

# DNA barcoding Asian Pteridophytes (1): 1.5KP DNA-referenced Collections from Taiwan and Southeast Asian Countries

---

**Programme:**BIFA

**Project ID:** BIFA6\_010

**Project lead organization:**Institute of Molecular & Cellular Biology, National Tsing Hua University

**Project implementation period:**1/9/2021 - 28/2/2023

**Report approved:** 29/6/2022

## Narrative Midterm report

---

### Executive Summary

---

We already finished (1) primer design (2) candidate region test (4) selecting samples (5) DNA extraction (6) NGS workflow design (7) sequencing 1/2 of samples (8) uploading 1/3 geo-referenced records to GBIF (9) submission of three manuscripts regarding to methodology and findings of new species diversity. At this stage, we mostly focused on molecular DNA work and computation in the lab, and regular lab meeting had used to monitored these progresses. In addition, we published partial of our dataset and methodology in two journal articles.

### Progress against milestones

---

**Has your project published at least one dataset through GBIF.org?: Yes**

**Dataset published:**

Dataset	DOI
Improving knowledge of Asian pteridophytes through DNA sampling of specimens in regional collections	10.15468/nhsnn2
two_new_asian_lomariopsis	10.15468/cqucep

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

**Name of the workshop participant:**Chen-Wen Chen

**Certification obtained:** Advanced Badge

### Report on Activities

---

#### Activity progress summary

A. Complete:

- (1) Selecting 1.5KP Asian collections
- (2) Revaluating previously proposed DNA barcodes
- (3) Designing NGS protocol for DNA barcoding

B. New activities:

- (1) Using DNA barcode as genetic evidence for undescribed fern species:

Using our preliminary data of DNA barcodes, we have identified several undescribed fern species in Asian regions, so far including two *Lomariopsis* species and one *Hymenophyllum* species. We also prepared two manuscripts to describe these new species. Both were submitted to the journal PhytoKeys, one is already published and another is still under review. The two new *Lomariopsis* species can be seen in "Wu, Y.-H., C.-Y. Sun, A. Ebihara, N. T. Lu, G. Rouhan, and L.-Y. Kuo. 2021. Two new species in the fern genus *Lomariopsis* (Lomariopsidaceae) from East Asia. PhytoKeys 187: 161–176."

## (2) New universal primer sets for fern DNA barcodes:

We wrote scripts and found conserve priming sites across 250 fern plastomes. For the *trnL-F* region, we already tested the newly designed primer set and published these primer sequences in Wu et al. (2022; "Integrating tissue-direct PCR into genetic identification: an upgraded molecular ecology way to survey fern field gametophytes. Applications in Plant Sciences: e11462.")

## C. Ongoing:

### (1) PCR and NGS sequencing:

We already finished PCR and sequencing ~74% and ~50% of samples. The delay of this part was due to the slowdown procedure of DNA extraction because lab works were affected by the recent COVID-19 situation in Taiwan. Nonetheless, we slightly re-designed our lab workflow, and this part can be finalized soon during June 2022.

### (2) Compiling 1.5 KP dataset & data mobilization:

We already submitted geo-referenced data of ~1/3 of samples to GBIF, but their sequences are still not yet submitted to GenBank. Besides, herbarium works for digitization of their voucher collections were also delayed due to the COVID-19 situation. For these remaining works, we also hope to adjust the workflow. Herbarium staffs in TAIF may help in collecting the geo-referenced dataset in the near future. In our new schedule, this part will be completed during this September.

### (3) Within country workshop to broadcast the workflow & DNA barcode information:

This workshop was planned to be held at May 28 but postponed due to the COVID-19 situation. We are now contacting the co-organizers to plan a new date for this workshop. Hopefully, we can complete this part later this year.

All above delays little affect the following progress.

## Completed activities

### Activity name: Selecting 1.5KP Asian collections

---

**Description:** Sampling and DNA extraction is already finished. PCR and sequencing are on going.

**Start Date - End Date:** 1/9/2021 - 28/2/2022

**Verification Sources:**

<https://docs.google.com/spreadsheets/d/1rxMHcscDVWqXHNp2EWF7u5sWtHI9bO5k/edit?usp=sharing&ouid=107650169824774415794&rtpof=true&sd=true>

### Activity name: Revaluating DNA barcode regions

---

**Description:** Revaluation amplification success of different proposed regions. We finally chose *trnL-F* and *rbcL* for downstream works because of high implication success.

**Start Date - End Date:** 1/10/2021 - 1/11/2021

**Verification Sources:** <https://drive.google.com/file/d/1LvJk1Qn8p202y-si86NW3PbKjQQ3X5pJ/view?usp=sharing>

### Activity name: Designing NGS protocol for DNA barcoding

---

**Description:** Designed and tested indexed-PCR primer sets. Designed pipeline for illumina and PacBio amplicon sequencing.

**Start Date - End Date:** 1/11/2021 - 1/12/2021

**Verification Sources:** <https://drive.google.com/file/d/1I9R5zpAY-SADfgAHTiQVUyyDSIWqvV8b/view?usp=sharing>

## Report on Deliverables

---

## Deliverables progress summary

We already uploaded two occurrence datasets to GBIF (DOI: 10.15468/nhsnn2 and DOI: 10.15468/cqucep). These cover ~1/3 of our planned collections.

### A. New deliverables:

(1). Two accepted journal articles:

(1.1) "Two new species in the fern genus *Lomariopsis* (Lomariopsidaceae) from East Asia":  
Using our preliminary data of DNA barcodes, we identify two undescribed *Lomariopsis* fern species in Asian regions, and described these new species in this paper. DOI: 10.3897/phytokeys.187.77035

(1.2) "Integrating tissue-direct PCR into genetic identification: an upgraded molecular ecology way to survey fern field gametophytes":

We designed new universal primers of one of DNA barcodes candidates, plastid trnL-F region, based on fern plastome sequences. we published these primer sequences in this paper. DOI: 10.1002/aps3.11462.

(2) Pipeline for demultiplexing NGS data:

The scripts of our pipeline are provided on github platform, <https://github.com/PowerBarcoder>

### B. Delayed deliverables:

The delays of the full dataset of 1500 pteridophytes contain (1) molecular works in the lab, (2) NGS sequencing, (3) herbarium works, and are mainly due to the COVID-19 situation. The first and second parts will be completed soon during the coming June. The third part is planned to finished during this October.

## Progress towards deliverables

### Title: Improving knowledge of Asian pteridophytes through DNA sampling of specimens in regional collections

---

**Type:** Dataset

**Status update:** The ferns, lycophytes and seed-free vascular plants commonly described as pteridophytes exhibit hyperdiversity in the insular vegetation that often characterizes Asian floras. Despite harboring biodiversity hotspots, these plants and their georegions have been poorly surveyed, particularly in Southeast Asia, where one third of the world's pteridophyte species are concentrated. More than 60 per cent of the approximately 4,500 species lack georeferenced records in GBIF and only 6 per cent have been DNA barcoded. This project aims to increase the available knowledge on Asian pteridophytes by compiling a georeferenced occurrence dataset that includes images, DNA barcodes and other vouchering information from thousands of recent collections, building on the efforts of the Taiwan Pteridophyte Research Group and its Southeast Asian collaborators. The project team will set up a workflow incorporating next-generation sequencing for 1,500 Asian pteridophyte specimens from selected collections in Taiwan, Vietnam, the Philippines, Malaysia and other Southeast Asian countries that can fill in taxonomic and geographic gaps and represent Asian pteridophyte diversity. Mobilization of and access to these vouchered and georeferenced DNA-derived records will advance further research into the biogeography of pteridophytes and other terrestrial vegetation and support the development of novel approaches to monitor biodiversity along the spatiotemporal scale, including metabarcoding of the invisible diversity held in soil and spore banks.

**Dataset scope:** Certain areas among South East Asia.

**Expected number of records:** 556

**Data holder:** Chien-Wen Chen

**Data host institution:** Taiwan Forestry Research Institute

**Sampling method:** TBC

**% complete:** 37

**DOI:** 10.15468/nhsnn2

**Expected date of publication:**

### Title: two\_new\_asian\_lomariopsis

---

**Type:** Dataset

**Status update:** Two East Asian *Lomariopsis* (Lomariopsidaceae, Polypodiales) species, *Lomariopsis moorei* and *Lomariopsis longini*, which were previously misidentified as *L. spectabilis*, are here described as new species based on evidence from morphological characters and a molecular phylogeny. The two species differ from the three other described species in East Asia by their venation, pinna shapes, and perine morphology. A phylogeny based on a combined dataset of three chloroplast regions (rbcL+ rps4-trnS + trnL-L-F) showed that *L. moorei* and *L. longini* each formed a well-supported monophyletic group which was distantly related to both *L. spectabilis* and the other morphologically similar East Asian species, *L. boninensis*.

**Dataset scope:** East Asian *Lomariopsis*

**Expected number of records:** 16

**Data holder:** Li-Yaung Kuo

**Data host institution:** Taiwan Biodiversity Information Facility

**Sampling method:** Morphological and phylogenetic analyses, and re-identification of previous collections in herbaria.

**% complete:** 100

**DOI:** 10.15468/cqucep

**Expected date of publication:**

---

### **Title: New East Asian Lomariopsis species**

---

**Type:** Other

**Description:** Accepted journal article for two new East Asian Lomariopsis species.

**Sources of verification:** doi: 10.3897/phytokeys.187.77035

---

### **Title: Newly designed universal primer set for fern DNA barcode**

---

**Type:** Other

**Description:** A journal article with publishing a newly designed primer set for fern DNA barcode region, trnL-F.

**Sources of verification:** doi: 10.1002/aps3.11462

---

### **Title: Pipeline for demultiplexing NGS data**

---

**Type:** Other

**Description:** Collections of scripts for our NGS pipeline "PowerBarcoder"

**Sources of verification:** <https://github.com/PowerBarcoder>

---

## **Communications and visibility**

---

We already made our datasets and methodology partially public in different platforms and journal articles. We hope to broadcast to research community soon via workshops, data papers, and an additional methodology paper.

---

## **Monitoring and evaluation**

---

### **Monitoring and evaluation findings**

Sample selection by different partners has been completed in a good progress. Communication of lab work procedure was also great, in which we sometimes used online meetings and emails to alleviate effects due to the COVID-19 circumference. However, regarding slowdown of lab and herbarium works, we had rearranged our schedule, and published a portion of our dataset and methodology in journal articles or other platforms first. In addition, we timely troubleshoot mistakes first among different temperate outputs, and thus we can highly reduce time needing to fill in missing gaps at the final stage.

---

## **Impact of COVID-19 pandemic on project implementation**

---

Because of the COVID-19 pandemic, our lab and herbarium works were delayed. Nonetheless, we had first prepared datasets and the corresponding manuscripts using some of our preliminary results, which are included in two published journal articles and one manuscript under review. We also made our methodology partially public, and these shall enhance the demonstration of our DNA barcoding workflow.

---

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.



