

Tapping two decades of ecological research in PNG: Creating a framework for data mobilisation at New Guinea Binatang Research Center

Programme: BID
Project ID: BID-PA2020-007-INS
Project lead organization: New Guinea Binatang Research Center
Project implementation period: 1/9/2021 - 28/2/2023
Report approved: 1/7/2022

Narrative Midterm report

Executive Summary

To date our project has focussed on preparing New Guinea Binatang Research Centre's GBIF Officer, Mr Phillip for his role as primary data handler. Mr Philip is steadily achieving greater competency in his role but still requires assistance with technical issues. We estimate that the GBIF Officer will be ready for totally independent work in the coming months as progress over the last two months has been significant. The team are also pleased with the progress we have made towards cleaning and formatting some of our datasets.

Progress has been monitored via monthly Zoom meetings between Mr Philp, Dr Butterill, and Dr Redmond. This involves a progress summary given by Mr Philip, followed by an advisory and troubleshooting session where Dr Butterill and Dr Redmond provide feedback and assistance with any issues.

Mr Philip experienced a number of unfortunate delays when attempting to secure his GBIF Biodiversity Data Mobilization digital badge. This had a knock on effect of delaying the project timeline, but we believe we have made the necessary adjustments to ensure project delivery moving into the second half of the project.

We do not require any official changes to the project at this moment, however we are working internally on reassessing our workflow in light of recent delays. We are setting up additional internal support system for the GBIF Officer within BRC facilities where our senior staff on the ground in Papua New Guinea can assist and monitor efforts more closely than is possible remotely from Europe.

Progress against milestones

Has your project published at least one dataset through GBIF.org?: No

Dataset published:

Dataset	DOI
Sap-sucking herbivores (Auchenorrhyncha: Hemiptera) reared from Papua New Guinean rainforest	

Has at least one member of your project team received certification following the BID capacity enhancement workshop?: Yes

Name of the workshop participant: Frank Philip

Certification obtained: Basic Badge

Report on Activities

Activity progress summary

Training resident GBIF officer:

This has been the area of focus during the first half of the project. Mr Frank Philip has greatly increased his proficiency in recent months, despite having experienced challenges along the way. He has received advice and training from Mr Penniel Lamei, Dr Phillip Butterill and Dr Conor Redmond. This training and supervision has been primarily based around data management, use of BRCs internal database 'Insecta Nuigini', data checking, working with taxa specialists at BRC, quality control, networking with data authors for confirmation/clarification of data issues, using GBIF toolkits and interfaces, developing teaching activities and improving time management and goal orientated workflows. Mr Philip has since improved greatly in these areas and has developed more robust timelines moving forward into the second half of the project. We have also now assigned some senior staff to be on-site supervisors for Mr Philip, which will further increase the team efficiency and increase speed of dataset preparation. Mr Philip has dedicated the majority of his time to data checking, quality assessment, formatting and cleaning datasets. As these datasets have come from a variety of authors, Mr Philip has found data quality control to be a labour intensive task and at present is the primary source of project delays. He has also liaised with taxa team leaders throughout BRC who are assisting him with data cleaning, while Mr Philip is explaining GBIF protocols.

Training for BRC taxa leaders and key staff:

Delayed partially. Rather than full training, two primary taxa specialists have received briefings and introductory training on formatting and data cleaning. This knowledge is being applied currently for the data quality checks by the taxa specialists.

Cause: Ensuring GBIF officer is adequately prepared for knowledge comprehensive sharing exercises with all taxa specialists and team leaders.

New Expected

Date: July 2022

This activity is still in its early phases where Mr Philip is liaising with individual taxa leaders regarding data formatting and quality control. We will host an official training exercise once we are happy that the GBIF Officer is fully capable of leading this exercise. We estimate this will occur in July following our upcoming student training course.

Quality control and refinement of BRC datasets:

This activity has been the most labour intensive element to date given that the large number of authors and diversity of datasets. We estimate that this task requires an additional 4-6 weeks to complete based on recent progress. In response to these delays, data manager Dr Butterill and project manager Dr Redmond have stepped in to assist with this process.

Review of DNA barcodes:

Work towards this goal is ongoing and progressing. To date we have been focussing on Lepidoptera barcodes within the Barcode of Life Data (BOLD) system. We have been working with external colleagues in the Smithsonian Institution, Washington DC to verify data integrity and quality.

Review of photography and improvements:

The BRC GBIF Officer has been tasked with this activity and it is scheduled to occur concurrently with the preparation and submission of each dataset as relevant. However to date this activity has not been relevant, estimated start date for this is now July 2022.

Mobilisation of BRC datasets:

To date we have mobilised our 'Madang Sap Suckers' dataset. This dataset was used as our primary training dataset for our GBIF Officer training exercises. This activity was supervised by data manager Dr Butterill. We had hoped to mobilise additional datasets by now but we decided to slow down this process for the sake of quality control and data robustness. Moving forward we have a series of time bound uploading goals and remain confident all datasets will be mobilised in the second half of the project. We believe the decision to delay uploads was the correct one in light of internal challenges we faced with data handling proficiencies.

Student Training Course Workshop:

Advanced date: From August 2022, to June 2022

The student training course is scheduled to take place June 2022. Dr Redmond will host a data analysis program which will see GBIF Officer share his experience and knowledge with participants in a lecture/workshop. In addition to the timing of the course, the nature of the course has also changed. Rather than hosting our usual international student course, this year will be exclusively Papua New Guinean students. The reason for this is COVID travel issues and liabilities for international students from Czech Republic. However we are very excited to host this new form of training course, in particular because this will be an inaugural training course targeting female biologists. This is a USAID funded program entitled "Empowering female researchers in research and conservation".

In person or online workshop with other national agencies:

Scheduled for later in the year

Mobilisation of EntomID-PNG database & FRI Insect Database

Delayed

Cause: Training of GBIF Officer delays with knock on effect to these activities.

New Expected Date: Sept 2022 onward

Completed activities

Report on Deliverables

Deliverables progress summary

Datasets Mobilised

Madang Sap Suckers.

This is our first dataset to be mobilised. BRC coordinated with SPREP who will host this dataset. Our GBIF officer and Data Manager coordinated efforts to ensure correct data formatting and data quality.

Remaining Datasets

Delayed

Cause: Training of GBIF Officer delays with knock on effect to these activities.

New Expected Date: Ongoing from now until Jan 2023

GBIF promotion and workshop at student course

Advanced

Now scheduled for June 2022, advanced from August 2022.

Promotion and workshop with members of key national institutions:

Scheduled for closer to project closeout

GBIF Officer

Partially Completed

Mr Philip has made significant progress in this role but we must categorise this deliverable as only partially completed as management believe that additional supervision and advise is required before he is ready for independent work. Nevertheless we are confident that this stage will be reached in short order, estimating July 2022.

"Guilds" Data Paper

Scheduled for closer to project closeout

Progress towards deliverables

Madang Sap Suckers

Dataset type: Occurrences

Dataset scope: A relatively small dataset of sap-sucking insects (Auchenorrhyncha: Hemiptera) reared from rainforest vegetation

Number of records: 1,857

Data holder: SPREP

Data host institution: SPREP

% complete: 95%

Status update: The dataset & metadata has been submitted to SPREP via their IPT, and mapping to Darwin Core completed. We expect SPREP to do a quick review before publishing

DOI:

Expected date of publication: 2022-06-03

Lowlands Betadiversity

Dataset type: Occurrences

Dataset scope: A large dataset of reared caterpillars (Lepidoptera)

Number of records: 75,000

Data holder: New Guinea Binatang Research Centre

Data host institution: SPREP

% complete: 55%

Status update: The data is in good condition, but will be an Occurrence type rather than Sample Event due to data structure. The main challenge is to update specimen IDs that have arisen from extensive barcoding projects - this work has begun and is progressing nicely, but slowly.

DOI:

Expected date of publication: 2022-07-01

Madang Adult Leaf Chewers

Dataset type: Occurrences

Dataset scope: Chewing insect larvae (Coleoptera, Phasmida, Orthoptera) reared from rainforest vegetation

Number of records: 32,000

Data holder: New Guinea Binatang Research Centre

Data host institution: SPREP

% complete: 75%

Status update: Will be Occurrence rather than Sampling Event type. Cleaning underway and progressing nicely. This will most likely be the next dataset.

DOI:

Expected date of publication: 2022-07-01

Simbu Lepidoptera

Dataset type: Occurrences

Dataset scope: Caterpillars (Lepidoptera) feeding on rainforest vegetation

Number of records: 13,000

Data holder: New Guinea Binatang Research Centre

Data host institution: SPREP

% complete: 60%

Status update: Changed to Occurrence rather than Sampling Event type. The main challenge is to update specimen IDs that have arisen from extensive barcoding projects, which is progressing steadily.

DOI:

Expected date of publication: 2022-08-01

Madang Lepidoptera

Dataset type: Occurrences

Dataset scope: Large rearing project of rainforest caterpillars

Number of records: 49,000

Data holder: New Guinea Binatang

Data host institution: BRC

% complete: 35%

Status update: Changed to occurrence dataset. Involves updating of IDs based on barcoded specimens, which is in progress

DOI:

Expected date of publication: 2023-01-16

Madang Leaf Miners

Dataset type: Occurrences

Dataset scope: Large-scale rearing of rainforest leaf miners

Number of records: 28,000
Data holder: New Guinea Binatang Research Centre
Data host institution: SPREP
% complete: 40%
Status update: Occurrence dataset. Update of record IDs in progress
DOI:
Expected date of publication: 2022-12-12

Madang Root Chewers

Dataset type: Occurrences
Dataset scope: Experimental rearing of root-feeding beetles
Number of records: 2,500
Data holder: New Guinea Binatang Research Centre
Data host institution: SPREP
% complete: 75%
Status update: Occurrence dataset. IDs & data cleaning in progress
DOI:
Expected date of publication: 2022-08-15

Madang Bark Beetles

Dataset type: Occurrences
Dataset scope: Host specificity of tropical insects feeding on angiosperms – Wood borers (Coleoptera)
Number of records: 2,000
Data holder: New Guinea Binatang Research Center
Data host institution: SPREP
% complete: 80%
Status update: Data cleaning nearing completion
DOI:
Expected date of publication: 2022-08-29

Wanang Ants

Dataset type: Occurrences
Dataset scope: Diversity of tropical arboreal communities – Ants (Hymenoptera)
Number of records: 3,000
Data holder: New Guinea Binatang Research Center
Data host institution: SPREP
% complete: 55%
Status update: Data cleaning in progress
DOI:
Expected date of publication: 2022-09-12

Mt. Wilhelm Ants

Dataset type: Occurrences
Dataset scope: Betadiversity of tropical ant communities – Ants (Hymenoptera).
Number of records: 3,500
Data holder: New Guinea Binatang Research Center
Data host institution: SPREP
% complete: 60%
Status update: Data cleaning in progress
DOI:
Expected date of publication: 2022-09-26

PNG plants

Dataset type: Occurrences
Dataset scope: Plants of PNG, from multiple ecological projects over the last 20 years.
Number of records: 15,000
Data holder: New Guinea Binatang Research Center
Data host institution: SPREP
% complete: 40%
Status update: IDs updates and cleaning in progress
DOI:
Expected date of publication: 2023-02-13

Wanang Lepidoptera

Dataset type: Occurrences
Dataset scope: Successional diversity of tropical caterpillars (Lepidoptera).
Number of records: 17,000
Data holder: New Guinea Binatang Research Center

Data host institution: SPREP
% complete: 55%
Status update: ID updates & data cleaning progressing
DOI:
Expected date of publication: 2022-10-03

Wanang Wood Borers

Dataset type: Occurrences
Dataset scope: Successional diversity of tropical wood-boring insects (Coleoptera).
Number of records: 15,000
Data holder: New Guinea Binatang Research Center
Data host institution: SPREP
% complete: 50%
Status update: ID updates & data cleaning in progress
DOI:
Expected date of publication: 2022-10-17

Other deliverables

EntomID-PNG Database

Description: A database of museum insect specimens (predominantly Coleoptera, Lepidoptera and Hymenoptera) collected by various institutions all over PNG.
% complete: 15%
Status update: Work has started in that a plan has been developed. Work will continue when time permits
Sources of verification: N/A

Communications and visibility

We will begin focussing on communications once we have established our internal mobilisation capacity. This will take the form of publishing information on our BRC webpage, Facebook page and on our project page. This is scheduled for the second half of the project.

Monitoring and evaluation

Monitoring and evaluation findings

We implemented monthly meetings between the project manager, data manager and GBIF Officer to monitor progress. At these meetings the GBIF Officer provided overviews and reports on progress, while the project and data manager provided guidance and troubleshooting support. We experienced a set back to the project timeline due to issues with our GBIF Officer achieving the Mobilization digital badge. Despite a slow start, we believe our GBIF officer has seen significant improvements in proficiency over the last couple of months. The team are pleased to see this and believe this is the key for successful project implementation.

We acknowledge there has been delays to the project timeline as a result, however we believe that taking extra time to establish a solid foundation is a worthwhile task which will ultimately bring higher data quality, which is a top priority for us.

The team has implemented a number of changes to project monitoring. We have taken the decision to bring in additional BRC management to help with supervision and guidance. This will be cofunded by BRC.

We acknowledge that we somewhat underestimated the complexities of formatting a large variety of datasets, all of which were authored by different individuals or teams. One issue we have encountered is contacting the original authors and support staff to verify data quality issues.

Steps have also been taken to improve team communication. We experienced some delays arising from cross-continental collaboration between Europe and PNG. Again however we have taken steps to improve this, specifically by bringing additional BRC management.

BRC are confident that we can deliver a productive second half of the project, which will be launched from the solid foundation we have taken time to develop.

Impact of COVID-19 pandemic on project implementation

COVID-19 has brought about delays to BRCs programs generally. It's impact on this project has been limited, but manifested primarily through inefficient remote team interactions and difficulties with attaining work/travel Visas for PNG (for European based project members Dr Redmond and Dr Butterill).

GBIF leads the Biodiversity Information for Development (BID), a programme funded by the European Union. The programme provides supplementary support for activities addressing the needs of regional researchers and policymakers through mobilization and use of biodiversity data.

