production, better cooperation and coordination of scientific activities in ABNJ, transfer or sharing of marine science infrastructure (e.g. research vessels, observatories or advanced laboratories), and the sharing of technology, data, collections and scientific knowledge (Cincin-Sain et al., 2018). Importantly, many non-monetary benefits incur substantial costs, although who should bear these remains unclear at present and discussions on options for funding mechanisms are only beginning.

S3.0 Definition of MGRs in the ILBI

The present draft of the ILBI includes three alternative definitions proposed for MGRs in Article 1:

[8. "Marine genetic material" means any material of marine plant, animal, microbial or other origin containing functional units of heredity.]

[9. Alt. 1. "Marine genetic resources" means any material of marine plant, animal, microbial or other origin, [found in or] originating from areas beyond national jurisdiction and containing functional units of heredity with actual or potential value of their genetic and biochemical properties.]

[9. Alt. 2. "Marine genetic resources" means marine genetic material of actual or potential value.]

These definitions can be compared with the definitions in the Convention of Biological Diversity (CBD) Article 2:

"Biological resources" includes genetic resources, organisms or parts thereof, populations, or any other biotic component of ecosystems with actual or potential use or value for humanity.

"Genetic material" means any material of plant, animal, microbial or other origin containing functional units of heredity.

and

"Genetic resources" means genetic material of actual or potential value.

Compared to the definition of biological resources in the CBD, the various definitions for MGRs in the ILBI intend to distinguish MGRs from other living resources in the ocean, such as fish or other organisms harvested for food. The three alternative definitions differ with 9. Alt. 2, which specifically defines MGR only as genetic material of actual or potential value compared to the other two definitions. The alternatives 8. and 9. Alt. 1 refer to any material coming from marine organisms containing functional units of heredity, with the latter appearing to offer a broader definition that includes genetic and biochemical properties. Depending on the interpretation of the wording of these paragraphs, the former could be interpreted as excluding derivatives from marine organisms as forming MGRs. "Derivative" was defined in the Nagoya Protocol Article 2(e) to mean "a naturally occurring biochemical

compound resulting from the genetic expression or metabolism of biological or genetic resources, even if it does not contain functional units of heredity".

In all the definitions above, some terms are specifically defined (e.g., genetic material, DNA or RNA), whilst others lack a clear definition (e.g. actual or potential value; Blasiak et al., 2020a). Importantly, once a molecule has been identified for a specific use, manufacturing of chemically synthesised versions or analogues to the molecule often follows, to avoid impacts such as over-harvesting/exploitation of wild populations or other associated environmental damage (Blasiak et al., 2020a; see recent example of use of shark squalene in vaccines; Sutch-Dagget, 2020). Researchers often use various biotechnological approaches to produce useful quantities of material. The discovery of MGRs can inspire innovation in the production of new types of molecules or materials, some of which are useful to humankind. More to the point, derivatives, rather than sequence data, are generally what is utilised in development of natural products, therefore are key in the value chain of MGR and their downstream commercialisation. In contrast, the majority of marine scientific research, such as biodiversity research utilising DSI, is generally focussed on the sequence data itself (rather than derivatives), e.g. COI barcodes, metabarcoding etc. Therefore, it is important to include derivatives in any definition of MGR.

The requirement for "functional units of heredity" in some of the proposed definitions of MGR is problematic when compared to the scientific definition. Compounds produced by marine organisms include secondary metabolites and enzymes that marine organisms may contain or release on to their surface or into the environment. For example, the biological cement or glue some organisms, such as barnacles, use to attach themselves to hard substrata (e.g. Liang et al., 2019) illustrates a potentially useful biochemical released by marine organisms that may not in itself contain units of heredity. In this example, the range of approaches used to study the composition of such biological glues includes DNA sequencing, but also studies of the chemical properties, structure, and amino acid sequence of various proteins secreted in the glue matrix (proteomics; Liang et al., 2019). Another problem with the term is that sections of a genome may be labelled as 'non-functional'- i.e. not coding for proteins- what is often referred to as 'non-coding' DNA. However, excluding such molecules would not be scientifically valid as large parts of the genome comprises of "non-coding DNA" which has been found to be functional as well as RNA molecules which also have a variety of functions (Pennisi, 2012; Li, 2019).

S4.0 National and International Cruise Notification

Europe

The Ocean Facilities Exchange Group (OFEG; http://www.ofeg.org/np4/home.html) provides the schedules for ocean-going research vessels for France, Germany, Netherlands, Norway, Spain and the United Kingdom. At the time of writing, because the COVID-19 pandemic has affected European cruise scheduling, some of the links to national vessel programmes were broken (e.g. Spain), or lacked information (e.g. France), or would provide information only by request (e.g. Netherlands). In general, national providers make upcoming cruise schedules publicly available. For the United Kingdom (UK), the OFEG site linked through to the Natural Environment Research Council (NERC) Marine Facilities Planning website (https://www.marinefacilitiesplanning.com/). This gives current and future cruise programmes for the UK vessels (*RRS Discovery*, *RRS James Cook*, *RRS Sir David Attenborough* and *RRS James Clark Ross*), including information on the cruise number, the Principal Scientist, and a map of the planned route. This information is also currently provided on the same website for the Spanish Research National Council (CSIC) fleet (*Sarmiento de Gamboa*,

Hesperides, *Garcia del Cid*) and the Netherlands Institute of Ocean Research (NIOZ) (*Pelagia*, *Navicula*).

In Germany, the portal for cruise planning (https://www.portal-forschungsschiffe.de/en/cruise-planning) provides cruise schedules for up to two years ahead for eight vessels (R/V *Polarstern* [Fig. 1A], R/V *Meteor*, R/V *Sonne*, R/V *Maria S Merian*, *Poseidon*, R/V *Alkor*, R/V *Heincke*, R/V *Elizabeth Mann Borgese*). The portal also gives a useful overview of cruise proposals accepted for funding, but not yet programmed into cruise scheduling.

Norway provides a cruise calendar for four vessels (R/V *GO Sars*, R/V *Johan Hjort*, R/V *Kristine Bonnevie* and R/V *Kronprins Haakon)*, through the cruise website of the Institute of Marine Research (IMR, https://toktsystem.imr.no/calendars). The vessel calendar projects forward one year, reflecting an annual vessel-time application deadline in August, for cruises starting from August the following year.

North America

In the USA, the University National Oceanographic Laboratory System (UNOLS) provides annual schedules for most of the US research fleet, along with area of operations (https://strs.unols.org/public/Search/diu_all_schedules.aspx) and shiptime request forms (https://www.unols.org/document/example-ship-time-request). Oversight and coordination of ship and National Deep-Submergence Facility vehicle schedules is conducted in conjunction with NSF and other funding agencies, and through the UNOLS Deep-Submergence Science Committee (https://www.unols.org/committee/deep-submergence-science-committee-dessc). UNOLS operation of global ocean class ships includes Woods Hole Oceanographic Institution (WHOI), who also release future schedules for its vessels (R/V *Atlantis* [Fig. 1C], R/V *Neil Armstrong*, R/V *Tioga*) and deep-submergence technology (submersible *Alvin*, remotely-operated vehicle [ROV] *Jason,* and autonomous underwater vehicle [AUV] *Sentry* [Fig. 1H]) for 2020. Searchable data archives, and past schedules for some vessels as far back as 2002 (https://www.whoi.edu/what-we-do/explore/vessels/vessels-schedules/). NOAA maintains an online platform that shows information on location, both past and present, of all 16 of ships in the NOAA fleet (https://shiptracker.noaa.gov/Account/Login?ReturnUrl=%2FHome%2FMapLoading). Additionally, NOAA publishes the schedules of its research vessel cruises online (https://sdat.noaa.gov/Account/Login?ReturnUrl=%2FHome%2FSchedule), however, this system is only available to NOAA users. The NOAA Office of Ocean Exploration and Research (OER) maintains the OER Digital Atlas, a map-based portal that provides access to all data from previous OER-supported expeditions going back to 1999. While these expeditions mostly focus on US waters, they do include limited surveys in adjacent countries, as well as in ABNJ. Additionally, NOAA OER announces its plans for upcoming expeditions on NOAA Ship Okeanos Explorer online (https://oceanexplorer.noaa.gov/explorations/explorations.html). GO-VESSEL, a US- led hydrography programme, also provides forward planning information on participating US cruises (see below).

South America

As an example of opportunities for ABNJ studies in developing countries, in Chile there are currently only two research vessels with ocean-going capacity, R/V *Abate Molina* and AGS *Cabo de Hornos* (Fig. 1B), which are managed by the National Fisheries Development Institute (IFOP) and the Chilean Navy, respectively. There are no formal publications of schedules for the use of these platforms, but they are available for the local and international scientific community through direct contact with their managers. Availability for the local

scientific community of AGS *Cabo de Hornos* is through specific calls by the Chilean Commission for Scientific and Technological Research (ANID) and also by the CIMAR program of the National Oceanographic Committee (CONA), but is almost completely focused on jurisdictional areas. Other local funding instruments are usually insufficient to cover the long-range expeditions needed to explore ABNJ. Therefore, in spite of Chile´s large maritime territory (~4 million km^2) and the vast ABNJ abutting its long coast and oceanic territories (~22 million km^2) in the southeast Pacific, except for occasional oceanographic research performed underway, mainly during transit to oceanic islands, there are very limited examples of local exploration of ABNJ.

Asia

In Japan most marine research vessels are operated by the Japan Agency for Marine-Earth Science and Technology (JAMSTEC). Their fleet, facilities and equipment as well as chartering rates are published (https://www.jamstec.go.jp/e/about/equipment) and availability of facilities and equipment can be directly requested.

In India information on seagoing research vessels, equipment, completed and ongoing projects can be found under https://www.nio.org/ .

Australia

The Australian Institute of Marine Science (http://www.aims.gov.au) has a live cruise tracker for the RV *Solander* and RV *Cape Ferguson*. The Marine National Facility (https://mnf.csiro.au) provides information on RV *Investigator*, its past, present and scheduled cruises, including information how to apply for primary or co-user shiptime.

International Programmes

The International Oceanographic Data and Information Exchange (IODE), under the Intergovernmental Oceanographic Commission of UNESCO (IOC), provides cruise data (https://www.iode.org/index.php?option=com_content&view=article&id=509&Itemid=100309) via the Oceanic site run by the University of Delaware (https://www.researchvessels.org/vessel_gen.asp). This site includes links to cruise data from Argentina, Australia, Belgium, Bermuda, Brazil, Chile, Denmark, Finland, France, Spain, Germany, India, Israel, Japan, South Korea, the Netherlands, Norway, Sweden, Turkey, the United Kingdom and the USA. These records run from 1990 to 2016 but lack of funding has resulted in discontinuation of further updating of the site (Douglas White pers. comms). However, the infrastructure still exists for development of an international system of cruise notification for ABNJ, giving information on past cruises as well as current and future cruise opportunities. Links from the IODE site to the Partner for Observation of the Global Oceans (POGO) website for records of cruises are currently broken. However, the POGO site itself provides access to reports from past cruises (http://www.pogo-oceancruises.org/international-research-vessel-cruise-programmes).

The International Research Ship Operators (IRSO; https://irso.info/) forum, a group of 49 research vessel operating organisations from 30 countries, represents more than 100 vessels that meet annually to exchange information and resolve common problems. The meetings typically address the following:

- Best practice, design, and operation of research vessels and associated scientific equipment
- Exchange of vessel time and equipment between countries
- Benchmarking and co-operation in support of marine research

- Developments in national research fleets

IRSO also developed a Code of Conduct for Marine Scientific Research Vessels and contributed to the founding of the Oceanic website noted above. They also provide links to regional research vessel organisations including the European Research Vessel Operators (ERVO) which is now Eurocean (https://www.eurocean.org/np4/home), OFEG, and UNOLS. Eurocean includes information on the current European research fleet (http://www.rvinfobase.eurocean.org/), but provides no scheduling information.

Eurofleets (https://www.eurofleets.eu/) provides funding for access to research vessels (14), AUVs (2), ROVs (7) and telepresence (1) facilities either as principal investigators, co-investigators or for remote access to samples or data. The programme especially encourages applications from nations with limited, or no, access to research vessels and other marine infrastructure. The programme includes infrastructure from a number of European countries as well as Bermuda, Canada, New Zealand and the USA. Information is provided for each vessel or other infrastructure, including geographical location of operation in future years.

The Global Ocean Vessel-based Hydrographic Investigations Program (GO-VESSEL) coordinates a global programme of sampling for physical oceanography, carbon cycle studies, biogeochemistry, and observations on ecosystems (https://www.go-vessel.org/index.html). The GO-VESSEL programme uses repeat observations of survey lines to provide approximately decadal resolution of changing inventories of ocean heat, freshwater, carbon, oxygen, nutrients, and transient tracers, spanning across ocean basins from coast to coast and full depth (top to bottom). Ten nations participate in the programme, including: Australia, Canada, France, Germany, Japan, Norway, South Africa, Spain, Sweden, United Kingdom and the USA. Some cruises occur annually or biennially and the GO-VESSEL website provides details of these and other cruises through to 2027, including the geographic area of operations (https://www.go-vessel.org/CruisePlans.html).

Privately funded research vessels

Privately funded research vessels are being offered as platforms for research to marine scientists globally free of charge. These include a number of ocean and global class research vessels run by philanthropic organisations including the RV *Falkor*, run by the Schmidt Ocean Institute, the EV *Nautilus* run but the Ocean Exploration Trust, and the forthcoming RV *REV Ocean* run by REV Ocean. The Schmidt Ocean Institute lists cruise schedules annually, as well as all past cruises and has ensured that all data arising from expeditions is made publicly available (https://www.schmidtocean.org). The Ocean Exploration Trust also announces its plans for upcoming expeditions on E/V Nautilus online and enables participation in cruises through live telepresence (https://nautiluslive.org/expedition). REV Ocean will similarly announce forthcoming cruise schedules and links to past cruise report via its website (https://www.revocean.org/). Again, this organisation will have an open data policy.

S5.0 The technical, financial and other challenges of the long-term storage and curation of biological material from ABNJ

All types of preservation of biological material require maintenance, for example, topping up specimen jars with preservatives to replace volatile preservation fluids that evaporate over time. Freezers and other cryopreservation facilities also require constant maintenance and connection to uninterruptible power supplies as well as batteries or emergency generators to ensure samples stay at low temperature during power cuts. Alarm systems are also often included, which are typically linked to curator's phones or email to enable transfer of

samples to back-up systems should freezers fail. Microbial culture collections are particularly challenging in terms of maintenance. Different microorganisms require special preservation procedures in order to ensure optimal viability, storage and quality (World Federation for Culture Collections, 2010). It is also notable that marine microorganisms are difficult to culture so considerable effort employing novel approaches may be necessary to keep them in viable cultures (Joint et al., 2010). Institutions holding microbial culture collections must ensure they have sufficient resources to handle new accessions, cope with supply of cultures and other services to researchers, as well as preservation, routine maintenance and viability testing (World Federation of Culture Collections, 2010). Cultures should be preserved by at least two different methods, at least one of which should be freeze drying or preservation at ultralow temperatures (-140°C or lower; World Federation of Culture Collections, 2010). If only freezing is available for cultures they should be stored as duplicates in separate freezers and or the most important or irreplaceable strains should be stored as duplicates at a separate site (World Federation of Culture Collections, 2010). The World Federation for Culture Collections has a long established World Data Centre for Microorganisms which provides data resources for microbial culture collections including Culture Collections Worldwide which provides metadata for more than 700 culture collections in more than 70 countries and the Global Catalogue of Microorganisms which provides access to information on more than 380,000 strains held by culture collections globally (Wu et al., 2017).

The Natural History Museum, London (NHM) illustrates an institution where researchers deposit marine biological samples. This institution houses some 80 million biological and mineralogical specimens (NHM, 2019) with the Department of Zoology housing about 29 million specimens including fish and invertebrates of marine origin (Rainbow, 2009). The Natural History Museum also houses about 350,000 seaweed samples mostly as herbarium specimens but also including 20,000 slides (Rainbow, 2009). The museum also houses the Molecular Collections facility, with the capacity to store 2 million 0.5ml cryovials to preserve molecular samples at low temperature. This facility includes a range of -20°C and -80°C freezers and three -196°C liquid nitrogen tanks, as well as humidity control cabinets (https://www.nhm.ac.uk/our-science/collections/molecular-collections.html). The facility supports externally funded projects, including the ABYSSLINE collections of deep-sea samples from ABNJ in the Pacific (see Rabone et al., 2019, Glover et al., 2016). The curation of the NHM collections currently run at more than £25 million per annum; although support costs, currently running at more than £17 million (NHM, 2019) are also likely to contribute to these (e.g. building maintenance etc.). The total income of the Natural History Museum is currently more than £86 million (NHM, 2019). Supporting access to biological specimens from the ocean therefore represents a significant cost requiring consideration in developing arrangements for sharing benefits from samples of ABNJ.

The care and curation of biological collections in museums and other institutions has long been undervalued and increasingly so in recent times in science (Thessen et al., 2019). Scientists and their institutions are judged through the production of scientific papers and by obtaining funding through grants, whilst curatorial activities are often viewed as non-productive (Thessen et al., 2019). Part of the under-valuing of curatorial work stems from the fact that taxonomy and curation go hand-in-hand- and taxonomy has suffered something of an existential threat in recent decades (e.g. Dupérré, 2020). The storage and curation of samples that may represent MGR providing opportunities for monetary and non-monetary benefits, however, are largely subsidised by institutions hosting collections (Rabone et al., 2019). Archiving of samples is a long-term and costly endeavour (many collections include types which need to be maintained in perpetuity) and with the erosion of public funding to major institutions such collections are under threat and increasingly looking to support

through grants. However, the grant-funding model of science is fundamentally at odds with curation which requires sustainable long-term funding streams (World Federation for Culture Collections, 2010; Collins et al., in review). Clearly, proper archiving of material is a costly and laborious endeavour currently practiced most extensively by developed countries. There is also a danger that the practice of archiving ABNJ specimens and samples in wealthy nations (even with generous loan policies) can be perceived as a form of colonialism. Solutions that engage or build capacity in less developed nations (e.g., digital archives, shared facilities in LDCs etc.), need to be discussed and developed (Collins et al., In review).

S6.0 Facilities required for marine biodiscovery

Marine biodiscovery can be conducted at various levels of sophistication, but a lab fully equipped to process MGR into biologically active products or biotechnological enzymes would require the following:

Facilities for storage and extraction

Materials collected from marine environments may be animals (vertebrate/invertebrate), algae, plankton, sediments etc. Whole organisms are sometimes used or alternatively, researchers can culture and isolate microorganisms from large variety of habitats, including other organisms. The latter commonly occurs now that researchers understand that much of the microbiome of an organism produces much of its chemistry. Organisms are stored in fridges/freezers or in solvents, whereas microorganisms can be preserved in a variety of ways (freeze dried/glycerol slopes). Highly prescribed preservation conditions are necessary to maintain the integrity of the organisms, requiring reliable cold storage. Whole organisms or microbial fermentations are extracted using organic solvents, requiring access to bulk, pure solvent supply (Fig. S5).

Microbiology and molecular biology facilities

Isolation of microorganisms from marine organisms or sediments requires basic microbiology facilities such as a clean bench and access to an autoclave (Fig. 8). Scale-up fermentation can be carried out using room temperature or temperature-controlled shakers. Molecular biology requirements depend on application - for identification of organisms, in particular microorganisms, amplifying sections of DNA requires access to primers, enzymes, and thermal cyclers. Most laboratories send DNA away to be sequenced at large, low cost, facilities to obtain high-quality data, therefore resulting in little need for in-house sequencing capacity. Whole genome sequencing is now fast and relatively cheap for microorganisms (although potentially out of reach for resource-limited states), but data analysis requires skill in bioinformatics. Many training opportunities now exist to learn to use open source bioinformatics software (e.g. see free bioinformatic software at: http://marnixmedema.nl/software.html). Researchers working in marine biodiscovery often prioritise organisms for further work based on the ‘chemical talent’ inherent in their biosynthetic genes. The discovery and manipulation of targeted enzymes requires more sophistication, and facilities will depend on the application.

Biological activity testing

Many types of benchtop assays give reliable data for prioritising extraction fractions, but even a simple microbiology lab enables the determination of antibacterial potential of an extract. Working with parasites to test extracts and pure compounds is more challenging, but model organisms have been genetically modified to act as a prescreen before testing against the parasite itself. Working with human cells can require additional facilities (e.g. CO_2 incubators) but may undertake prescreening and utilise many cell lines before

broadening to a wider range of cells. Biochemical/mechanistic assays require access to enzymes/reagents, and a mechanism such as a UV spectrophotometer or plate reader to analyse the output. Liquid handling robots and other types of automation, as well as image analysis software, can add sophistication but these tools are expensive and require specialist skills and training, and are out of reach to all but the most well-resourced labs.

Compound isolation and identification

Bioactive compound isolation is most often achieved by prioritising fractions of purification steps using bioactivity data (Fig. S5). Purification equipment can be as simple as a glass column containing chromatographic media, but good separation requires a medium or high-pressure liquid chromatograph. Alternatively, the use of liquid chromatography-mass spectrometry together with chemoinformatics software (e.g. the Global Natural Products Social Molecular Networking web-based ecosystem for sharing data on tandem mass spectrometry data – see: https://gnps.ucsd.edu/) also enables visualisation of the 'chemical talent' of an organism. Following isolation of a pure compound, x-rays can determine its structure if crystalline, or structure can be determined using spectroscopic equipment such as nuclear magnetic resonance spectrometers and mass spectrometers. Equipping a purification lab can cost tens of thousands of US dollars, and a spectroscopic lab can easily run into hundreds of thousands of dollars. Running, maintenance, and breakdown costs add significantly to this cost (Figure S5).

Figure S5. Laboratory research on MGR. (a) Early career scientist evaporating solvent from fraction of sample extraction (M Jaspars). (b) Early career scientist with nuclear magnetic resonance machine (M Jaspars). (c) Cultures of microorganisms (M Jaspars).

S7.0 Voyages of discovery, the *Tara* Ocean and *Malaspina* Expeditions

Samples were size fractionated on board the sailing vessel *Tara*, preserved specifically for genomics studies, and returned to land for sequencing of genomes (including from single cells) and genes expressed within planktonic communities (Karsenti et al., 2011; Pesant et al., 2015; Alberti et al., 2017; Seeleuthner et al., 2018). The impressive scientific knowledge gathered by this programme included the fact that more than 80% of the genes discovered from the sampled organisms were previously unknown to science (Sunagawa et al., 2015). The genomic information also revealed previously undescribed groups of microbial eukaryotes as well as large numbers of undescribed species from known groups of organisms (de Vargas et al., 2015). This expedition, along with the Malaspina expedition (Duarte, 2015) that sampled the oceans down to 5000 m depth, tripled the number of known viral populations from the ocean and doubled the number of bacterial and archaeal virus genera (Roux et al., 2016). This number was later expanded 12-fold (195,728 viral populations) through more in-depth sequencing of mesopelagic samples (zone between 200m and 1,000m depth) and inclusion of samples from the *Tara* Oceans Polar Circle expedition (Gregory et al., 2019). The analyses of bathypelagic microbial metagenomes from the Malaspina Expedition formed the basis for the recently published Malaspina Deep-Sea

Gene Collection. The collection consists of mostly novel genes (71%), and its analysis has uncovered 11 potential new phyla greatly expanding our understanding of the functional diversity and metabolic versatility of deep ocean microbial taxa (Acinas et al., in press). *Tara* Oceans also identified 116 million genes expressed by planktonic organisms across the global ocean (Carradec et al., 2018; see also Salazar et al., 2019). The programme implemented an open data archiving strategy from its inception and nucleic acid sequence data were deposited in the International Nucleotide Sequence Database Collaboration (INSDC, http://www.insdc.org/about) via the European gateway to the INSDC, the European Nucleotide Archive (ENA, http://www.ebi.ac.uk/ena) at the European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI; Alberti et al., 2017). All environmental and biogeochemical measurements for each *Tara* Oceans sample were deposited at the PANGAEA database.

Samples were size fractionated on board the sailing vessel *Tara*, preserved specifically for genomics studies, and returned to land for sequencing of genomes (including from single cells) and genes expressed within planktonic communities (Karsenti et al., 2011; Pesant et al., 2015; Alberti et al., 2017; Seeleuthner et al., 2018). The impressive scientific knowledge gathered by this programme included the fact that more than 80% of the genes discovered from the sampled organisms were previously unknown to science (Sunagawa et al., 2015). The genomic information also revealed previously undescribed groups of microbial eukaryotes as well as large numbers of undescribed species from known groups of organisms (de Vargas et al., 2015). This expedition, along with the Malaspina expedition (Duarte, 2015) that sampled the oceans down to 5000 m depth, tripled the number of known viral populations from the ocean and doubled the number of bacterial and archaeal virus genera (Roux et al., 2016). This number was later expanded 12-fold (195,728 viral populations) through more in-depth sequencing of mesopelagic samples (zone between 200m and 1,000m depth) and inclusion of samples from the *Tara* Oceans Polar Circle expedition (Gregory et al., 2019). The analyses of bathypelagic microbial metagenomes from the Malaspina Expedition formed the basis for the recently published Malaspina Deep-Sea Gene Collection. The collection consists of mostly novel genes (71%), and its analysis has uncovered 11 potential new phyla greatly expanding our understanding of the functional diversity and metabolic versatility of deep ocean microbial taxa (Acinas et al., in press). *Tara* Oceans also identified 116 million genes expressed by planktonic organisms across the global ocean (Carradec et al., 2018; see also Salazar et al., 2019). The programme implemented an open data archiving strategy from its inception and nucleic acid sequence data were deposited in the International Nucleotide Sequence Database Collaboration (INSDC, http://www.insdc.org/about) via the European gateway to the INSDC, the European Nucleotide Archive (ENA, http://www.ebi.ac.uk/ena) at the European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI; Alberti et al., 2017). All environmental and biogeochemical measurements for each *Tara* Oceans sample were deposited at the PANGAEA database.

References

Abdelkrim, J., Aznar-Cormano, L., Buge, B., Fedosov, A., Kantor, Y., Zaharias, P., et al., (2018). Delimiting species of marine gastropods (Turridae, Conoidea) using RAD sequencing in an integrative taxonomy framework. *Mol. Ecol.* 27, 4591-4611.

Acinas, S.G., Sánchez, P., Salazar, G., Cornejo-Castillo, F.M., Sebastián, M., Logares, R., et al. (2019). Metabolic architecture of the deep ocean microbiome. *BioRxiv*, 635680.

Balbar, A.C., Metaxas, A. (2019). The current application of ecological connectivity in the design of marine protected areas. *Global Ecol. Cons.* 17, e00569.

Bjørn-Yoshimoto, W.E., Ramiro, I.B.L., Yandell, M., McIntosh, J.M., Olivera, B.M., Ellgaard L., et al., (2020). Curses or cures: a review of the numerous benefits versus the biosecurity concerns of conotoxin research. *Biomedicines* 8, 235.

Bruno, S., Coppola, D., di Prisco, G., Giordano, D., Verde, C. (2019). enzymes from marine polar regions and their biotechnological applications. *Mar. Drugs* 17, 544.

Cardás, J.B., Deconinck, D., Márquez, I., Torre, P.P., Garcia-Vazquez, E., Machado-Schiaffino, G. (2020). New eDNA based tool applied to the specific detection and monitoring of the endangered European eel. *Biol. Conserv.* 250, 108750.

Carr, M.H., Robinson, S.P., Wahle, C., Davis, G., Kroll, S., Murray, S. (2017). The central importance of ecological spatial connectivity to effective coastal marine protected areas and to meeting the challenges of climate change in the marine environment. *Aquat. Conserv.* 27 (Suppl. 1), 6-29.

Carradec, Q., Pelletier, E., Da Silva, C., Alberti, A., Seeleuthner, Y., Blanc-Mathieu, R., et al., (2018). A global ocean atlas of eukaryotic genes. *Nat. Commun.* 9, 373.

Carvalho, G.R., Hauser, L. (1994). Molecular genetics and the stock concept in fisheries. *Rev. Fish Biol. Fisher*. 4, 326-350.

Cincin-Sain, B., Vierros, M., Balgos, M., Maxwell, A., Kurz, M., Farmer, T., et al. (2018) Policy Brief on Capacity Development as a Key Aspect of a New International Agreement on Marine Biodiversity Beyond National Jurisdiction (BBNJ). Published by the Global Ocean Forum, Washington DC, USA, 85pp. Available at: https://globaloceanforumdotcom.files.wordpress.com/2018/08/policy-brief-on-bbnj-capacity-development-aug-2018.pdf

Dalmaso, G.Z.L., Ferreira, D., Vermelho, A.B. (2015). Marine extremophiles: a source of hydrolases for biotechnological applications. *Mar. Drugs* 13, 1925-1965.

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., et al., (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science*, 348, 1261605.

Dupérré, N. (2020). Old and new challenges in taxonomy: what are taxonomists up against? *Megataxa* 1, 59–62.

Frankham, R., Ballou, J.D., Briscoe, D.A. (2010) *Introduction to Conservation Genetics, 2nd Edition*. Cambridge University Press, Cambridge, UK, 618pp.

Gasol, J.M., Kirchman, D.L. (2018). The evolution of microbial ecology of the ocean. In: Gasol, J.M., Kirchman, D.L. (Eds.) *Microbial Ecology of the Ocean 3rd Edition*. John Wiley & Sons, Hoboken, New Jersey, USA, p 1-46.

Gregory AC, Zayed AA, Conceiçãco-Neto N, Temperton B, Bolduc B, Alberti A, et al. (2019). Marine DNA viral macro- and microdiversity from pole to pole. *Cell* 177, 1109-1123.

Hikida, Y., Kimoto, M., Hirao, I., Yokoyama, S. (2017). Crystal structure of Deep Vent DNA polymerase. *Biochem. Bioph. Res. Co.* 483, 52-57.

Hunter, M.E., Meigs-Friend, G., Ferrante, J.A., Kamla, A.T., Dorazio, R.M., Diagne, L.K., et al., (2018) Surveys of environmental DNA (eDNA): a new approach to estimate occurrence in Vulnerable manatee populations. *Endanger. Species Res*. 35, 101-111.

Iglésias, S.P., Toulhoat, L., Sellos, D.Y. (2010). Taxonomic confusion and market mislabelling of threatened skates: important consequences for their conservation status. *Aquat. Conserv.* 20, 319-333.

Joint, I., Mühling, M., Querellou, J. (2010). Culturing marine bacteria – an essential prerequisite for biodiscovery. *Microb. Biotechnol*. 3, 564-575.

Li, J., Liu, C. (2019). Coding or noncoding, the converging concepts of RNAs. *Front.Genetics* 10, 496.

Liu, Q., Zhang, Y., Wu, H., Liu, F., Peng, W., Zhang, X., et al., (2020). A review and perspective of eDNA application to eutrophication and HAB control in freshwater and marine ecosystems. *Microorganisms* 8, 417.

Midwestern University (2020). Clinical Pipeline Marine Pharmacology. https://www.midwestern.edu/departments/marinepharmacology/clinical-pipeline.xml Accessed 9/10/2020.

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

Morris, R.M. (2018). Metatranscriptomics and metaproteomics: elucidating marine microbial ecosystem functions. In: Gasol, J.M., Kirchman, D.L. (Eds.) *Microbial Ecology of the Ocean 3rd Edition*. John Wiley & Sons, Hoboken, New Jersey, USA, p 123-148.

Moura, C.J., Lessios, H., Cortés. J., Nizinski, M.S., Reed, J., Santos. R.S., et al., (2018). Hundreds of genetic barcodes of the species-rich hydroid superfamily Plumularioidea (Cnidaria, Medusozoa) provide a guide toward more reliable taxonomy. *Sci. Rep.-UK* 8, 17986.

Natural History Museum (2019) *Speaking to the World: Annual Review 2018/19*. Natural History Museum, London, UK, 56pp.

Newman, D.J., Cragg, G.M. (2016). Natural products as sources of new drugs from 1981 to 2014. *J. Nat. Products* 79, 629-661.

Pedrós-Alió, C., Acinas, S.G., Logares, R., Massana, R. (2018). Marine microbial diversity as seen through high-throughput sequencing. In: Gasol, J.M., Kirchman, D.L. (Eds.) *Microbial Ecology of the Ocean 3rd Edition*. John Wiley & Sons, Hoboken, New Jersey, USA, p 47-98.

Penisi E (2012) ENCODE Project writes eulogy for junk DNA. Science 337: 1159-1161.

Pleijel, F., Jondelius, U., Norlinder, E., Nygren, A., Oxelman, B., Schander, C., et al., (2008). Phylogenies without roots? A plea for the use of vouchers in molecular phylogenetic studies. *Mol. Phylogenet. Evol*. 48, 369-371.

Rogers, A.D. (2019). *The Deep: The Hidden Wonders of Our Ocean and How We Can Protect Them.* Wildfire Publishers, London, UK. 358pp.

Roux, S., Brum, J.R., Dutilh, B.E., Sunagawa, S., Duhaime, M.B., Loy, A., et al., (2016). Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature* 537, 689-697.

Ruppert, K.M., Kline, R.J., Rahman, M.S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecol. Conserv*. 17, e00547.

Salazar, G., Paoli, L., Alberti, A., Huerta-Cepas, J., Ruscheweyh, H.-J., Cuenca, M., et al., (2019). Gene expression changes and community turnover differentially shape the global ocean metatranscriptome. *Cell* 179, 1068-1083.

Sarmiento, F., Peralta, R., Balmey, J.M. (2015). Cold and hot extremozymes: Industrial relevance and current trends. *Front. Bioeng. Biotechnol.* 3, 148.

Schneider, C.W., Popolizio, C.R., Saunders, G.W. (2019). Collections from the mesophytic zone off Bermuda reveal three species of Kallymeniaceae (Gigartinales, Rhodophyta) in genera with transoceanic distributions. *J. Phycol.* 55, 415-424.

Seeleuthner, Y., Mondy, S., Lombard, V., Carradec, Q., Pelletier, E., Wessner, M., et al., (2018). Single-cell genomics of multiple uncultured stramenopiles reveals underestimated functional diversity across oceans *Nat. Commun.* 9, 310.

Sunagawa, S., Coelho, L.P., Chaffron, S., Kultima, J.R., Labadie, K., Salazar, G., et al., (2015). Structure and function of the global ocean microbiome. *Science* 348, 1261359.

Takagi, M., Nishioka, M., Kakihara, H., Kitabayashi, M., Inoue, H., Kawakami, B., et al., (1997) Characterization of DNA polymerase from *Pyrococcus* sp. strain KOD1 and Its application to PCR. *App. Environ. Microb*. 63, 4504-4510.

Terpe, K. (2013). Overview of thermostable DNA polymerases for classical PCR applications: from molecular and biochemical fundamentals to commercial systems. *App. Microbiol. Biot.* 97: 10243-10254.

Thessen, A.E., Woodburn, M., Koureas, D., Paul, D., Conlon, M., Shorthouse, D.P., et al., (2019). Proper attribution for curation and maintenance of research collections: metadata recommendations of the RDA/TDWG Working Group. *Data Sci. J.* 18, 1–11.

World Federation for Culture Collections (2010) *World Federation For Culture Collections Guidelines for the Establishment of Collections of Cultures or Microorganisms. 3rd Edition*. World Federation for Culture Collections, Brussels, Belgium, 19pp.

Wu, L., Sun, Q., Desmeth, P., Sugawara, H., Xu, Z., McCluskey, K., et al., (2017). World data centre for microorganisms: an information infrastructure to explore and utilize preserved microbial strains worldwide. *Nucleic Acids Res.* 45, D611-618.

Zea, S., Henkel, T.P., Pawlik, J.R. (2014). *The Sponge Guide: a picture guide to Caribbean sponges*. 3rd Edition. Available online at www.spongeguide.org. Accessed on: 2020-09-03.