**Methods**

*Scanning electron microscopy (SEM)*

Eighty-three *Guadua* foliage leaf specimens were imaged following modified procedures based on those of Clark (1990). Samples of both the abaxial and adaxial surface of each leaf were imaged. To remove epicuticular waxes that obscure viewing of micromorphology, foliage leaf specimens were submerged in xylene twice for approximately six minutes and agitated at least three times during each treatment. At the end of each xylene treatment, samples were removed from xylene, rinsed with distilled water, and allowed to air dry. Once treatment was complete, samples were attached to sample holders with double-sided carbon tape and taken to the Roy J. Carver High Resolution Microscopy Faculty at Iowa State University to be coated with a thin layer of platinum using a Cressington 208HR sputter coater. A Hitachi SU4800 FE-SEM field emission scanning electron microscope was used to take SEM images of each sample at 200x and 900x magnifications.

*Data visualization and analysis*

The presence or absence of 29 micromorphological features was recorded for each specimen in a binary dataset (Supplemental Data). Data were analyzed in R (R Core Team 2017). When visualizing and analyzing our data, we considered three factors: habitat, habit, and region. We hypothesized that micromorphology of foliage leaves would be different for specimens from savanna-associated species when compared to specimens from forest and river associated species. We also hypothesized that micromorphology of foliage leaves would not be significantly different for specimens belonging to small arching species (typically less than 10 meters tall) compared to specimens from species that grow to greater lengths, including both tall and erect species and leaning/climbing species. Finally, we hypothesized that micromorphology of foliage leaves would not differ based on region in which the leaves were collected (Mexico, Central America, Andes, or Eastern South America).

The R packages readxl (Wickham and Bryan 2019) and tidyverse (Wickham et al. 2019) was used to assist in importation, visualization, and analysis of data. Prior to graphical visualization, a correlation analysis was run to determine whether the presence or absence of certain features was strongly correlated with the presence or absence of other features. The R package ggcorrplot (Kassambra 2019) was used to visualize the correlation table. Principal coordinates analysis (PCoA) was used to visualize our data. To run PCoA, we used the R packages ade4 (Chessel et al. 2004; Dray and Dufour 2007; Dray et al. 2007; Bougeard and Dray 2018; Thioulouse et al. 2018) and vegan (Oksanen et al. 2020). A binary distance matrix was required to conduct the PCoA and further analyses; simple matching coefficient was used to obtain the distance matrix, as this method takes into account both presence and absence of different features; presence or absence of micromorphological features are both important to our analyses.

To analyze our data, a factorial multivariate analysis of variance (MANOVA) was used. The basic equation for the MANOVA was Y = β0 + β1(Habitat) + β2(Habit) + β3(Region) + ε, with each of the 29 dependent variables treated as separate “Y”, but placed on a matrix so that the linear model could be run on each variable. The y-intercept for each dependent variable is β0, while β1 is the coefficient for habitat, β2 the coefficient for habit, and β3 the coefficient for region. Residual error not accounted for by the model parameters is represented by ε; error distribution was binomial. Permutation of randomized residuals in a linear model was used to conduct the factorial MANOVA with the R packages RRPP (Collyer and Adams 2018; 2019), and geomorph (Adams et al. 2021; Baken et al. 2021; Collyer and Adams 2018; 2021). The factorial MANOVA was run with 1000 permutations using the ordinary least squares estimation method and type I sums of squares and cross products.

Model comparison was conducted using a log likelihood test. Alternative models to the full model (including habitat, habit, and region), included models with just habit and habitat, just habit and region, and just habitat and region, as well as each of these three variables on its own. The model with the highest log likelihood value was selected as the model that best fit the data. The R packages pander (Daróczi and Tsegelskyi 2021) and readr (Wickham and Hester 2021) were required to conduct and view the log likelihood values, in addition to RRPP and geomorph.

**Results**

*Data visualization*

Strong correlations (correlation coefficient > 0.7) were typically found for silica bodies of the same type on both surfaces and presence of papillae with ridged saddle-shaped silica bodies (Figure 1). Excluding correlations for silica body types found on both surfaces of the leaf, the strongest correlations (correlation coefficient = 0.85) were found for presence of papillae on the long cells on the bulliform cells and interstomatal zones of the adaxial surface and presence of adaxial ridged saddle-shaped silica bodies, as well as presence of papillae on bulliform cells of the adaxial surface and presence of papillae on the long cells of the interstomatal zone on the adaxial surface (Figure 1).

When used to visualize our data, principal coordinates analysis (PCoA) indicated that based on the top two principal coordinate axes, specimens from savanna-associated habitats typically had different foliage leaf micromorphology than those belonging to forest of river-associated habitats (Figure 2). A PCoA colored by habit further indicated that based on the top two principal coordinate axes, specimens with a small arching habit typically had different foliage leaf micromorphology than those with leaning and climbing or tall and erect habits, which based on the first two coordinate axes of the PCoA appeared nearly indistinguishable from each other (Figure 3). However, other than the limited range for Andean specimens, there appeared to be few patterns in the PCoA based on region using the first two principal coordinate axes (Figure 4). Eigenvalues were obtained from the PCoA and used to construct a scree plot we to show the amount of variation explained along each principal coordinate number (Figure 5). The scree plot indicates that principal coordinate 1 is responsible for the majority of variation in the PCoA, followed by a steep drop in variation. The “elbow” of the scree plot, indicating the number of principal coordinates necessary to retain to capture the majority of variation, appears a to be located at approximately the fourth or fifth principal coordinate. By the 20th principal coordinate, variation is essentially negligible.

*Data analysis*

*Guadua* foliage leaf micromorphology of specimens from species associated with savanna habitats was significantly different than that of foliage leaf micromorphology of species associated with forest and riverine habitats (factorial MANOVA lm.rrpp, p = 0.001, F = 9.43, Table 1). Additionally, foliage leaf micromorphology of specimens with a small arching habit was typically different than those with a big and erect or climbing habit (factorial MANOVA lm.rrpp, p = 0.001, F = 24.97, Table 1). The foliage leaf micromorphology of at specimens belonging to at least one region differed significantly from those belonging to other regions (factorial MANOVA lm.rrpp, p = 0.003, F = 1.65, Table 1); however, due to small sample size, post hoc analysis to determine which region(s) significantly differed from others was not possible.

Using model selection, it was determined that the model considering habit and habitat but excluding region best fit the data, as this model had the highest log likelihood value (logLik = 3784, Table 2). The model considering all three variables (habit, habitat, and region) had the second highest log likelihood score( logLik = 3620, Table 2), followed by the model considering only habitat and region (logLik = 3558, Table 2), then only habit and region (logLik = 3556, Table 2), and finally habitat (logLik = 3484, Table 2), habit (logLik = 3483, Table 2), or region only, respectively (logLik = 3064, Table 2).

Table 1. Factorial multivariate analysis of variance for *Guadua* foliage leaf micromorphology, using residual randomization permutation procedure: randomization of null model residuals, considering the factors habit, habitat, and region.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | df | SS | MS | Rsq | F | Z | p-value  Pr (>F) |
| Habit | 1 | 3.1976 | 3.1976 | 0.21459 | 24.9672 | 5.1177 | 0.001 |
| Habitat | 1 | 1.2081 | 1.2081 | 008108 | 9.4332 | 5.6533 | 0.001 |
| Region | 3 | 0.6335 | 0.2112 | 0.04251 | 1.6488 | 2.7194 | 0.003 |
| Residuals | 77 | 9.8615 | 9.8615 | 0.1281 | 0.66181 |  |  |
| Total | 82 | 14.9007 |  |  |  |  |  |

Table 2. Log likelihood values (logLik), residual principal components number (residual.pc.no), penalty score, and Akaike information criterion (AIC) score for model comparison using residual randomization permutation procedure.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | logLik | residual.pc.no | penalty | AIC |
| Habit, habitat, and region | 3620 | 28 | 1218 | -6021 |
| Habit and habitat | 3784 | 29 | 1044 | -6524 |
| Habitat and region | 3558 | 29 | 1160 | -5957 |
| Habit and region | 3556 | 28 | 1160 | -5953 |
| Habitat | 3484 | 28 | 986 | -5982 |
| Habit | 3483 | 28 | 986 | -5981 |
| Region | 3064 | 26 | 1102 | -5026 |

**Figures**

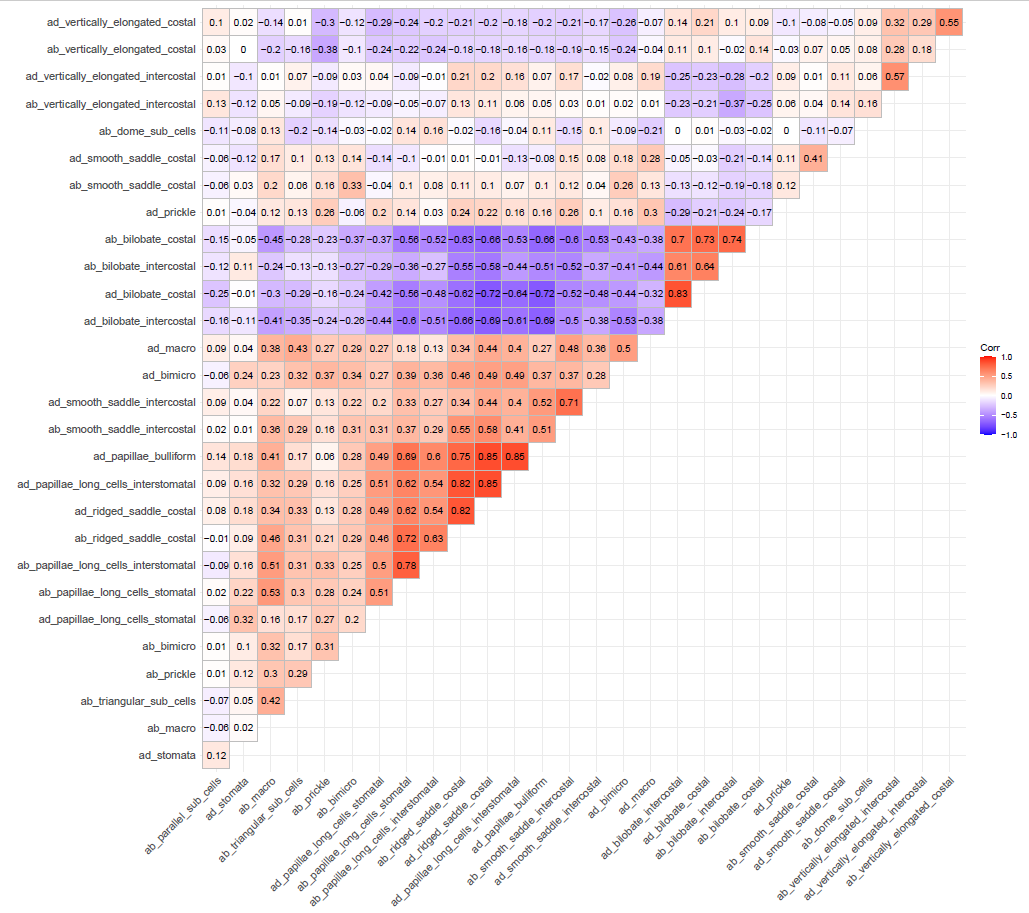


Figure 1. Correlation of presence/absence of various micromorphological features on *Guadua* foliage leaves. See data dictionary (Supplemental Data) for variable codes.

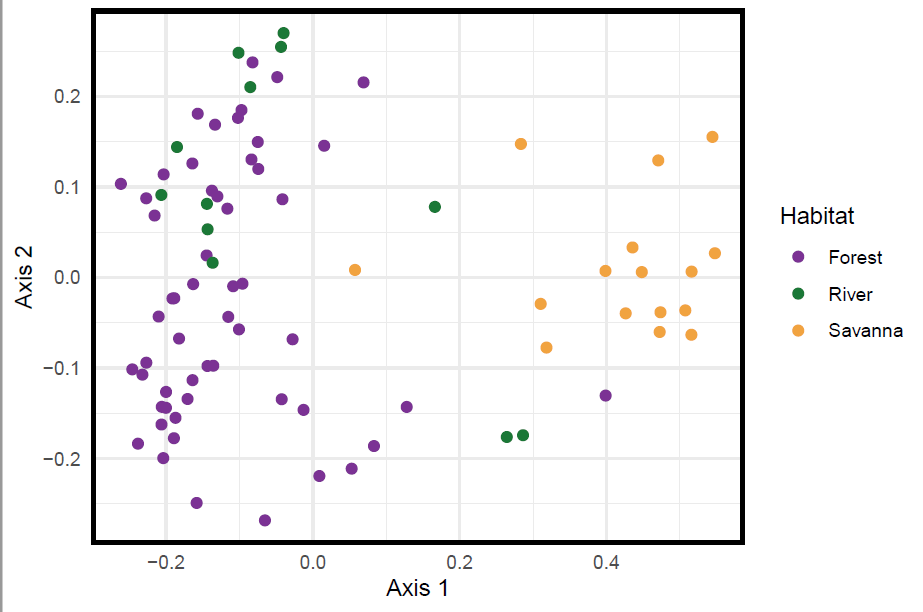


Figure 2. Principal coordinates analysis of *Guadua* foliage leaf micromorphology based on typical habitat associated with each specimen.

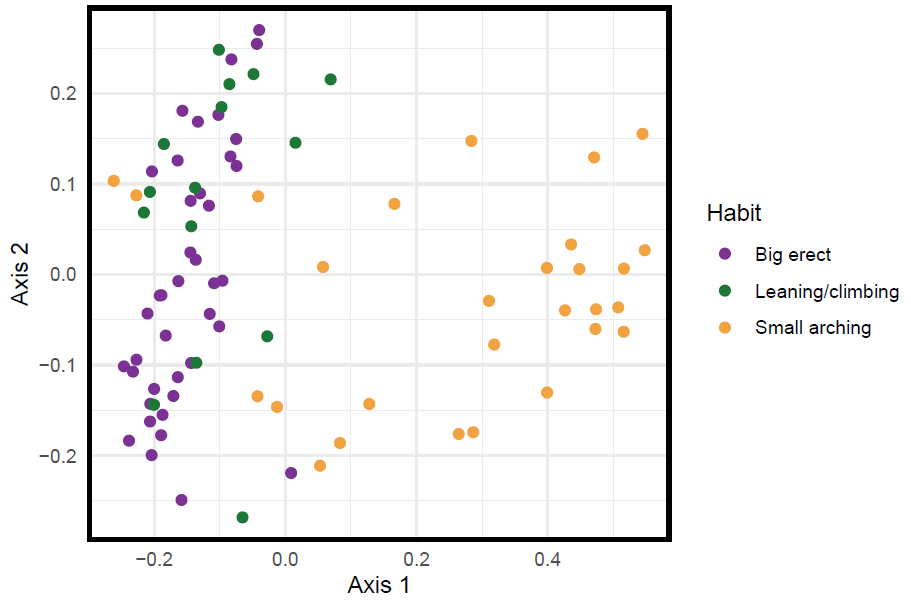


Figure 3. Principal coordinates analysis of *Guadua* foliage leaf micromorphology based on typical habit associated with each specimen.

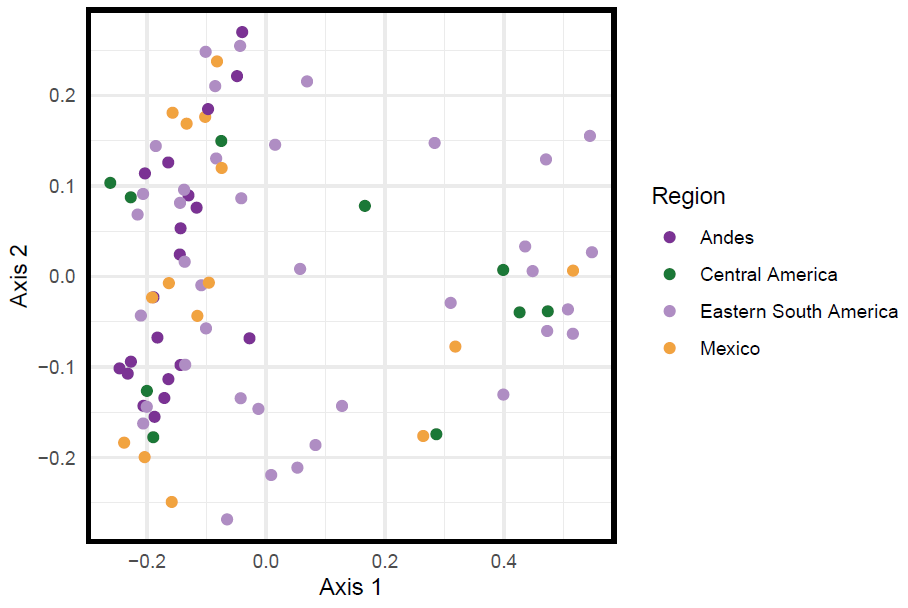


Figure 4. Principal coordinates analysis of *Guadua* foliage leaf micromorphology based on region in which each specimen was collected.

**References**

Adams D, Collyer M, Kaliontzopoulou A, Baken E (2021). “Geomorph: Software for geometric morphometric analyses. R package version 4.0.”

Baken E, Collyer M, Kaliontzopoulou A, Adams D (2021). “gmShiny and geomorph v4.0: new graphical interface and enhanced analytics for a comprehensive morphometric experience.”

Bougeard S, Dray S (2018). “Supervised Multiblock Analysis in R with the ade4 Package.” Journal of Statistical Software, 86(1), 1–17. doi: 10.18637/jss.v086.i01.

Chessel D, Dufour A, Thioulouse J (2004). “The ade4 Package – I: One-Table Methods.” R News, 4(1), 5–10. <https://cran.r-project.org/doc/Rnews/>.

Clark LG (1990) *Chusquea* sect. *Longiprophyllae* (Poaceae: Bambusoideae): A New Andean Section and New Species. Systematic Botany 15:617–634. https://doi.org/10.2307/2419158

Collyer M, Adams D (2018). “RRPP: An R package for fitting linear models to high‐ dimensional data using residual randomization.” https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.13029.

Collyer M, Adams D (2019). “RRPP: Linear Model Evaluation with Randomized Residuals in a Permutation Procedure. R package version 0.4.0.” <https://CRAN.R-> project.org/package=RRPP.

Collyer M, Adams D (2021). “RRPP: Linear Model Evaluation with Randomized Residuals in a Permutation Procedure, R package version 0.6.2.” <https://cran.r-> project.org/package=RRPP.

Daróczi G, Tsegelskyi R (2021). pander: An R 'Pandoc' Writer. R

package version 0.6.4. https://CRAN.R-project.org/package=pander

Dray S, Dufour A (2007). “The ade4 Package: Implementing the Duality Diagram for Ecologists.” Journal of Statistical Software, 22(4), 1–20. doi: 10.18637/jss.v022.i04.

Dray S, Dufour A, Chessel D (2007). “The ade4 Package – II: Two-Table and K-Table Methods.” R News, 7(2), 47–52. <https://cran.r-project.org/doc/Rnews/>.

Ellis RP (1979) A procedure for standardizing comparative leaf anatomy in the Poaceae. II. The epidermis as seen in surface view. Bothalia 12:641–671. <https://doi.org/10.4102/abc.v12i4.1441>

International Committee for Phytolith Taxonomy (ICPT), Neumann K, Strömberg CAE, et al (2019) International Code for Phytolith Nomenclature (ICPN) 2.0. Annals of Botany 124:189–199. https://doi.org/10.1093/aob/mcz064

Kassambara A (2019). ggcorrplot: Visualization of a Correlation Matrix using 'ggplot2'. R package version 0.1.3. https://CRAN.R-project.org/package=ggcorrplot

version 0.1.3. https://CRAN.R-project.org/package=ggcorrplot

Oksanen J, Blanchet G, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin P, O'Hara R, Simpson G, Solymos P, Stevens M, Szoecs E, Wagner H (2020). vegan:Community Ecology Package. R package version 2.5-7. https://CRAN.R-project.org/package=vegan.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Thioulouse J, Dray S, Dufour A, Siberchicot A, Jombart T, Pavoine S (2018). Multivariate Analysis of Ecological Data with ade4. Springer. doi: 10.1007/978-1-4939-8850-1.

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). “Welcome to the tidyverse.” Journal of Open Source Software, 4(43), 1686. doi: 10.21105/joss.01686.

Wickham H, Bryan J (2019). readxl: Read Excel Files. R package

version 1.3.1. <https://CRAN.R-project.org/package=readxl>

Wickham H, Hester J (2021). readr: Read Rectangular Text Data. R

package version 2.0.2. <https://CRAN.R-project.org/package=readr>