Intergovernmental Oceanographic Commission
Workshop Report No. 293

First PacMAN Scientific workshop

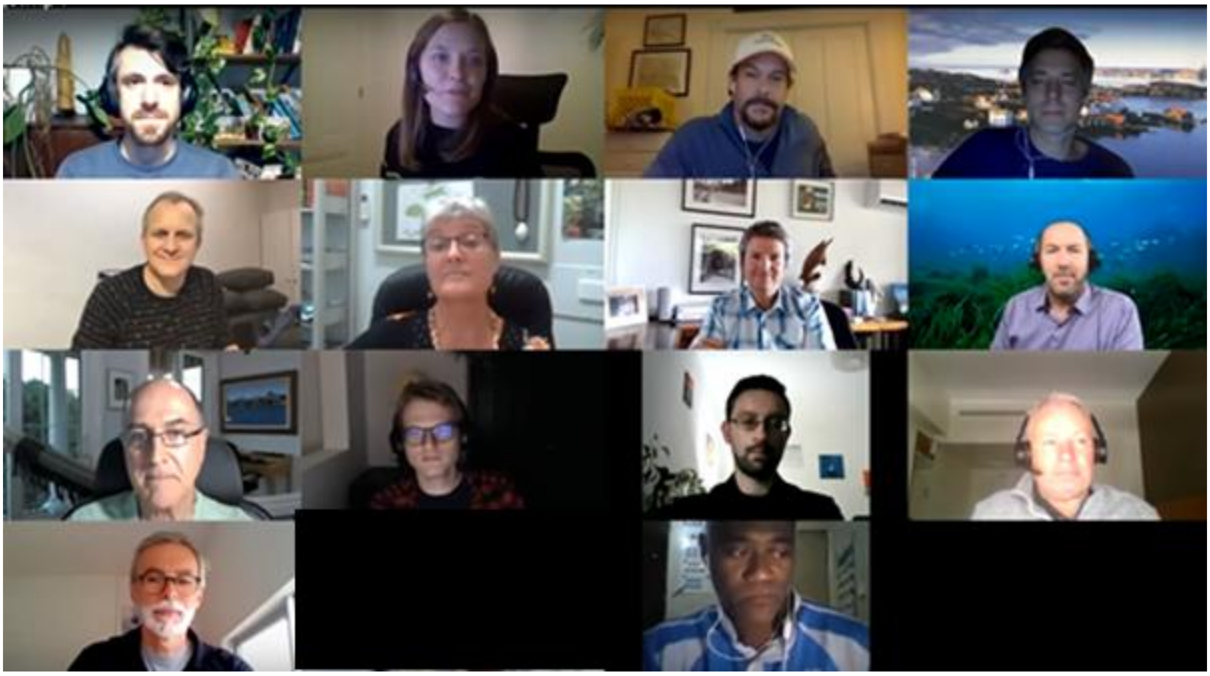
Online
23-24 November & 30 November-1 December 2020

UNESCO

IODE/PacMAN

Online, 23-24 November & 30 November-1 December 2020

English only



Group picture (Names from left to right, from top to bottom): Dr. Pascal Hablutzel, Dr. Saara Suominen, Dr. Chris Meyer, Assoc. Prof. Matthias Obst, Mr. Ward Appeltans, Dr. Gilianne Brodie, Dr. Neil Davies, Assoc. Prof. Craig Sherman, Dr. Frank Muller-Karger, Mr. Pieter Provoost, Dr. Pier Luigi Buttigieg, Mr. Kevin MacKay, Dr. Nic Bax, Mr. Joape Ginigini. In addition participants not in this picture: Mrs. Paayal Kumar and Mr. John Deck

For bibliographic purposes, this document should be cited as follows:
IODE/First PacMAN Scientific workshop, 23-24 November & 30 November-1 December 2020.
IOC Workshop Reports, 293, UNESCO, 21 pp. (English)

Published in 2021
by the United Nations Educational,
Scientific and Cultural Organization
7, place de Fontenoy, 75352 Paris 07 SP
(IOC/2021/WR/293)

Table of Contents

OPENING THE SESSION 5

SESSION 1: INTRODUCTIONS 5

- Introduction to PacMAN 5
- Introduction to the team in Fiji 5
- Introduction to the scientific committee 6

SESSION 2: Monitoring plan 9

- Survey on non-indigenous species (NIS) 9
- Sampling plan 10
  - Initial port survey 10
  - Sample types 11
- eDNA analysis 11

SESSION 3: Data management 12

- Sequence data management 12
- Taxonomic identification 12
- Access and Benefit sharing 13

SESSION 4: Synthesis and workplan 13

Annexes 14

- Annex 1 Meeting Agenda 14
- Annex 2 Participants list the First PacMAN Scientific Workshop, 2020 **Error! Bookmark not defined.**

Executive Summary

The PacMAN (Pacific Islands Marine Bioinvasions Alert Network) project is a 3-year project started in August 2020, which aims to build a monitoring plan for marine invasive species for small Island developing states utilizing eDNA analyses. The first PacMAN scientific workshop brought together the scientific experts of the advisory board of the project, with expertise in invasive species, marine ecology, molecular ecology and scientific data management, to discuss their experiences as well as the possible approaches for the monitoring plan and data management of the project. In total during three online sessions of 2 hours there were 18 participants, including the project team as well as two observers from the Biosecurity Authority of Fiji.

During the first session, introductions to the project, the team, the stakeholders and the facilities at the University of South Pacific started the discussion. Each scientific expert shortly presented their experience and field of expertise. The presentations included designs of invasive species monitoring in Australia and Belgium as well as scientific research around the use of molecular tools for the purpose of monitoring. The occurrence records available from Fiji currently were introduced by the South West Pacific OBIS node. The global ARMS (Autonomous Reef Monitoring Structures) program as well as the European network of ARMS deployments for biodiversity monitoring and the standards related to this were introduced. In addition, the GEOME (the Genomic Observatories Metadatabase) platform for scientific data management and MBON (Marine Biodiversity Observation Network) initiatives were introduced, the existing initiatives for FAIR compliant data management in Moorea were discussed and the session was closed with a short discussion on the recent developments around standardizing data management protocols for biodiversity and omics data.

The second and the third sessions of the workshop were reserved for free discussion around the monitoring plan as well data management plan for the project. Due to a lack of previous studies and incomplete reference databases in Fiji, the importance of baseline biodiversity surveys was highlighted. In addition to developing a metabarcoding approach for surveying the biodiversity in sampling sites, it was agreed that there is a need to collect a list of target risk species that can directly guide management efforts at site. The necessary sampling design, and choices for the molecular assays and sequencing platform were discussed. A full data management plan will need to be developed before the start of sampling, including decisions on how sequence data will be processed, who owns the samples and data, as well as how they can be accessed after sampling.

Finally, the meeting decided on an action plan that will initiate the project and the development of the monitoring plan. A comprehensive needs assessment of the major local stakeholders will be made to evaluate the best approach for the community in Fiji. A list of target species based on literature review of surveys done in other locations in the Pacific will be made and based on these identified lists the completeness of the genetic reference databases will be evaluated. A decision tree for the different use cases for the monitoring plan and the resulting decision support tool will be made to aid communication with the main local stakeholders.

The meeting participants agreed that after these actions and a draft monitoring and data management plan were drawn up, a follow-up synthesis meeting could be held for further discussion on the specific scientific details of the monitoring plan.

1 OPENING THE SESSION

Mr Ward Appeltans (OBIS/PacMAN project manager) opened the session with welcoming words to all participants. Dr Saara Suominen (OBIS/PacMAN Scientific Officer) introduced the objectives and agenda of the workshop and the practical matters involved with the virtual meeting. The meeting brought together the scientific experts who expressed interest in contributing to PacMAN and the project coordination team based at the OBIS Secretariat (Belgium) and the University of the South Pacific (Fiji). The goals of the meeting were to (i) introduce everyone to the PacMAN project, (ii) getting familiar with the projects, networks and systems of the scientific experts and (iii) provide advice on the different elements of the national monitoring plan including technical and scientific aspects, and more specifically "sampling, sample handling, molecular work, data management and policies and bioinformatics pipeline.

2 SESSION 1: INTRODUCTIONS

2.1 Introduction to PacMAN

Dr Saara Suominen provided an introduction to the PacMAN project. The objective of PacMAN (Pacific islands Marine bioinvasions Alert Network) is to provide a national monitoring plan for identifying and monitoring the presence of invasive aquatic species (IAS) in the SIDS (small islands developing states) in the South Pacific. The monitoring plan and trial periods will be developed in collaboration with Fiji, with a possibility to be applied to other island states as well. The scientific knowledge, tools and services will directly support the development and implementation of national strategic action plans for the control and management of marine invasive species.

The main approaches of PacMAN include a comprehensive needs assessment of the local community in terms of monitoring of marine invasive species. Monitoring will be based on environmental DNA (eDNA analyses) and rely on risk species identifications and initial port surveys. The sequence data derived from sampling will be processed and stored in a standardized manner to ensure the fulfilment of the FAIR principles. Standard bioinformatics pipelines and data management practices will be reviewed and adopted as seen fit. Based on the output data stored in the Ocean Biodiversity Information System (OBIS), a decision-support tool will provide the observations in a user-friendly dashboard indicating the presence of invasive species. Early warnings will be generated based on observations in nearby adjacent areas and will provide a service for local managers and decision-makers to set up targeted monitoring programmes. In addition, to ensure the longevity of the program, training for local scientists and environmental managers will be provided in collaboration with the OceanTeacher Global Academy (OTGA).

2.2 Introduction to the team in Fiji

The University of the South Pacific (USP) in Suva, Fiji is the local implementing partner for the PacMAN Project. Dr Gilianne Brodie (USP/PacMAN marine science advisor) discussed preliminary knowledge of the stakeholder needs in Fiji. USP is the main institution of higher learning for the Pacific region, owned by governments of the island states. It has 14 campuses, with the region spreading across 33 million square kilometers of ocean. The identified stakeholders include local experts in ports, shipping, biosecurity and the environment, as well as local beneficiaries, the science experts, regional organisations as well as international organisations. The main national stakeholders in Fiji include the Fiji Ministry of Waterways and Environment (MoWE), the Biosecurity Authority of Fiji (BAF), the Maritime Safety Authority (MSA) of Fiji and the University of the South Pacific (USP). At the moment, preliminary meetings have been held, but a more comprehensive needs assessment will be conducted to capture in detail the core needs of the community, as well as long term needs. The national stakeholders expect the program to benefit the local community directly, strategically align with existing initiatives, directly involve their knowledge and experience as well as support meeting existing commitments. The monitoring plan should not significantly increase their day to day workload and responsibilities. More specifically, there is need for supporting the national strategic action plan for the management of ship biofouling, and alignment to existing Invasive Alien Species projects of BAF. Most importantly it needs to be ensured that the benefits of PacMAN will be sustained beyond the life of the project. In terms of the USP, there are wishes for getting baseline surveys of the harbor areas, storing data for long-term use by stakeholders, showcasing new molecular technologies as well as supporting the development of in-country expertise and facilities.

Mr Joape Ginigini (PacMAN project manager at the University of the South Pacific in Suva, Fiji) introduced his team and the equipment available from the Pacific Natural Products Research Centre. The core functions of the research centre include increasing technical capacity of the local community to valorise local flora and fauna, and to capture benefits of the local ecological knowledge. Recently, the unit has also diversified to blue carbon research as well ocean acidification research. Currently data is collected to a bioprospecting database started in 2018. The key facilities in the molecular lab in USP include a preparatory room, where DNA extracts and buffers are prepared, and analytical balances, water baths, as well as consumables and reagents. There is a PCR room with a gradient PCR as well as isotemperature PCR. The laboratory has both a Qbit and nanodrop devices for DNA quantifications as well as all necessary equipment for gel electrophoresis and imaging. In addition in a third room there is access to a biosafety cabinet level 2 as well as a centrifuge that can take 48 samples, a homogenizer, and set-ups for 96 well plate analyses. In addition there is a 96-well plate UV plate reader, which is planned to be used in the future. Finally the USP also has ARMS plates which are available for use from previous research projects.

2.3 Introduction to the scientific committee

Each member of the scientific committee gave introductions on themselves as well as their experience with regards to the activities related to the objectives of the PacMAN project.

Dr Nic Bax (CSIRO, Australia) gave a brief overview of the Australian Centre for Research on Introduced Marine Species (CRIMP, 1995-2005) with an emphasis of the role of genetic probes in testing the efficacy of the domestic shipping Decision Support System in detecting species in the port where ballast water was taken up. CRIMP undertook a series of national port surveys to establish the invasive species baseline, developed risk assessments to identify the most likely high impact invasive species most likely to arrive next in Australia (and in the SW Pacific generally), worked with the Australian government and State of Victoria to provide Decision Support Tools to manage ballast water on incoming ships. These tools were then tested with genetic probes developed through RT PCR to identity invasive species at very low concentration. CRIMP also developed public information tools to support response to new pest incursions. Dr Bax stressed the importance of prioritising the invasive species that are being looked for in biosecurity situations, where there needs to be an option of a timely response if the data are to be used. Risk assessment is a key first step and guiding principle for designing projects aimed at reducing the threats or impacts of marine invasive species.

Assoc. Prof. Craig Sherman (EcoGenetics Lab, Deakin University, Australia) gave an overview of the use of molecular tools for marine biosecurity work in south eastern Australia. This includes traditional sanger sequencing for taxonomic verification, phylogeography and population genetic surveys (for identifying source populations, introduction pathways, range expansions, sources and sinks, genetic bottlenecks), species-specific quantitative and digital PCR for targeted species detections and metabarcoding approaches for monitoring and surveillance. With targeted probes for single species detection, these need to be validated in silico and in vivo to ensure they are species-specific and there is no cross reactivity against native taxa, or species already present in an area. It is also important to establish eDNA decay rates and how it is dispersed by local hydrodynamics, especially for marine environments where complex hydrodynamic movements can cause eDNA to be detected some distance from where it was released. Combining high resolution biophysical models with an understanding of eDNA decay ties can provide an detection shadow over which the eDNA may have originated from. Metabarcoding approaches provide an alternative surveillance tool that allows for the monitoring of a broader range of taxon and therefore may allow for the detection of a range of different introduced species. However, the major issue is being able to identify invasive sequences within the data sets generated. This arises because many invasive species have no reference sequences in the data bases, or the loci used do not provide species level taxonomic resolution (i.e. invasive and closely related taxa have the same sequence for the metabarcoding gene that has been sequenced). Australia is currently developing a reference sequence database for both potentially invasive and native taxa using a genome skimming approach. The idea is to recover full mitochondrial genomes (which are often targeted for metabarcoding) and partial nuclear genomes using taxonomically verified samples. This is so that the most appropriate loci can be targeted that provide the best taxonomic resolution. Ideally, all data for reference databases should be made publicly available with associated metadata on the source of specimens and the taxonomic verification process.

Mr Kevin Mackay (NIWA, New Zealand) introduced OBIS (Ocean Biodiversity Information System) and the Southwestern Pacific Regional OBIS Node (SWPRON). SWPRON is already established as a mechanism for publishing biological and environmental datasets for the southwestern Pacific region into the global OBIS and GBIF networks. This is achieved by providing the resources to convert existing datasets to the Darwin Core standard, or working directly with data providers to generate, and harvest from, streams of Darwin Core standard data from current sampling programmes. Current records in OBIS for Fijian waters indicate that only sporadic sampling events have been ingested into the public portals, most of which are based on fish and mollusc specimens from museum collections. SWPRON is currently working with Biosecurity New Zealand and the Ministry for Primary Industries to publish the results from the current port biosecurity surveys to the OBIS network.

Dr Pascal Hablützel (Flanders Marine Institute, Belgium) provided an overview of the GEANS project (https://northsearegion.eu/geans/). The GEANS consortium aims to harmonise and consolidate metabarcoding for routine sampling of biological diversity across North Sea countries. It uses pilot studies for validation of genetic tools and methods, and creates synergies and assures comparability through transnational co-operation. Furthermore, the consortium develops a policy decision framework including fit for purpose choice of genetic tools and protocols, helping to translate genetic results for interpretation by stakeholders. One of the pilot studies within GEANS focuses on monitoring of non-indigenous species using a metabarcoding approach.

Dr Chris Meyer (Smithsonian National Museum of Natural History) introduced the Global ARMS Program, a standardized sequence-based approach to census complex benthic marine communities. While initially developed through the Census of Marine Life's CREEFs program for use on coral reefs, over 2,000 ARMS units are now in use all over the globe from the poles to the tropics and from the intertidal to submersible depths. ARMS is short for Autonomous Reef Monitoring Structures. These devices act as identical, pre-fabricated “homes” to attract marine species and mimic the complexity of benthic habitats in order to capture a representative sample of local marine biodiversity. They are anchored to the sea floor for a period of time (generally 1-3 years), then recovered to see who “moved in”. The ARMS are taken apart, size-fractioned, the plates are photographed and then scraped. Each fraction is analyzed using various metabarcoding approaches with potentially multiple amplicon libraries. Each ARMS documents 100s to 1000s of species. ARMS are currently being used to look at the effect of marine protected areas, anthropogenic influences (pollution), ocean acidification, depth gradients, and as detectors for invasive species. Because they are standardized, they can compare place to place or a particular location over time to determine trends and shifts in marine communities. The Global ARMS portal serves as a centralized supplier, a clearinghouse for protocols, and to provide training opportunities for partners.

Assoc. Prof. Matthias Obst gave a status update for the European ARMS-MBON. A network of ARMS samplers were deployed around Europe in 2018-2020, to have timelines of biodiversity. More than 130 ARMS have been deployed, three fractions will be collected and three genetic markers will be sequenced. ARMS are used especially for hard bottom monitoring, which is notoriously difficult otherwise. Short-term deployments are used for alien species scans, while long-term deployments are used for building timelines. Due to standardized nature, there are many factors that need to be taken into account while planning sampling campaigns, which this project has formulated in detail. Images are a major part of the ARMS sampling, and are used to also develop species lists. Samples are also stored for future use and possible developments in sequencing technologies. So far, the results of this project has shown for example that invasive species, while found in ports, are not present in marine protected areas, showing how important these areas are. Data management for the European ARMS project is aiming to link images, raw sequences, and fastq files, as well as the species lists that come from different sources. The project suggests building predictive ability with biodiversity data such as “biological weather forecasts”, throughout these broad observation networks. The challenges related to these projects relate to the long-term goals; how can the project build motivation in the community to continue observations after the end of the project. An issue with these observations is a time-lag between sampling and obtaining results, which needs to be minimised. In addition, the need for a wide range of expertise can be a challenge for the project development.

Mr John Deck (University of California at Berkeley) gave an introduction to the Genomic Observatories Metadatabase (GEOME). This project started with the Moorea biocode project, which revealed the challenges in working across different research groups and universities. One of the main functions of GEOME is to provide instant feedback on sample data contained in spreadsheets, which would allow management of errors in input data. A main objective is to support researchers to comply with global metadata standards as well as use persistent identifiers, and work in project spaces. When users input data, they are guided to generate spreadsheet templates for metadata. In addition they can add many different types of data like image data as well as sequence data, which is linked to external data storage sites like NCBI or BOLD. There is discussion of linking eDNA modules to subsamples in the project space.

Dr Frank Muller-Karger (University of South Florida, USA) provided an overview of the Marine Biodiversity Observation Network (MBON). MBON links groups that collect observations in the world's ocean and provides advice on how to track marine species, their abundance and biomass, and how marine habitats are changing. MBON is working with the Global Ocean Observing System (GOOS), the Ocean Biodiversity Information System (OBIS), and the Ocean Best Practices System (OBPS) of the Intergovernmental Oceanographic Commission (IOC), to define the Essential Biodiversity Variables (EBVs) and Essential Ocean Variables (EOVs) that need to be measured locally and by each State for UN assessments, to meet the UN Sustainable Development Goals, and to address targets and goals defined under the Convention on Biological Diversity. This network seeks to serve the community and provide guidelines for the UN Decade of Ocean Science for Sustainable Development 2021-2030. (MBON website: http://marinebon.org)

Dr Neil Davies (Gump South Pacific Research Station, Moorea, French Polynesia; UC Berkeley) reflected on progress in genetic approaches to determining the pathways of marine bioinvasions since the first Marine Bioinvasions Conference in 1999. He reviewed global efforts to implement crucial data standards for observations of biodiversity at the molecular level, such as those of the Genomic Standards Consortium (GSC) and GEO Biodiversity Observations Network (BON) through the Genomic Observatories Network initiated in 2012. These efforts, including programs such as the Moorea Biocode Project and Ocean Sampling Day, have contributed to growing databases of well-contextualized genetic information, which combined with other data types and sources of evidence, feed into increasingly useful ecological models. He emphasized the importance of science-community dialogue in these processes to ensure that new technologies and scientific understanding respond to societal needs in the Pacific Islands and address social-ecological challenges at local as well as regional scales.

Dr Pier Luigi Buttigieg (BioECO panel, GOOS), introduced BioECO essential ocean variables and a recently started data working group. At the moment, standards are lacking in many areas of biological oceanography which are now being developed. There are efforts to start combining MixS and Darwin Core standards used by the community at large. This is needed especially to combine practices in biodiversity research with the community working with genetic data. As well for the data to be retrievable and accessible for the community after projects are finished, and to retain links to protocols used in the data processing workflows. Currently, systems to collect best practices for the community are being developed, as well as large initiatives to build a global community for biodiversity observation with environmental DNA.

3 SESSION 2: Monitoring plan

The best practices for the monitoring plan were discussed based on preliminary stakeholder assessments and shared expertise in sampling, sample handling and molecular work.

3.1 Survey on non-indigenous species (NIS)

Major challenges related to eDNA surveys include incomplete reference databases. Therefore, it is important to perform baseline surveys of biodiversity at the study site to be able to follow the introduction of new species into the area (CEMP indicators), as well as understand the completeness of the reference database, and if more work is needed to support this. In this section we will discuss the possible existing knowledge on marine invasive species in Fiji and surrounding areas, as well as recommendations on how to best categorize genetic data and build reference databases.

Dr Bax recommended a careful consideration of the focal point of the project, also corroborated by Dr Neil Davies later in the discussion. It was noted that significantly different approaches will be needed if the focus is to improve the biosecurity capacity of Fiji instead of focusing on adding scientific information on the status of biodiversity in the sampling sites. Most of the scientific information that will be gained from the proposed collections and metabarcoding analyses will not be relevant to biosecurity, or even when relevant, some of the information will not be able to be acted upon in a timely fashion or within the resources available. For example, if there is a six month delay between collection and results, then it would often be too late to undertake an early eradication, but might be sufficient to develop new management to reduce further spread. Additionally, there will be many detections of previously unknown species for Fijian waters, whose origin will not be clear.

Due to this point, the need to focus on developing a list of target risk species, and their comprehensive risk assessments was discussed. Several lists from Australia exist that could be used as a starting point for this work. The most noteworthy species could then be followed by quantitative PCR assays, which are faster and more sensitive for detecting low-abundance species. It was noted however, that this would require considerable work in the development of species-specific primers for each of the species in question, that would be reliable also in the context of the local biodiversity in Fijian waters. In addition, this approach would not enable the detection of new possible invasive species in the area, and would not be transferable to other high-risk regions. It would also remove the opportunity to develop pipelines for the addition of genetic data to OBIS.

On the other hand, it would also be possible to combine both approaches in the long term. Most likely, sampling and DNA extraction would be the same processes. Once the DNA is acquired and it is ensured that it is safely stored for the long term, further analyses can always be made in case the interest arises. It could be possible for example to complement the targeted gene surveys with periodical metabarcoding of the DNA extract to build baseline data, and information for the purpose of long-term monitoring. General changes in the community composition on the long-term could act as indicators of an unbalance in the ecological communities, and therefore used to induce more detailed biosurveys of the sampling site. A good operational first step could therefore be the detection of a small list of top risk species, which could be expanded to detecting the unknowns at a later stage. Conversely, it is also possible to adopt the opposite approach; A baseline detection approach would be in place, and a set of trigger conditions for more sensitive techniques at a set of sites. Simultaneously, randomised and periodic sensitive technique use should occur as a matter of course.

The OBIS team suggested that the project should carefully consider and consult with the local stakeholders which approach would be more beneficial, and align with the needs and existing initiatives in the area.

3.2 Sampling plan

3.2.1 Initial port survey

Ports are highly complex environments with multiple different ecosystems both above and below the waterline. It is important to choose sampling locations based on discussions with the port authorities. What needs to be taken into account is proximity to high-risk carriers like berthings for international vessels, recreation vessels, cruise liners, industry (gas and oil), or aquaculture while making sure that their activity is not hindered. Another important factor to take into account is the local hydrodynamics, to plan representative sampling points. Therefore, ports or sampling points that are sheltered from the open ocean may be the best for initial detections. In addition. it should be considered if sampling points also outside of the port could be analysed to compare the species assemblages to areas further away from human activity. Port characteristics can be logged based on the Helcom port survey protocol and port information sheet (Helcom/OSPAR). Several port survey reports from Australia were listed . Dr Gilianne Brodie also mentioned that the hydrodynamics in the port has been defined by colleagues at USP, and could help us in defining the sampling grid.

In addition, it was noted that any sampling campaign should follow statistically sound sampling plans, to be able to draw reliable conclusions from the sequencing results derived from these samples.

3.2.2 Sample types

While eDNA samples of the water column alone can provide a wealth of information, sampling of different materials at site is required for the full description of the local community. The different needed sample substrates in the port were discussed, namely settlement plates, plankton samples, filtered water, and sediment. However, it was noted that sampling of several different substrates can quickly get out of hand, and may not be possible to replicate in several ports. Therefore, it was suggested that PacMAN would focus on a few sample types only.

In the case of settlement plates, the group noted that the sample size must somehow account for averaging out patchiness that can introduce high variability between samples. For example, it was noted that based on experience standard ARMS plates often have reproducible community composition, while the smaller CAUs (calcification accretion units), produce more stochastic results. Therefore a size of plate somewhere in between these two may be best to balance ease of sampling and community detection. Generally, however, simple settlement plates were strongly recommended (also developed by OSPAR). These could be standardized for a very simple sampling setup. It was suggested to come up with a simple sample handling protocol. For example, the plates could be collected to zip-lock bags, and eDNA extracted by simple sonication. The idea here being to have a simple, replicable sampling setup that could be easily distributed to many sites. However, monitoring settlement plates was also mentioned to be possibly too late for very early detection, and therefore other methods should be considered for the list of the main target risk species.

If samples were to be taken directly from biofilms on existing structures in the sampling area, it was noted that floating objects appear to have more initial settlers than objects that were attached to the seafloor. Dr Brodie noted that USP has buoys and floating wharfs in the lagoon across the harbor, which might serve as an easily accessible sampling point. A possibility to sample biofouling from vessels from hull cleaning facilities was also suggested, and it was noted that these facilities do exist in Suva.

Ballast water sampling was not recommended for several reasons; obtaining the samples can be difficult, and may not represent a relevant picture of what is important for the introduction of invasive species. It would be hard to distinguish eDNA arising from live vs. dead organisms. In addition due to more efficient regulations in terms of ballast water exchange and treatment, this problem has decreased, and is not expected to be a major pathway for invasive species in the future.

3.3 eDNA analysis

Choices for the metabarcoding studies were also briefly discussed. It was recommended that the COI gene be used as a biomarker, as it is the only widely used gene region with extensive reference databases (BOLD) that can provide classification to the species level. However, due to biases in PCR amplification, other markers may be needed to complement this. The 18S gene tends to work well, but lacks taxonomic resolution. However, if something turns up that is very different from native taxa, then this can be used as an indicator for a need for further studies. In the case of certain groups of organisms like fish or algae, other markers may be more useful, like 23S or 12S ribosomal rRNA gene fragments. It should be expected that in any sample examined from Fiji, the biodiversity will be considerable, and can far exceed that, which is found in temperate environments. Therefore, it is also necessary that the sequencing plan accounts for obtaining enough depth in the sequencing. This can be checked with the use of rarefaction curves, and could be optimised during the trial sampling and sample analysis runs. However, it should be remembered that new invasives would likely be found in the rare biosphere, and therefore even reaching a species plateau does not mean that all invasives were necessarily detected.

4 SESSION 3: Data management

The data management plan (DMP) drawn up in the project document was used as a basis for discussion. Specific parts of the DMP include (i) metadata management, (ii) bioinformatics pipelines, (iii) taxonomic classification, (iv) ASV repository and (v) data packaging for OBIS. Most importantly, the DMP must include answers for questions like which permits are needed for sampling and analyses, who owns the data as well as how will its release be managed, to ensure that possible risks associated with invasive species detections are evaluated by the main stakeholders. The project should develop a clear DMP document before any actions are started. One possibility is to link PacMAN with the recently started FAIR Islands project, and use PacMAN as a use case for the protocols developed there. The FAIR Islands project aims to develop the best case scenarios for DMP for location-based research, and therefore PacMAN fits in this plan perfectly. It was suggested that the DMP should reflect the two possible approaches the project may take: a separate approach for species-specific detections vs data resulting from metabarcoding studies. More detailed discussions are presented in the next sections. DMP parts that were not discussed in detail included: the bioinformatics pipeline and the decision support tool.

4.1 Sequence data management

One of the major bottlenecks in the current data management pipelines is how sequence data will be dealt with. PacMAN will act as a pilot for the addition of sequencing data to the OBIS database, and several decisions need to be made with OBIS to go forward. The importance of ASV repositories was discussed. This would allow a central database where analysed sequences could be stored and sequence classifications/clusters could be managed before ingesting occurrence data to OBIS. An ASV repository is being planned for the Smithsonian, which could be linked to PacMAN as well. In addition ftoor adding data to OBIS it is important to follow the development of the MixS extension for the Darwin Core format, which would allow recording all relevant metadata with each sequence to OBIS.

4.2 Taxonomic identification

Currently, for the CO1 gene, the main authoritative database is the Barcode of Life database (BOLD). This is a curated collection of sequences from voucher specimens around the world. BOLD also bins sequences based on their similarity to the vouchered sequences. If PacMAN could benefit from the bins and bin IDs provided by BOLD it could provide a link to confirmed taxonomic classifications directly. Currently, it is not possible to submit sequences from environmental samples to BOLD, but if the binning algorithm could be sourced, it could be possible to bin local sequences in the same manner. In any case, it is recommended that BOLD be one of the databases used for taxonomic annotation of CO1 sequences.

Another important question relating to taxonomic classification is how the project will deal with unknown sequences that can not be named. It is expected that due to the large diversity of the tropical marine environment, as well as undersampling in the area, a major proportion of the sequences will not be classified to a high taxonomic resolution. Simply naming these sequences based on the closest known taxonomy will likely cluster many sequences, where it is clear based on the sequences themselves that they are not a part of the same group/species. An example following problematic haplotypes of snails in the terrestrial realm was cited here. In this case, bin IDs based on the BOLD database could partially help in reliable classification. In any case it should be noted that classification to linnean names in OBIS, which is currently required, needs to be more flexible to function with sequence data additions.

4.3 Access and Benefit sharing

It was noted that for ecological research, due to the low amounts of sample needed, as well as limited amount of bioprospecting potential of biomarker sequences, the regulations related to ABS and the permissions needed may not be as stringent as in other research cases. The extent of sampling permissions will be investigated by the team in Fiji.

Major decisions to be made in connection with the DMP include how possible material samples will be dealt with, in case these are collected. The marine collection in Fiji could act as the local storage facility for any collected specimens/samples. After these samples are stored however, it needs to be decided who can have access to the samples. Most importantly for PacMAN, clear decisions need to be made on who will own the sequence data and how it will be shared. There are a few possibilities on how open data will be: either fully private, fully open, or filtered based on specific criteria. For a project under the auspices of IOC-UNESCO, the IOC Oceanographic Data Exchange Policy (www.iode.org/policy) applies which states that Member States shall provide timely, free and unrestricted access to all data, associated metadata and products generated under the auspices of IOC programmes. However, there is room for discussion what timely means and we need to evaluate any potential risks that may arise for the local community from publishing the data (including uncertainties and possible misinterpretation), as well as considering that possibility of generalizing or temporarily holding off on releasing accurate information of endangered species observations

5 SESSION 4: Synthesis and workplan

A number of action items were identified as necessary to make further progress in the project.

1. A comprehensive needs assessment of the major local stakeholders
2. A list of target species based on literature review of surveys done in other locations on the Pacific, including North Australia, Samoa, Guam, Pearl Harbour, Hawaii
3. An evaluation of the completeness of the genetic reference databases based on these identified lists
4. A decision tree for the different possibilities of the monitoring plan as well as use cases for the monitoring plan.

The virtual synthesis session was postponed to allow for creating clear action plans and documents to discuss for the next session. Also a visual representation of the possible routes for monitoring and analyses will be developed to communicate the project possibilities and plans to stakeholders prior to the advisory board meeting.

1. Annexes

1. Annex 1 Meeting Agenda

Agenda

Session 1: Mon 23 November 2020 8-10.30 pm UTC

Introduction to PacMAN, coordinating team, scientific experts and exchange of preliminary information of the local stakeholders.

Timeline

8-8.30 Welcome and Introduction

8.00-8.05	Welcome words	Ward Appeltans
8.05-8.10	Introduction to PacMAN	Saara Suominen
8.10-8.15	Introduction to the USP team	Joape Ginigini
8.15-8.30	Preliminary stakeholder needs in Fiji	Gilianne Brodie

8.30-9.10 Monitoring design

8.30-8.40	*Monitoring design and genetic probes for invasive species in Australia*	Nic Bax
8.40-8.50	*Molecular tools for detecting marine pests*	Craig Sherman
8.50-9.00	*Southwest Pacific Regional OBIS node*	Kevin MacKay
9.00-9.10	*Geans Project*	Pascal Hablutzel

9.10-9.15 5 minute break

9.15-9.35 ARMS program development

9.15-9.25	*Global ARMS Program and Ocean DNA Initiative*	Chris Meyer
9.25-9.35	*Status update for the European ARMS-MBON*	Matthias Obst

9.35-10.15 Data management initiatives

9.35-9.45	*Biodiversity Data Management*	John Deck
9.45-9.55	*Marine Biodiversity Observation Network (MBON)*	Frank E Muller-Karger
9.55-10.05	*TBA*	Neil Davies
10.05-10.15	*TBA*	Pier Luigi Buttigieg

Session 2: Tue 24 November 2020 8-10 pm UTC

Monitoring plan: based on preliminary stakeholder assessment and shared expertise in the project discuss sampling, sample handling and molecular work. We will start with reviewing again the current knowledge of stakeholder needs in Fiji. This will be followed by an open session and will allow discussion into the major points that are needed for building a functioning monitoring plan. The following questions can be used to guide the discussion:

Free discussion on the monitoring plan:

- What are the most important pitfalls that may be encountered during the initial biodiversity survey?
- What are the main recommendations for the initial monitoring plan?
  - Timeline
  - Sampled substrates
  - Analysed marker genes and primers?

- What is the best approach to building an invasive Species reference database
  - Existing databases in the Pacific and sequences?

- Opportunities for collaboration to build a standard method?

Session 3: Mon 30 November 2020 8-10 pm UTC

Monitoring plan: Discuss elements of Data management plan and policies and bioinformatics pipeline. This will be an open session and will allow discussion into the major points that are needed for building a functioning monitoring plan. The following questions can be used to guide the discussion:

- What are the possible approaches to manage data
    - Platforms/existing workflows at research institutions
- What is needed to incorporate sequencing data to the data management workflow?
    - Data formats, important steps/considerations?
- What are the existing bioinformatics pipelines?
- What are the most important considerations for the bioinformatics pipeline
    - OTU/ASV algorithm?
- What are the policies that have to be taken into account and what are the best options?
    - Data policy
    - Nagoya Protocol
    - Other policies?

Session 4

Synthesis, discussion of outcomes, agreement on work plan and wrap-up. In this session the PacMAN team will discuss a synthesis of the discussions in both sessions 2 and 3. Additions to each report can be made at this point as well as discussion regarding the next steps.

Annex 2 Participants list the First PacMAN Scientific Workshop, 2020

Scientific Experts

Mr. Kevin MACKAY
Marine Data Manager
Environmental Information
National Institute of Water and Atmosphere
Private Bag 99940, Newmarket
Wellington
Auckland 1149
New Zealand

Dr. Pier Luigi BUTTIGIEG
Data Scientist
Helmholtz Metadata Collaboration
GEOMAR | Helmholtz Centre for Ocean Research Kiel
Duesternbrooker Weg 20
24105 Kiel
Germany

Christopher MEYER
Research Zoologist and Curator
Invertebrate Zoology
National Museum of Natural History, Smithsonian Institution
12th and Constitution Ave., NW
Washington DC, 20560-0163
United States

Dr. Pascal HABLÜTZEL
Scientist
Research Department
Vlaams Instituut voor de Zee
Wandelaarkaai 7
8400 Oostende
Belgium

Assoc. Prof. Matthias OBST
Scientist
Department of Marine Sciences
Göteborg University, Department of Marine Ecology
Box 461
40530 Gothenburg
Sweden

John DECK
Research Programmer
Berkeley Natural History Museums

1007 Valley Life Sciences Building
Berkeley, California 94720
United States

Dr. Frank MULLER-KARGER
Professor
College of Marine Science, Institute for Marine Remote Sensing
University of South Florida, College of Marine Science
140 7th Ave. South, Saint Petersburg, FL 33701
St. Petersburg, Florida 33701
United States

Dr. Neil DAVIES
Director
UC Gump South Pacific Research Station
BP-244
UCB
98728 Maharepa Moorea
French Polynesia

Dr. Nic BAX
Oceans & Atmosphere
CSIRO Oceans and Atmosphere Information and Data Centre, Hobart
GPO Box 1538
Hobart TAS 7001
Australia

Assoc. Prof. Craig SHERMAN
Associate Professor
School of Life and Environmental Sciences
Deakin University
Deakin University, Pigdons Road
Waurn Ponds VIC 3216
Australia

PacMAN project team

Dr. Gilianne BRODIE
Deputy Director IAS
Institute of Applied Sciences
University of the South Pacific, Faculty of Science Technology & Environment, Institute of Applied Sciences
Lower Laucala Campus
Suva
Fiji

Mr. Joape GINIGINI
Senior Scientific Officer, Institute of Applied Sciences
Institute of Applied Sciences
The University of the South Pacific
The University of the South Pacific

Institute of Applied Sciences, USP Marine campus,
Suva
Fiji

Mrs. Paayal KUMAR
Researcher
Pacific Natural Products Research Centre
Institute of Applied Sciences
University of the South Pacific
Lower Campus
Suva
Fiji

Mr. Ward APPELTANS
Project Manager OBIS, GOOS Biology & Ecosystems, IOC Capacity Development
UNESCO / IOC Project Office for IODE
Wandelaarkaai 7
Pakhuis 61
8400 Oostende
Belgium

Mr. Pieter PROVOOST
OBIS Data Manager
UNESCO / IOC Project Office for IODE
Wandelaarkaai 7
Pakhuis 61
8400 Oostende
Belgium

Dr. Saara SUOMINEN
Scientific Officer
OBIS
UNESCO / IOC Project Office for IODE
Wandelaarkaai 7
Pakhuis 61
8400 Oostende
Belgium