

Mobilising data and building capacity for identification and monitoring of ecologically important insect taxa in Malaysia and Singapore

Programme:BIFA

Project ID: BIFA6_032

Project lead organization:Asian School of the Environment

Project implementation period:1/11/2021 - 31/3/2023

Report approved: 31/5/2023

Narrative Final report

Executive Summary

We have published 4 datasets, one checklist, one occurrence, one sampling-event and one resource metadata dataset. To date, we have met four of our project collaborators and presented on GBIF and our project to them in addition to discussing how we can integrate their collections into our project.

After improved understanding of our data, we changed our dataset deliverables once prior to the mid-term reporting and also once prior to the final report. This was discussed with Lily and she confirmed this with the GBIF Secretariat. Because most of our data are derived from a similar pool of sources, it has been difficult to decide how to split certain datasets. Our data are essentially nested Checklist>Occurrences>Sampling event. However, GBIF seems to nest Occurrences in Sampling event which means that data end up duplicated or the Occurrence data incomplete with only data without protocols in this dataset, which is misleading.

We also faced challenges in publishing the remaining deliverables for several reasons: (i) complications and delays were met during the sequencing portion of the project. We have only just received the data from sequencing companies and have yet to interpret them. This means that the sequencing datasets can only be published at a later date. ii) the Singapore dung beetle data lacks accurate species IDs and reliable data. This means that until further verification is completed, we cannot publish an accurate dataset and have instead published a resource metadata dataset to inform data users of the incoming datasets post-project.

Progress against milestones

Has your project completed all planned activities?: No

Rationale: Plans to travel to Helsinki and to Japan to visit collections and taxonomists were impacted by Covid-19. Despite this, we still have strong intentions to visit both places in the future to confirm species identifications and to update these datasets.

Has your project produced all deliverables?: No

Rationale: We have 4 deliverables that need to be moved to 'post-project deliverables': (i) molecular sequences of dung beetles of Sabah (.ext), (ii) molecular sequences of dung beetles of Singapore (.ext) - These deliverables could not be completed owing to complications during molecular work and duration it takes for sequencing companies to get back to us. At present, we have only just received the data back from the sequencing companies and are just beginning to interpret the data. The next stage would be to assemble the molecular datasets to complete these deliverables. (iii) Checklist of dung beetles of Singapore (.core), (iv) Occurrence of dung beetles of Singapore (.ext) - These deliverables could not be met due to unreliable specimen IDs and unreliable data. Until we confirm the species IDs with taxonomic experts, we are unable to authenticate the bulk of the data we have presently and hence are not comfortable with publishing with partial dataset which may mislead data users.

Activity implementation summary

We have completed majority of the activities that we have described in our project proposal but will here explain why certain activities were not achieved or will be delayed.

Activities

Project set up, training & BIFA workshop - Marx Yim (Project Officer) received his certification (Advanced badge) on 6 Apr 2022 following the Asia Data mobilisation course and held meetings with our project collaborators from Lee Kong Chian Natural History Museum, Singapore, Helsinki Museum of Natural History, Forest Research Centre (FRC), and Universiti of Sabah (UMS) in Sabah to consolidate the plan to implement project - *Completed*.

Collate datasets from literature and specimens from Singapore - Completed data digitisation process from specimens and literature held in various collections and sources - *Completed*.

DNA barcoding of Singapore and Sabah specimens - Collected key specimens from our collaborators in Singapore and Sabah, Malaysia and sequenced specimens to acquire sequencing data - *Completed*. Deliverables associated to this activity will be completed post-project.

Visits to Sabah and collation of datasets - Visited FRC and UMS in Sabah in May 2022 and Oxford University Museum of Natural History (OUMNH) in March 2023 to catalogue and digitise their collection. We also gave presentations of our GBIF project to our collaborators and obtained specimens for sequencing during these visit. We also learned of additional datasets that could be included into our published datasets. We did not get to visit Helsinki due to Covid-19 but instead visited OUMNH as an extra activity. Another extra future activity is that we intend to visit Teruo Ochi and Masahiro Kon, dung beetle taxonomists in Japan in 2023/2024 to help with identifications and further publications of these datasets - *Completed*.

Construction of user-friendly keys and guides for Singapore and Malaysia - At present, we have a draft for the book containing the user-friendly key and guide for the dung beetles of Sabah, Malaysia. This requires further editing and corrections by our collaborators and editors before publishing - *Partial completion*.

Data input and publishing in GBIF - We have published 4 out of the 7 datasets originally suggested and we intend to continue working fully on the creation and publication of the remaining datasets post-project - *Partial completion*.

The delays can be explained in two ways:

- Molecular Sequencing datasets - These deliverables could not be completed owing to complications during molecular work and duration it takes for sequencing companies to get back to us. At present, we have only just received the data back from sequencing companies and are just beginning to interpret the data. The next stage would be to assemble the molecular datasets to complete these deliverables. We hope to publish these molecular datasets by the end of 2023.
- Singapore datasets - These deliverables could not be met due to unreliable specimen IDs and unreliable data. Until we confirm the species IDs with taxonomic experts, we are unable to authenticate the bulk of the data we have presently and hence are not comfortable with publishing with partial dataset which may mislead data users. We have instead published a resource metadata dataset in place of a checklist (.core) and occurrence (.extension). We hope to publish these Singapore datasets after we have visited taxonomists Teruo Ochi and Masahiro Kon in Japan and also after referencing type specimens from collections in London, Oxford and Paris, sometime in 2024.

Completed activities

Activity name: Publication of 'Taxonomic Checklist of the Dung beetles (Coleoptera: Scarabaeidae: Scarabaeinae) in Sabah'

Description: This dataset was published and registered on GBIF on 13 April 2022.

Start Date - End Date: 1/11/2021 - 31/3/2023

Verification Sources: <https://www.gbif.org/dataset/c41029b7-b950-48f3-be69-60267043e020>

Activity name: Publication of 'Occurrence Records of Dung Beetles (Coleoptera: Scarabaeidae: Scarabaeinae) from Sabah, Malaysia Borneo (I)'

Description: This dataset was published and registered on GBIF on 29 March 2023.

Start Date - End Date: 1/11/2021 - 31/3/2023

Verification Sources: <https://www.gbif.org/dataset/822991fc-0532-46b0-8dda-d17dd6cdb8df>

Activity name: Publication of 'Occurrence Records of Dung Beetles (Coleoptera: Scarabaeidae: Scarabaeinae) from Sabah, Malaysia Borneo (II)'

Description: This dataset was published and registered on GBIF on 29 March 2023.

Start Date - End Date: 1/11/2021 - 31/3/2023

Verification Sources: <https://www.gbif.org/dataset/164f242c-55fd-4227-97bb-28c031c4e3d4>

Activity name: Publication of 'Records of Dung Beetles (Coleoptera: Scarabaeidae: Scarabaeinae) of Singapore'

Description: This dataset was published and registered on GBIF on 29 March 2023

Start Date - End Date: 1/11/2021 - 31/3/2023

Verification Sources: <https://www.gbif.org/dataset/d79b210b-07bb-4cb8-9545-fb9a78b455f9>

Activity name: Collaborator meeting & institution visit (Universiti Malaysia Sabah)

Description: This meeting is in connection with the BIFA GBIF project Mobilizing data on ecologically important insects in Malaysia and Singapore project funded by the Biodiversity Information Fund for Asia (BIFA).

On the 8 of April 2022, Dr Eleanor Slade (Project Investigator) met with our project collaborators, Prof Monica Suleiman and Puan Petherine Jimbau at ITBC Universiti Malaysia Sabah (UMS) to discuss our BIFA GBIF project and to discuss how to integrate the Borneensis dung beetle collections.

In May 2022, Dr Eleanor Slade, Marx Yim (Project Officer) and Ong Xin Rui (PhD student) will spend a week at UMS. We will give a presentation on our BIFA GBIF project to our UMS colleagues and work with them to examine the collections to determine the mobilisation efforts necessary.

Start Date - End Date: 8/4/2022 - 8/4/2022

Verification Sources: <https://www.gbif.org/event/799de3-c9cb-4e91-93e0-b7b5404/gbif-bifa-meeting-with-collaborator-universiti-malaysia-sabah>

Activity name: Collaborator meeting & institution visit (Institute of Tropical Biology and Conservation - Borneensis Collection)

Description: This meeting is in connection with the BIFA GBIF project Mobilizing data on ecologically important insects in Malaysia and Singapore project funded by the Biodiversity Information Fund for Asia (BIFA).

Between the 23 to 27 May 2022, Dr Eleanor Slade (PI), Marx Yim (Project Officer), Ong Xin Rui (PhD student) and Chiew Li Yuen (Postdoc researcher) visited the Institute of Tropical Biology and Conservation (ITBC) at Universiti Malaysia Sabah (UMS) for meetings with collaborators and to examine the Borneensis dung beetle collections.

On one of the days, we were invited to speak to the staff and students of ITBC in a knowledge sharing seminar. Marx Yim spoke about our GBIF project "Mobilising data on ecologically important insects in Malaysia and Singapore" where he shared more about the objectives, function and application of GBIF and also about our project in-depth.

This was followed by a closed-door meeting with the staff of ITBC where we discussed progress and ongoing processes of the project and shared our findings on how best to integrate the BORNEENSIS dung beetle collections into the project. The meeting also involved understanding the resources available at ITBC and how it can be utilised effectively for the project. We also discussed best practices in digitisation workflows following the examination of their collection. Among those present were Mdm Puan Petherine Jambau, Dr. Liew Thor Seng, Dr. Evan Quah Seng Huat and Dr. Ng Ting Hui

Start Date - End Date: 23/5/2022 - 27/5/2022

Verification Sources: <https://www.gbif.org/event/015238-69a8-41e5-8a20-15618e1/gbif-bifa-meeting-with-institute-of-tropical-biology-and-conservation-borneensis-collection>

Activity name: Collaborator meeting & institution visit (Forest Research Centre Sepilok,)

Description: This meeting is in connection with the BIFA GBIF project Mobilizing data on ecologically important insects in Malaysia and Singapore project funded by the Biodiversity Information Fund for Asia (BIFA).

Between the 1 to 3 June 2022, Marx Yim (Project Officer) visited the Forest Research Centre (FRC) Sepilok in Sandakan, Sabah for meetings with collaborators and to examine their dung beetle collections.

On one of the days, he was invited to speak to the staff and researchers of FRC in a knowledge sharing seminar. Marx Yim spoke about our GBIF project "Mobilising data on ecologically important insects in Malaysia and Singapore" where he shared more about the objectives, function and application of GBIF and also about our project in-depth.

This was followed by a closed-door meeting with Dr. Arthur Chung where we discussed progress and ongoing processes of the project and shared our findings on how best to integrate the FRC dung beetle collections into the project. The meeting also involved understanding the resources available at FRC and how it can be utilised effectively for the project. We also discussed best practices in digitisation workflows following the examination of their collection.

Start Date - End Date: 1/6/2022 - 3/6/2022

Verification Sources: <https://www.gbif.org/event/b8dacf-05c0-497c-822c-5a3a465/gbif-bifa-meeting-with-forest-research-centre-sepilok-sabah>

Activity name: Collaborator meeting & institution visit (Oxford University Museum of Natural History)

Description: This meeting is in connection with the BIFA GBIF project Mobilizing data on ecologically important insects in Malaysia and Singapore project funded by the Biodiversity Information Fund for Asia (BIFA). This visit is in place of our museum visit to Helsinki.

Between the 13 to 17 March 2023, Marx Yim (Project Officer) visited the Oxford Museum of Natural History in Oxford for meetings with collaborators and to examine their dung beetle collections. Key specimens were catalogued and marked for future digitisation and imaging efforts. Meetings were held with curators.

Start Date - End Date: 13/3/2023 - 17/3/2023

Verification Sources: <https://teelabntu.wixsite.com/home/post/examining-the-dung-beetle-collections-at-oxford-university-museum-of-natural-history>

Report on Deliverables

Production of Deliverables - Summary

We have completed about half of the deliverables that we have described in our project proposal but will here explain why certain deliverables were not achieved or will be delayed:

Sabah

(Published) Checklist of the dung beetles of lowland Sabah, Malaysia (Checklist - Core)

(Published) Dung beetle distribution records for Sabah (Occurrence - Core)

(Published) Dung beetles of lowland Sabah database (Sampling-event - Core)

(Not published - expected publication: 2023) DNA sequence data for dung beetles of Sabah

Singapore

[NEW] (Published) Records of dung beetles from Singapore (Resource metadata)

(Not published - expected publication: 2024) Checklist of the dung beetles of Singapore (Checklist - Core)

(Not published - expected publication: 2024) Dung beetle distribution records for Singapore and Bornean Malaysia (Occurrence - Extension)

(Not published - expected publication: 2024) DNA sequence data for dung beetles of Singapore (Sequencing - Extension)

The delays can be explained in two ways:

Molecular Sequencing datasets - These deliverables could not be completed owing to complications during molecular work and duration it takes for sequencing companies to get back to us. At present, we have only just received the data back from sequencing companies and are just beginning to interpret the data. The next stage would be to assemble the molecular datasets to complete these deliverables. We hope to publish these molecular datasets by the end of 2023

Singapore datasets - These deliverables could not be met due to unreliable specimen IDs and unreliable data. Until we confirm the species IDs with taxonomic experts, we are unable to authenticate the bulk of the data we have presently and hence are not comfortable with publishing with partial dataset which may mislead data users. We have instead published a resource metadata dataset in place of a checklist (.core) and occurrence (.extension). We hope to publish these Singapore datasets after we have visited taxonomists Teruo Ochi and Masahiro Kon in Japan and also after referencing type specimens from collections in London, Oxford and Paris, sometime in 2024.

Production of deliverables

Title: Taxonomic Checklist and Occurrences of the Dung beetles (Coleoptera: Scarabaeidae: Scarabaeinae) of Sabah, Malaysia

Type: Dataset

Status update: Checklist (.core) Reference (.extension) To be added Molecular (.extension) Name change of dataset title yet to be reflected.

Dataset scope: Dung beetles}

Expected number of records: 192

Data holder: GBIF, TEE Lab - NTU/ASE, Eleanor Slade, Marx Yim, Xin Rui

Data host institution: Asian School of the Environment

Sampling method: Data extraction from taxonomical and ecological literature, published and unpublished datasets and specimen collections. The data was entered according to the Darwin Core Archive (DwC-A) formats and controlled vocabulary was used. Data was cleaned using OpenRefine and Excel and validated through 'GBIF species-lookup tool' and 'GBIF data validator'.

% complete: 100

DOI: <https://doi.org/10.15468/r8s697>

Expected date of publication:

Title: Occurrence Records of Dung Beetles (Coleoptera: Scarabaeidae: Scarabaeinae) from Sabah, Malaysia Borneo (I)

Type: Dataset

Status update: Sampling-event (.core) To be updated as needed.

Dataset scope: Dung beetles, Scarabaeinae, Sabah, Sampling-event}

Expected number of records: 15585

Data holder: GBIF, TEE Lab - NTU/ASE, Eleanor Slade, Marx Yim, Ong Xin Rui

Data host institution: Asian School of the Environment

Sampling method: Data extraction from taxonomical and ecological literature, published and unpublished datasets and specimen collections. The data was entered according to the Darwin Core Archive (DwC-A) formats and controlled vocabulary was used. Data was cleaned using OpenRefine and Excel and validated through 'GBIF species-lookup tool' and 'GBIF data validator'.

% complete: 100

DOI: <https://doi.org/10.15468/kdt8vd>

Expected date of publication:

Title: Occurrence Records of Dung Beetles (Coleoptera: Scarabaeidae: Scarabaeinae) from Sabah, Malaysia Borneo (II)

Type: Dataset

Status update: Occurrence (.core) To be updated as needed.

Dataset scope: Dung beetles, Scarabaeinae, Sabah, Occurrence}

Expected number of records: 882

Data holder: GBIF, TEE Lab - NTU/ASE, Eleanor Slade, Marx Yim, Ong Xin Rui

Data host institution: Asian School of the Environment

Sampling method: Data extraction from taxonomical and ecological literature, published and unpublished datasets and specimen collections. The data was entered according to the Darwin Core Archive (DwC-A) formats and controlled vocabulary was used. Data was cleaned using OpenRefine and Excel and validated through 'GBIF species-lookup tool' and 'GBIF data validator'.

% complete: 100

DOI: <https://doi.org/10.15468/gutk64>

Expected date of publication:

Title: Records of Dung Beetles (Coleoptera: Scarabaeidae: Scarabaeinae) of Singapore

Type: Dataset

Status update: To be converted to Checklist (.core) and to include occurrence (.extension)

Dataset scope: Dung beetles, Scarabaeinae, Singapore, Resource-metadata}

Expected number of records: 0

Data holder: GBIF, TEE Lab - NTU/ASE, Eleanor Slade, Marx Yim, Ong Xin Rui

Data host institution: Asian School of the Environment

Sampling method: Data extraction from taxonomical and ecological literature, published and unpublished datasets and specimen collections. The data was entered according to the Darwin Core Archive (DwC-A) formats and controlled vocabulary was used. Data was cleaned using OpenRefine

and Excel and validated through 'GBIF species-lookup tool' and 'GBIF data validator'.

% complete: 100

DOI: <https://doi.org/10.15468/6rq74u>

Expected date of publication:

Title: Taxonomic Checklist, Occurrence & Sequences of the Dung beetles of Singapore

Type: Dataset

Status update: This data in this dataset is in need of validation prior to publishing. We are in touch with taxonomists from Japan and will proceed with publication as soon as possible. Format: Checklist (.core) + Occurrence (.extension) + Molecular (.extension)

Dataset scope: Dung beetles, Scarabaeinae, Singapore, Checklist, Occurrence}

Expected number of records: 4000

Data holder: GBIF, TEE Lab - NTU/ASE, Eleanor Slade, Marx Yim, Ong Xin Rui

Data host institution: Asian School of the Environment

Sampling method:

% complete: 50

DOI:

Expected date of publication: 2023-12-31

Title: Sequences of the Dung beetles (Coleoptera: Scarabaeidae: Scarabaeinae) of Sabah, Malaysia

Type: Dataset

Status update: This dataset is in progress.

Dataset scope: Dung beetles, Scarabaeinae, Sabah, Molecular}

Expected number of records: 50

Data holder: GBIF, TEE Lab - NTU/ASE, Eleanor Slade, Marx Yim, Ong Xin Rui

Data host institution: Asian School of the Environment

Sampling method: We used the NGS sequencing approach.

% complete: 50

DOI:

Expected date of publication: 2023-12-31

Impact of COVID-19 pandemic on project implementation

The project has achieved most of its goals despite covid restrictions. We have made travel to malaysia to have meetings with project partners in UMS and FRC in May and June of 2022 and we made an additional trip to Oxford University Museum of Natural History (OUMNH) in March 2023 in place of Helsinki due to Covid-19. We do not expect this to impact our outputs and deliverables.

Events

GBIF-BIFA Meeting with Collaborator, Lee Kong Chian Natural History Museum

Dates: 2021-11-29 - 2021-11-29

Organizing institution: Lee Kong Chian Natural History Museum - Asian School of the Environment

Country: Singapore

Number of participants: 5

Comments:

Website or sources of verification: <https://www.gbif.org/event/c47137-e62c-42c7-85ec-80c72e8/gbif-bifa-meeting-with-collaborator-lee-kong-chian-natural-history-museum>

Events

GBIF-BIFA Presentation to Tropical Ecology & Entomology Lab, ASE

Dates: 2022-03-17 - 2022-03-17

Organizing institution: Asian School of the Environment, Nanyang Technological University

Country: Singapore

Number of participants: 12

Comments:

Website or sources of verification:

Events

GBIF-BIFA Presentation to Forest Ecology Lab, ASE

Dates: 2022-03-24 - 2022-03-24

Organizing institution: Asian School of the Environment, Nanyang Technological University

Country: Singapore

Number of participants: 20

Comments:

Website or sources of verification: <https://twitter.com/ElleSlade/status/1509124737523363846>

Events

GBIF-BIFA Meeting with Collaborator, Universiti Malaysia Sabah

Dates: 2022-04-08 - 2022-04-08

Organizing institution: Universiti Malaysia Sabah - Asian School of the Environment

Country: Sabah, Malaysia

Number of participants: 3

Comments:

Website or sources of verification: <https://www.gbif.org/event/799de3-c9cb-4e91-93e0-b7b5404/gbif-bifa-meeting-with-collaborator-universiti-malaysia-sabah>

Events

GBIF-BIFA meeting with Institute of Tropical Biology and Conservation - Borneensis Collection

Dates: 2022-05-23 - 2022-05-27

Organizing institution: Institute of Tropical Biology and Conservation - Borneensis Collection

Country: Sabah, Malaysia

Number of participants: 20

Comments:

Website or sources of verification: <https://www.gbif.org/event/015238-69a8-41e5-8a20-15618e1/gbif-bifa-meeting-with-institute-of-tropical-biology-and-conservation-borneensis-collection>

Events

GBIF-BIFA meeting with Forest Research Centre Sepilok, Sabah

Dates: 2022-06-01 - 2022-06-03

Organizing institution: Forest Research Centre Sepilok

Country: Sandakan, Sabah, Malaysia

Number of participants: 8

Comments:

Website or sources of verification: <https://www.gbif.org/event/015238-69a8-41e5-8a20-15618e1/gbif-bifa-meeting-with-institute-of-tropical-biology-and-conservation-borneensis-collection>

Events

GBIF-BIFA Meeting with Collaborator, Oxford University Museum of Natural History

Dates: 2023-03-13 - 2023-03-17

Organizing institution: Asian School of the Environment

Country: United Kingdom

Number of participants: 10

Comments:

Website or sources of verification: <https://teelabntu.wixsite.com/home/post/examining-the-dung->

Communications and visibility

Since the start of the project, we have communicated our progress through posts on Twitter, our Tropical Ecology and Entomology Lab website and have been working to create a Sabah dung beetle guidebook.

Monitoring and evaluation

Final Evaluation

The project implementation went smoothly, even despite Covid impacts. We managed to engage well with all project partners, and visited most partners and exchanged knowledge and collated datasets. Having a dedicated Project Officer, Marx Yim, facilitated the interactions and was crucial to the success of the project and producing the datasets and deliverables. The PHd student Ong Xin Rui has also contributed to the project and will use the datasets to explore species distributions and responses to land use change as a chapter in her PhD. She will directly use the published Sabah datasets from this project. This project has started to fill a large data and knowledge gap for the dung beetles of the region, and has helped forge stronger collaborations between those working on them.

Best Practices and Lessons Learned

The support from Lily and GBIF was very useful and was essential to us in making sure the datasets were published correctly. It was also crucial, especially in the beginning, for us to be able to navigate the system. However, it did at times feel a little micro-managed! It was not clear to use exactly how the datasets should be structured in the beginning, and we ended up splinting some datasets. The way the datasets are divided into .core and .ext can be confusing and the fact that sampling events cannot be an extension of occurrences does not make the most sense for our data, where the sampling events are nested within occurrence datasets, not the other way around. The system at times is not very user-friendly and I have concerns that the way we are forced to classify and publish data could lead to end-user errors (e.g. the occurrence dataset is not a true occurrence dataset as it only contains the subset of occurrences not in the sampling event data).

Post Project Activity(ies)

We will continue to work together with project partners to complete the project deliverables (sequencing datasets, Singapore datasets, and the guidebooks) post-project as outlined above.

Sustainability plans

The project has long-term impacts in terms of filling a large data gap on the dung beetles of Singapore and Malaysia and the database will hopefully act as the starting point for a larger database for dung beetles of the whole SE Asian region. The project will contribute to long-term capacity building in the region both institutionally and with the wider community. The impact of this will be measured by the uptake of the guides, and the number of times datasets, databases and papers are accessed, downloaded and cited.

GBIF leads the Biodiversity Information Fund for Asia (BIFA), a programme funded by the Ministry of the Environment, Government of Japan. The programme provides supplementary support for activities addressing the needs of regional researchers and policymakers through mobilization and use of biodiversity data.

