Plan Overview

A Data Management Plan created using DMPTool

Title: Micromorphological Analysis of Guadua

Creator: Elizabeth McMurchie

Affiliation: Iowa State University

Template: C-CHANGE Grass2Gas CAP

Project abstract:

The epidermal micromorphology of foliage leaves is a valuable tool for assessing taxonomic placement of Neotropical woody bamboos. In bamboos with similar macromorphological features, foliage leaf micromorphology can help distinguish different species and reveal potential evolutionary or ecological relationships. Characters such as the shape of silica bodies and the presence, abundance, placement, and shape of stomata and papillae on epidermal cells on different surfaces of the leaf are sometimes diagnostic of certain clades but can show dramatic variation even at the species level. Among the Neotropical woody bamboos, members of the subtribe Guaduinae, including the genus *Guadua*, are unusual in that most species have numerous stomata on the adaxial surface of the foliage leaf, in addition to the abaxial surface. Most *Guadua* have also been noted for having papillae associated with these adaxial stomata, although in some species papillae may also be found on long cells of the abaxial surface of foliage leaves. The presence of papillae on both surfaces of foliage leaves has historically been noted in *G. paniculata,* a widespread species that, unlike most Neotropical woody bamboos, most frequently occurs in open savanna habitats rather than forests or forest margins. *Guadua paniculata* and other species of *Guadua* found in open habitats share a suite of macromorphological characteristics, including shorter, narrower culms, culm leaves with subequal sheath and blade, relatively narrow foliage leaves, and shorter pseudospikelets compared to members of the genus found in more forested areas. However, whether patterns of variation in foliar micromorphology in *Guadua* align with the macromorphological variation based on habitat is not currently known. In this study, we sought to determine whether habitat and geographic location were significant factors in patterns of variation in *Guadua* foliage leaf epidermal micromorphology. Through a review of scanning electron micrograph (SEM) images of foliage leaves of different species of *Guadua*, we constructed a multivariate, binary dataset to record the presence or absence of micromorphological character states in individual specimens. To compare the similarity of the specimens, we converted the presence and absence data to distances to visualize the dataspace. Using the simple matching distances, we evaluated a linear model using permutation methods to determine potential causes of variation in foliar micromorphology. Preliminary results indicate that micromorphological patterns differ between species found in open habitats and those found in forested habitats.

ID: 69776

Start date: 09-15-2021

End date: 12-18-2021

Last modified: 09-19-2021

Micromorphological Analysis of Guadua

Documentation

Describe how your data is being documented. Provide details on metadata, data dictionaries, codebooks and README files.

Metadata:

- Species name, collector name, career number, and location of collection are recorded for each herbarium specimen used for scanning electron microscope (SEM) imaging
- SEM images were taken by Lynn G. Clark or Elizabeth McMurchie and will be labeled according to photographer
    - Autogenerated metadata files are associated with SEM images taken recently by Elizabeth McMurchie
    - Older images taken by Lynn G. Clark are scanned from Polaroid photographs and are accompanied by a single metadata file.

Data dictionary:
The following variables are to be used:

- Specimen: Specimen collector and number, with species name included
- Country: Country in which the specimen was collected
- Region: Region in which specimen was collected: Mexico, Central America, Andes, or Eastern South America
- Species: Species name
- Habit: Habit typical of species: Small arching, large and erect, or leaning and climbing
- Habitat: Habitat typical of species: Forest, river, or savanna
- General_habitat: Habitat typical of species, with river and forest categories combined to include savanna or general forest habitats.
- Ab_papillae_long_cells_stomatal: Abaxial: papillae on the long cells in the stomatal zone 0 = absent 1 = present
- Ab_papillae_long_cells_interstomatal: Abaxial: papillae on the long cells in the interstomatal zone: 0 = absent; 1 = present
- Ab_pickle: Abaxial: Prickle hairs: 0 = absent; 1 = present
- Ab_bimicro: Abaxial: Bicellular microhairs: 0 = absent; 1 = present
- Ab_macro: Abaxial: Macrohairs: 0 = absent; 1 = present
- Ab_bilobate_intercostal: Abaxial: Bilobate silica bodies in the intercostal zone: 0 = absent; 1 = present
- Ab_ridged_saddle_intercostal: Abaxial: Ridged saddle-shaped silica bodies in the intercostal zone: 0 = absent; 1 = present
- Ab_smooth_saddle_intercostal: Abaxial: Smooth saddle-shaped silica bodies in the intercostal zone: 0 = absent; 1 = present
- Ab_vertically_elongated_intercostal: Abaxial: Vertically elongated silica bodies in the intercostal zone: 0 = absent; 1 = present
- Ab_bilobate_costal: Abaxial: Bilobate silica bodies in the costal zone: 0 = absent; 1 = present
- Ab_ridged_saddle_costal: Abaxial: Ridged saddle-shaped silica bodies in the costal zone: 0 = absent; 1 = present
- Ab_smooth_saddle_costal: Abaxial: Smooth saddle-shaped silica bodies in the costal zone: 0 = absent; 1 = present
- Ab_vertically_elongated_costal: Abaxial: Vertically elongated silica bodies in the costal zone: 0 = absent; 1 = present
- Ab_triangular_sub_cells: Abaxial: Triangular stomata subsidiary cells: 0 = absent; 1 = present
- Ab_dome_sub_cells: Abaxial: Dome-shaped stomata subsidiary cells: 0 = absent; 1 = present
- Ab_parallel_sub_cells: Abaxial: Parallel-sided stomata subsidiary cells: 0 = absent; 1 = present
- Ad_stomata: Adaxial: stomata on the adaxial surface of foliage leaf blades: 0 = absent; 1 = present.
- Ad_stomata_freq: Adaxial: Frequency of stomata if present on the adaxial surface of foliage leaf blades: 0 = common; 1 = infrequent.
- Ad_papillae_long_cells_stomatal: Adaxial papillae on the long cells in the stomatal zone exclusive of bulliform cells 0 = absent 1 = present
- Ad_papillae_long_cells_interstomatal: Adaxial papillae on the long cells in the interstomatal zone: 0 = absent; 1 = present.
- Ad_papillae_overarch: Adaxial: Papillae on long cells of the intercostal zone adjacent to the stomates: 0 = not overarching the stomates; 1 = overarching the stomates.
- Ad_papillae_bulliform: Adaxial papillae on the bulliform cells: 0 = absent; 1 = present.
- Ad_pickle: Adaxial: Prickle hairs: 0 = absent; 1 = present
- Ad_bimicro: Adaxial: Bicellular microhairs: 0 = absent; 1 = present
- Ad_macro: Adaxial: Macrohairs: 0 = absent; 1 = present
- Ad_bilobate_intercostal: Adaxial: Bilobate silica bodies in the intercostal zone: 0 = absent; 1 = present
- Ad_ridged_saddle_intercostal: Adaxial: Ridged saddle-shaped silica bodies in the intercostal zone: 0 = absent; 1 = present
- Ad_smooth_saddle_intercostal: Adaxial: Smooth saddle-shaped silica bodies in the intercostal zone: 0 = absent; 1 = present
- Ad_vertically_elongated_intercostal: Adaxial: Vertically elongated silica bodies in the intercostal zone: 0 = absent; 1 = present

- Ad_bilobate_costal: Adaxial: Bilobate silica bodies in the costal zone: 0 = absent; 1 = present
- Ad_ridged_saddle_costal: Adaxial: Ridged saddle-shaped silica bodies in the costal zone: 0 = absent; 1 = present
- Ab_smooth_saddle_costal: Adaxial: Smooth saddle-shaped silica bodies in the costal zone: 0 = absent; 1 = present
- Ad_vertically_elongated_costal: Adaxial: Vertically elongated silica bodies in the costal zone: 0 = absent; 1 = present
- Ad_triangular_sub_cells: Adaxial: Triangular stomata subsidiary cells: 0 = absent; 1 = present
- Ad_dome_sub_cells: Adaxial: Dome-shaped stomata subsidiary cells: 0 = absent; 1 = present
- Ad_parallel_sub_cells: Adaxial: Parallel-sided stomata subsidiary cells: 0 = absent; 1 = present

Methods

- SEM imaging followed the procedures of Clark (1990) with modifications for recently-taken images following McMurchie et al. (in prep.)
- Images for each specimen will be analyzed and the presence or absence of the preceding thirty-six characters recorded, resulting in a binary dataset
- Will obtain distance matrix from binary dataset using simple matching coefficient
- Visualization and analysis will be conducted in R
- Data will be visualized with Principal Coordinates Analysis (PCoA)
- Data will be analyzed with Factorial MANOVA (multivariate analysis of variance) via RRPP (randomized residuals in a permutation procedure)
    - Comparisons will be made between habitat, habit, and location of collection

Codebook

On completion, a codebook will be provided containing relevant R code and packages.

README files

README files will accompany raw data and R script to describe variables and supplementary information for R scripts. A README file will also accompany SEM images for metadata purposes.

References:

McMurchie, E. B. J. Peterson, T. D. Leandro, X. Londoño, and L. G. Clark. A revision of *Chusquea* sect. *Serpentes* (Bambuseae, Bambusoideae, Poaceae) including two new species from South America. In prep.

Clark, L. G. 1990. *Chusquea* sect. *Longiprophyllae* (Poaceae: Bambusoideae): A new Andean section and new species. *Systematic Botany* 15: 617-634.

Data Collection

Describe the data you will collect or generate.
Provide the details on file formats, total storage space required, and data sources.

- Scanning electron microscope (SEM) images were taken of foliage leaf specimens of *Guadua*, and the images stored in TIFF format. All images fit on a 64 gb thumb drive.
    - Autogenerated metadata files accompany SEM images taken recently by Elizabeth McMurchie.
- Presence/absence of different micromorphological characters was recorded in a 24 kb Microsoft Excel File
- Analysis will take place in R, with R scripts estimated to be about 6 kb

Describe how the data will be collected or generated.

Examples: field datasheets, electronic data collection using sensors and dataloggers, unmanned aerial vehicles, data extraction from Google Earth or other systems, gSSURGO, copy records from the stakeholder record system, voice recording, video recording, pictures, surveys.

Data will be collected using visual analysis of SEM images (in TIFF format) of micromorphology of *Guadua* leaves. Presence/absence of various features will be recorded in a Microsoft Excel file.

Ethics and Legal Compliance

Describe expected ethical and legal issues including details about Institutional Review Board (IRB) approval and anonymization requirements.

There are no expected ethical or legal issues involved in this study.

Describe how you will manage copyright and Intellectual Property Rights (IP/IPR).

All photos are copyrighted to Elizabeth McMurchie or Lynn G. Clark. Spreadsheets and R scripts are copyrighted to Elizabeth McMurchie.

Storage and Backup

Describe where and how the data will be stored and backed up. Clearly indicate the primary storage location and the backup storage locations which should include at least 3 digital locations and at least 2 separate physical locations. State when the data will be backed up and by whom.

Physical locations

- Images, spreadsheets, and R scripts will be stored primarily on Elizabeth McMurchie's laptop computer
- A backup that is updated at least daily will be stored on an external hard drive.
- Images taken by Elizabeth McMurchie are also stored on the SEM computer at the Roy J. Carver High Resolution Microscopy facility
- Polaroid photos taken by Lynn G. Clark are also stored in 349 Bessey, with electronic copies on Lynn G. Clark's laptop

Digital locations

- Images, spreadsheets, and R scripts will be stored on the ISU Cybox and shared with lab members
- Images, spreadsheets, and R scripts will also be stored in an R project linked to a Github repository and updated at least once a day when changes have been made. This R project and Github repository are part of a course in data science in R
- A second Github repository and R project will be created to with images, spreadsheets, and R scripts upon completion of the data science in R course

Describe how you will manage access and security.

The Cybox folder is private and shared only with group members. The Github repository is currently public for educational purposes, so it can be shared with current professors. The Github repository that will be created on completion of the current data science in R course will be made private until publication of data, at which time it will become public.

Data Preservation

Describe which data have long term value and should be preserved. If the datasets are relatively small and easily stored, all of the data should be preserved. If the data are relatively large and space is limited, choices must be made about which data to preserve and which data to destroy.

For data generated from computer code, it will likely make more sense to keep scripts that generate the data rather than the data itself. In this case, care must be taken to include version number for all software used for reproducibility. State how this versioning will be handled.

All images, spreadsheets, and R scripts will be preserved using Dryad (part of the Dataone network), ensuring that they will be available indefinitely. The version of R used to create the R scripts will be recorded, to ensure that data analysis procedures are not lost with R updates.

Describe how the data selected for preservation will be made accessible and reusable for a long period of time.

All data, including images of specimens, will be preserved, as obtaining additional SEM images requires destructive sampling and is expensive. They will be made accessible through Dryad and ISU DataShare.

Data Sharing

Describe how your data will be shared internally with the Grass to Gas team.

Internally, data will be shared with the Lynn G. Clark lab through CyBox and the GitHub repository.

Describe default sharing agreements, e.g with the Grass to Gas team, statistical consulting service, etc.

There are no current default sharing agreements outside of the Lynn G. Clark lab.

Describe how the data will be shared post publication and post-award (e.g. ISU DataShare, PSU DataCommons, Dryad, etc.)

Data will be shared post publication with ISU DataShare and Dryad.

Roles and Responsibilities

Describe who is responsible for each aspect of data management including data collection, data validation, data analysis, and data publication. Personnel and responsibilities rotate so document changes in personnel roles as they happen.

Lynn G. Clark was responsible for imaging of some *Guadua* specimens, while others were imaged by Elizabeth McMurchie. Analysis of images, data analysis, and data publication are the responsibility of Elizabeth McMurchie, with validation by Lynn G. Clark. Lynn G. Clark is the PI for the lab and ultimately in charge of all procedures.

Additional Notes

Please provide addition notes about data management that are not addressed with the previous responses.

Planned Research Outputs

Data paper - "Evaluating patterns of variation in foliar micromorphology in Guadua (Poaceae: Bambusoideae: Bambuseae)"

This is intended to be part of a research paper, possibly published in the *Botanical Journal of the Linnean Society* or *Plant Systematics and Evolution*, possibly combined with information on *Guadua* lemma and palea micromorphology.

Planned research output details

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
Evaluating patterns of variation in foliar micromo ...	Data paper	Unspecified	Restricted	None specified		None specified	None specified	No	No