

Bee data mobilization from TIGER project

Programme:BIFA Project ID: BIFA6_009 Project lead organization:Chulalongkorn University, Department of Biology Project implementation period:1/11/2021 - 31/10/2022 Report approved: 15/12/2022

Narrative Final report

Executive Summary

The purpose of this project is to process, sort and identify, providing georeferencing data, highresolution images, and mobilize at least 3,000 processed occurrence records of bees specimens deposited at the QSBG, Thailand under the TIGER project to GBIF. The objectives were met with 3,393 verified records of bees from 5 families: Andrenidae (3), Apidae (1,359), Colletidae (15), Halictidae (1,903), and Megachilidae (113). Bees are identified to 33 described genera and 74 species despite 119+ morphospecies identified. More than 8,200 images from 2,029 specimens were captured. 3,156 out of 3,393 occurrence records are georeferenced.

Overall, the project implemented is a success despite many delays during to the pandemic. Important lesson learned from this project including (1) establishing a realistic goal (2) appropriate task allocation and synergy among staff members (3) frequent contacts with GBIF personnel is highly encouraged (4) scheduled meetings with GBIF management team will clarify many questions relating to the project. Post-project activities that will insure the project sustainability including (1) Continuing to update this published TIGER bee database in GBIF, since there are possibly more than 2,000 bee specimens left at QSBG that needed to be sorted, identified, and digitized (2) This TIGER bee dataset and our previous data from BIFA5_005 (DOI10.15468/tf4ejd) will provide backbone information for future work in DNA barcoding of Thai bees associated with data in iBOL and GBIF.

Progress against milestones

Has your project completed all planned activities?: Yes

Has your project produced all deliverables?: Yes

Report on Activities

Activity implementation summary

Traveling to QSBG and sorting specimens

We traveled to QSBG in Chiang Mai twice to obtain sufficient number of specimens for the project. First, on the 28th November to 2nd December 2021 and second on the 27th to the 31th of March 2022.

Specimens processing

More than 5,000 hymenopteran specimens retrieved from QSBG were pinned, dried, and attached with labels

Digitization I: Identification of the specimens, validating collecting data, and process these info into database

More than 3,393 occurrence records were sorted, identified and validated for TIGER bee specimens. This is 300+ more records beyond our initial pledge of 3,000 specimens.

Digitization II: Photographing of specimens and associated labels

Initially, we proposed to captured 3,000 bee specimens (4 images for 1 specimen). Currently we captured 8,200+ images from 2,029 specimens and will continue until we reach 3,393 specimens with 13,500+ images. The number of images that can be viewed in GBIF at the present is 8,200+ images from 2,029 specimens that were already uploaded.

Digitization III: Validating georeferencing data and mapping species distributions.

3,156 out of 3,393 occurrence records are georeferenced. This is 100+ more georeferenced records beyond the initial pledge of 3,000.

Digitization IV: Attaching QR codes to specimens and link to database

QR numbers were generated for each occurrence record. These codes will be uploaded later in the dataset. Attaching actual QR information to the specimens will be carried out after the submission of the final report. This delay initially stemmed from the pandemic situation and the lack of funding during the early days of the project, which a suitable equipment for printing the actual QR codes cannot be obtained on time. However, the QR numbers are already generated and the final step of attaching the QR codes to specimens should be completed in less than a month.

Deliveries of data: Transcription of data to GBIF platform and upload processed records

Since the beginning stage of the project, we uploaded our occurrence data twice in major batches. First attempt was on the 30th April 2022 for 1,687 occurrence records and the second attempt on the 31st October 2022 for a cumulative of 3,393 records. This is 400+ more records beyond our initial pledge of 3,000.

The dataset contains occurrence records (3,393) of bees (Anthophila) from 5 families: Andrenidae (3), Apidae (1,359), Colletidae (15), Halictidae (1,903), and Megachilidae (113). 3,393 records have been identified to a total of 33 described genera and 74 species: Andrenidae 1 genus 1 species, Apidae 17 genera and 40 species, Colletidae 2 genera and 2 species, Halictidae 9 genera and 24 species, and Megachilidae 6 genera and 7 species. However, the total number of morphospecies identified is 119+ suggesting more describing of new species and revision are essential for the future.

Submission of mid-term report

The mid-term report was submitted on the 30th April 2022 and was approved later by the GBIF secretariat.

Workshop

Workshop to showcase the Thai bee TIGER database to students and researchers interested in pollination ecology, conservation, bee taxonomy, and biodiversity informatics was initially schedule for November 2022; however, this intended workshop has a time conflict with the GBIF Asia Regional meeting on the 22nd-24th of November 2022. Therefore, we are planning to hold the actual or virtual workshop later on 2022 or early 2023.

BIFA Capacity enhancement virtual Workshop

Our technician, Mr. Nontawat Chatthanabun, attended the mandatory workshop on the 25th October-15th December 2022 and completed the training with the advance badge and certificate (see attached document).

Completed activities

Activity name: Traveling to the QSBG and sorting specimens

Description: Travel to the QSBG in Chiang Mai province to sort out bee specimens from remaining TIGER collection and bring specimens back to the lab in Bangkok **Start Date - End Date:** 1/11/2021 - 30/11/2021 **Verification Sources:** DOI10.15468/p6zfzc

Activity name: Specimens processing

Description: Bee specimens were pinned, dried, and labels attached. **Start Date - End Date:** 1/12/2021 - 31/12/2021 **Verification Sources:** DOI10.15468/p6zfzc

Activity name: Digitization I

Description: Identification of specimens, validating collecting data, and process info into the database **Start Date - End Date:** 1/1/2022 - 31/10/2022 **Verification Sources:** DOI10.15468/p6zfzc

Activity name: Digitization II

Description: Photographing bee specimens and associating labels **Start Date - End Date:** 1/1/2022 - 31/10/2022 **Verification Sources:** DOI10.15468/p6zfzc

Activity name: Digitization III

Description: Validating georeferencing data and mapping species distribution. **Start Date - End Date:** 1/1/2022 - 31/10/2022 **Verification Sources:** DOI10.15468/p6zfzc

Activity name: Digitization IV

Description: Attaching QR codes to specimens and link to database **Start Date - End Date:** 1/1/2022 - 31/10/2022 **Verification Sources:** DOI10.15468/p6zfzc

Activity name: Deliveries of data

Description: Transcription of data to GBIF platform and upload processed records **Start Date - End Date:** 1/5/2022 - 31/10/2022 **Verification Sources:** DOI10.15468/p6zfzc

Activity name: BIFA Capacity enhancement virtual workshop

Description: Mandatory workshop Start Date - End Date: 25/10/2021 - 15/12/2021 Verification Sources: in Report Attachments

Activity name: Submission of mid-term report

Description: Submission of mid-term report **Start Date - End Date:** 30/4/2022 - 30/4/2022 **Verification Sources:** See narrative mid-term report

Report on Deliverables

Production of Deliverables - Summary

The expected deliverable product of this project is a fully digitized occurrence dataset of 3,000 bee specimens in Thailand covering bees from families Andrenidae, Apidae, Collectidae, Halictidae, and Megachilidae. Fortunately, we managed to process 3,393 verified bee occurrence records and uploaded the dataset in GBIF under DOI10.15468/p6zfzc. However, images were captured from 2,029 specimens with each specimens comprising 4 images each (8,200+ images in total). After the submission of final report, we will continue producing images for the rest of 1,364 bee specimens to reach a total images of 13,500+ pictures. The cause of delay for image capturing in the later half of the project is due primarily to the pandemic situation and cases that surged in Thailand which resulted in major lockdown at university where the project has been carried out.

Production of deliverables

Title: Thai Bees from TIGER Project

Type: Dataset

Status update: 3,393 occurrence records with verified geo-referencing coordinates are completed with 8,200+ images from 2,029 bee specimens.
Dataset scope: Bees in families Andrenidae, Apidae, Colletidae, Halictidae, and Megachilidae from TIGER project in Thailand from 2006-2008}
Expected number of records: 3393
Data holder: Chulalongkorn University

Data host institution: Chulalongkorn University Sampling method: Malaise and pan traps % complete: 90 DOI: DOI10.15468/p6zfzc Expected date of publication:

Impact of COVID-19 pandemic on project implementation

The project has been conducted consistently from the beginning in November 2021; however, intermittent university shut-downs from COVID-19 situation has delayed our works briefly. The shut-down of the university prohibited us from gaining access to collection, photographic equipment, and computer including server for database storage. Nevertheless, we were able to compensate the time lost and keep up with our work to come up with 3,393 occurrence records of bees specimens in TIGER project, though the images captures came up short for 900+ specimens.

In conclusion, despite the obstacles we faced during the end of this pandemic, we are delighted that we achieve many milestones in our project to meet this deadline.

Communications and visibility

Occurrence record database of bees in TIGER project will be linked to the QSBG website (underdevelopment) to display our data to local Thai audiences. Taxonomically important records of certain species resulted from this work will be published in scientific journals to complement the database in the future. The processes and methods of digitization of bee occurrence records will be disseminated and shared with other research collections, universities, and institutions in Thailand through communications to the QSBG website and other social media.

In the end of 2022 or early 2023, a workshop on TIGER bee database is planned at Chulalongkorn University, Bangkok, which will bring together students and researchers interested in pollination ecology, conservation, and bee taxonomy to familiarize and learn how to efficiently utilize the information provided in the database, and also to advertise information regarding the importance of pollinators to the public. For further communication and dissemination of the project, we are more than happy to liaise with the GBIF Secretariat to help advertising of biodiversity data sharing to other Thai research institutions.

Monitoring and evaluation

Final Evaluation

Overall, the project implemented is a success with 3,393 occurrence records for bee specimens collected in Thailand from 2006-2008 published in GBIF with more than 8,200 high-resolution images. We are enthusiastic about this bee dataset to be available to the public and the scientific community, since this data represents the first collection of bee specimens collected from national parks and restricted areas in Thailand. The number of specimen records that we initially proposed (3,000) is a practical number of record that can be achieved in a year worth of effort. Despite the delay due the global pandemic situation, we are fortunate and consistent in executing the project to reach this current status.

The strength of this project implementation is the synergistic efforts among everyone involved in the database built up. Task allocation is very effective by assigning the right person to the right job. The recruitment of technicians working in the project with extensive background in biodiversity, photographing and computer is critical to the success of the work. In addition, the staff meetings were held every week. All of the staff worked in the same room and most of the time at the same time. This ambience created opportunity to address questions and discussion among technicians working on different aspects of the project.

Weaknesses observed throughout the project implementation period include factors involving shortage of computer and photographing equipment to process high quality images fast enough to meet the deadline. This internal problem directly delayed our progress in finishing images capturing of 3,933 specimens. Nevertheless, we have only 971 specimens left to capture to meet the pledge numbers and we will continue image capturing of all our records to 3,393 specimens.

The GBIF Secretariat and the management team have been very helpful throughout the implementation of the project. Email responding were answered in short time with practical info to help us tackled many technical issues. In addition, the virtual meetings are informative and essential to our

understanding of GBIF goals. All personnel working for GBIF are empathetic with good will. We would like to single out acknowledging Mr. Chihjen Ko for his proactive approach in communication with us through out the time working in this project. Mr. Ko has been very resourceful and promptly replying to all of our questions.

The only issue that we would like to convey to the GBIF management team is similar to last year comments in project BIFA5_005, which revolves around the mode of funding transfer. The bureaucratic process in Thai university hindered the efficiency for researchers to obtain monetary supports in time to meet many of their objective goals. This situation caused delays on the project because many necessary equipment can not be procured and payments to students helping the project were also affected. We are hoping that in the future, the BIFA grant will consider a direct transfer of funding to the project manager instead.

Best Practices and Lessons Learned

Lesson learned from this project can be described as follows:

(1) To accomplish a project relating to data mobilization and database construction, a realistic goal, i.e., number of records needed to be carefully and diligently plan. We were fortunate to complete our pledge in term of number of record because we did not overcommit in the beginning.

(2) Task allocation is one of the major components in the success of a group project. In addition, consistent and swift communication among staff can reduce number of difficulties that may have occurred. We are grateful with the working atmosphere that we have on this project. Every staff is dedicate to his/her specialized task and the work ethic was held.

(3) A consistent communication with GBIF management team whether scheduled or casual contacted provide many opportunities to discuss the progress, problem, and solution to project development throughout the year.

(4) Mandatory and voluntary meetings scheduled by the GBIF management team are highly essential even if part of the context is redundant sometimes. The opportunity to meet with GBIF team is invaluable and very informative. We would suggest everyone that receive BIFA grant in the future to not skip a single meeting.

Post Project Activity(ies)

Post project activities beside of finishing image capturing of the rest 971 bee specimens to update the database and attaching QR codes to every specimens in the collection are:

(1) Continuing to update this published TIGER bee database in GBIF, since there are possibly more than 2,000 bee specimens left at QSBG that needed to be sorted, identified, and digitized. Next year, we hope to complete the digitization of every bee specimens at the QSBG (wish us luck :)

(2) This TIGER bee dataset and our previous data from BIFA5_005 (DOI10.15468/tf4ejd) will provide backbone information for future work in DNA barcoding of Thai bees associated with data in iBOL (International Barcoding of Life) and GBIF. We are planning to recruit funding of this project using the dataset we provided to GBIF.

Sustainability plans

The digitization of bee occurrence records will be disseminated and shared with other research collections, universities, and institutions in Thailand through communications to the QSBG website via the Thai Bee Database (under development in http://61.91.64.2:8075) and other social media channels.

As mentioned earlier in Post Project Activities, we are expecting to continue sorting, identification, and digitization of the remaining bee specimens at the QSBG. In the future, we are planning to expand our collection beyond the number we earlier committed. As time progress, the number of specimens in the database will be updated and should GBIF permit, we will continue adding more occurrence records into the dataset to keep the dataset active and up to date with new information.

In addition, our current Thai bee dataset will provide a backbone information for a future work in DNA barcoding of Thai bees. We are planning to establish a barcoding project using the dataset we procured for GBIF to enhance the capability and performance of the database. The project is aimed to collaborate with iBOL to ascertain the quality of the data generated and broad public access.

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.

