Acta Botanica Brasilica - 35(3): 339-351. July-September 2021.
doi: 10.1590/0102-33062020abb0368

Quantitative morphometrics suggest that the widespread Neotropical *Humiria balsamifera* (Aubl.) St. Hil. is a species complex

Ana Sofia Sousa de Holanda[1*], Alberto Vicentini[1], Jerome Chave[2], Fernanda Antunes Carvalho[3] and Charles E. Zartman[1]

Received: July 31, 2020
Accepted: October 20, 2020

ABSTRACT

Humiria balsamifera is an infraspecific complex of high phenotypic variation and widely distributed in northern South America. Leaf traits are traditionally considered the most relevant taxonomic characters for varietal level delimitation in the group. However, substantial phenotypic overlap among vegetative characters complicates taxonomic diagnoses in this complex. The objective of this study was to quantify and analyze phenotypic variation among individuals of the complex at a continental scale using uni- and multivariable analyses to assess whether morphometric analyses detect discontinuities. Secondarily, these quantitative data were used to test whether phenotypic similarity was related to geographic distance. Twenty- five quantitative and 27 qualitative character traits did not overtly reveal a topology corresponding to traditional varietal classification, nor to geographic structure. However, petiole length alone revealed a definitive separation of *H. balsamifera* var. *guianensis* (together with another recognized species *Humiria wurdackii*, and variety *H. balsamifera* var. *laurina*) from the rest of the taxa. Our results highlight substantial morphological overlap among vegetative and reproductive characters including those used in identification keys, and no morphological discontinuities suitable for clearly separating taxa within the complex were encountered demonstrating a future need to integrate multiple sources of information, including molecular data, to resolve this complex.

Keywords: Amazon, morphological species concept, phenotypic variation, quantitative taxonomy, species delimitation, white-sands forest

Introduction

The tropics, which have been exposed to long periods of frequent climatic variation over large areas (Wang *et al*. 2017), are likely to harbor many poorly connected populations through propagule dispersal thus producing "species complexes": groups of individuals which present wide phenotypic variation, yet lack perceivable discontinuities among potential taxonomic units (Grube & Kroken 2000; Prata *et al*. 2018). Such groups of high phenotypic variation are relatively common and offer, due to their high phenotypic overlap, demonstrably difficult challenges to traditional classification (Jacobs *et al*. 2019; Damasco *et al*. 2019). Evolutionarily, such species complexes may be assemblages of distinct, albeit cryptic lineages, which reflect geographical variation in phenotypic characters across environmental

---

1 Programa de Pós-Graduação em Botânica, Instituto Nacional de Pesquisas da Amazônia, 69060-001, Manaus, AM, Brazil
2 Laboratoire Evolution et Diversité Biologique, Université Paul Sabatier, 31062, Toulouse, France
3 Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, 31270-901, Belo Horizonte, MG, Brazil

* Corresponding author: anasofiaufpa@gmail.com

gradients signaling incipient speciation processes: recently diverging populations coupled with incomplete reproductive isolation (Mayr 1982; Coyne & Orr 2004).

The epicenter of the *Humiria balsamifera* (Humiriaceae, Malpighiales) complex is located in the Amazon Basin where knowledge of the Region´s floristic mega-diversity remains considerably limited (Hopkins 2007; Steege *et al*. 2016; Cardoso *et al*. 2017; Steege *et al*. 2019; Hopkins 2019). Indeed, many species complexes may lie unrecognized as a series of poorly collected, related 'species' awaiting the attention of a taxonomist, and such examples of frustrating situations motivated the coining of the term 'ochlospecies' (the Greek root ochlos meaning 'irregular', but also 'troublesome' or 'annoying'; White 1962). Although such ochlospecies may be considered a nightmare for taxonomists (White 1998), in some plant groups studies integrating multiple levels of information (*e.g.*, Endara *et al*. 2018; Prata *et al* 2018; Damasco *et al*. 2019) have indeed succeeded in shedding light on challenging infra- and inter-specific delimitations. Ultimately, the task of identifying such species groups is the first step towards understanding the mechanisms and processes of speciation in the richest plant biome on the planet: the Amazon Basin.

Delimiting species in recently diverging lineages is complicated due to potential lack of correlation among phenotypes and genotypes (*i.e.*, homoplastic traits) resulting from, for example, ancestral polymorphisms, and/or weak pre- or post-zygotic barriers (Raxworthy *et al*. 2007). Such emerging differences confound interpretations of morphological species and force taxonomists to resort to a multiple evidence consensus as a means of justifying taxonomic affinities. Ideally, an integrative taxonomic approach to delimiting complex species groups, such as coupling quantitative phenotypic and genetic data, is necessary.

Tropical species complexes distributed across continental scales are common (Thorne 1972; Pinheiro *et al*. 2018). Advances in DNA sequencing in conjunction with biosystematic studies have shed light on mechanisms which contribute to recent divergences within some Amazonian complexes (*e.g.*, Esteves & Vicentini (2013): *Pagamea coriacea*, Holanda *et al*. (2015): two sympatric varieties of the *H. balsamifera* complex, Dexter *et al*. (2017): *Inga* (Fabaceae), Prata *et al*. (2018): *Pagamea guianensis* (Rubiaceae), and Damasco *et al*. (2019): *Protium cordatum* (Burseraceae)). However, little remains known as to how phenotypic and phylogenetic variation are correlated among the vast majority of Neotropical complexes. Indeed, central to understanding processes of speciation lie in quantitative evaluations of phenotypic variation within these complexes (Mace 2004; Bacon *et al*. 2012; Duminil *et al*. 2012; Jacobs *et al*. 2019).

Species are recognized through phenotypic differences based on discontinuities among groups of character traits. However, such differences are rarely determined through objective or quantitative methodologies (McDade 1995; Zapata & Jiménez 2011; Yang *et al*. 2014). Although molecular evidence is the most evolutionarily informative approach to verifying typical taxonomic classifications, it may miss valid species in widely distributed, recently derived lineages due to, among other things, incomplete lineage sorting (Sites & Marshall 2003). The exploration of character traits by using quantitative data allows for establishing hypotheses based on sub-specific classifications thus offering a framework from which phylogenetic hypotheses may be tested (Zapata & Jiménez 2011; Saraiva *et al*. 2015; Trofimov *et al*. 2016; Jacobs *et al*. 2019).

In species complexes characterized by high phenotypic variation, difficulties arise in identifying taxonomically, and evolutionarily relevant character traits (Atria *et al*. 2017). The most common approach to dealing with taxonomically problematic plant groups has been the use of multivariate statistics (Henderson & Martins 2002; Knudsen 2002; Bacon & Bailey 2006; Henderson 2006; 2011; Atria *et al*. 2017). Such tools refine and summarize phenotypic discontinuities where they may exist among closely related taxa. Furthermore, such approaches allow for greater efficiency in recognizing diagnostic characters key to sub-specific classification (Pinheiro & Barros 2007; Boratynski *et al*. 2013; Biye *et al*. 2016; Fernández *et al*. 2017).

Humiria is a small Neotropical genus in the Humiriaceae recognized by four species (*H. balsamifera*, *H. crassifolia*, *H. fruticosa*, and *H. wurdackii*) (Cuatrecasas1961). The genus consists of bushes, treelets and occasionally trees with pentamerous flowers, free petals, 20 uniseriate stamens of alternating sizes united by a basal tube; bisporangiate anthers with thickened connective tissue with tricomes; an intra-staminal disc surrounding the five-locular ovary ; woody endocarps with five apical foraminae; longitudinal germination valves containing between one to four fertile seeds. (Cuatrecasas 1961; Bove *et al*. 1997; Medeiros *et al*. 2015).

Humiria balsamifera is by far the most common, and widely distributed species in the genus occurring in sandy habitats such as *Campina* (low white-sands forest) and *Campinarana* (high white-sands forest) vegetation of the Amazon Basin, the *Cerrado* of central Brazil, as well sandstone hills throughout northern South America. In the Amazon, *H. balsamifera* is one of the most common species (Barbosa & Ferreira 2004; Ferreira 2009; Costa *et al*. 2020) in geomorphological formations associated with nutrient poor, sandy soils, which occupy ca. 7 % of the Amazon's surface area (Anderson 1981; Prance & Daly 1989; Prance 1996).

The genus *Humiria* is characterized by a wide range of morphological variation across its global distribution (Medeiros *et al*. 2015). Some extreme phenotypes within *H. balsamifera* were previously given species status, including *H. floribunda*, *H. guianensis*, *H. montana* and *H. parvifolia* (Urban 1877), but in the most recent monograph of Humiriaceae, Cuatrecasas (1961) lumped these into one single species thus recognizing *H. balsamifera* as a large, widespread species complex. Cuatrecasas (1961) also organized the species into 14 varieties and two forms, as recognized herein.

Leaf characters, such as form, size, width, presence/absence of pedicels, and trichomes, were considered the most taxonomically relevant characters in the *Humiria balsamifera* complex for determining varietal status (Cuatrecasas 1961). However, overlapping character trait variation blurs distinctions among varieties leaving doubts as to whether quantitative patterns of phenotypic variation in *H. balsamifera* would correspond to discrete, discontinuous taxonomic units within the complex.

The ecological and/or historical factors which contribute to morphological variation in this widespread Neotropical complex, *Humiria balsamifera*, remain elusive. Distinct phenotypic discontinuity among two sympatric varieties (*Humiria balsamifera* var. *guianensis* and *Humiria balsamifera* var. *balsamifera* f. *attenuata*) at fine scales was not explained by pollinator guild composition, nor results from inter-intra fertilization experiments suggesting that pre-zygotic barriers do not explain the phenotypic discontinuities among these two taxa. (Holanda *et al.* 2015).

The main goal here is to evaluate the classification of the *Humiria balsamifera* complex as conceptualized by Cuatrecasas (1961) by using a large dataset of quantitative and qualitative characters representing both vegetative and reproductive traits, in order to answer the following questions: 1) Can phenotypic discontinuities be detected among species and/or varieties of the *Humiria balsamifera* complex? 2) If so, which characters best represent such discontinuities? And finally, 3) Is phenotypic variation geographically structured?

Materials and methods

Study specimens

A total of 345 individuals from 6 countries and 69 localities were examined in this work. The majority of this material (210 specimens) represent herbarium material housed at INPA. Additional 135 specimens were collected during recent collecting trips across the Brazilian range of *H. balsamifera*. Varietal limits were established *a priori* based on identifications using Cuatrecasas (1961) as well as the use of comparisons made from his own confirmations of specimens included in this study. Type and isotype specimens for 11 sub-specific taxa were also measured from NY and US Herbaria based on high-resolution images when available. Measurements of physical specimens were done by using a stereomicroscope, and digital pachymeter on rehydrated flowers, and dried leaves. A total of 52 morphological characters were evaluated for each specimen, of which 27 were quantitative (eight vegetative and 19 reproductive), and 25 qualitative (nine vegetative and 16 reproductive) (Tab. S1 in supplementary material). The list with species names and varieties as well as the number of samples per specimen used in these analyses are found in Table 1.

We included six varieties of *H. balsamifera* (*H. balsamifera* var. *balsamifera*, *H. balsamifera* var. *floribunda*, *H. balsamifera* var. *guianensis*, *H. balsamifera* var. *laurina*, *H. balsamifera* var. *parvifolia*, and *H. balsamifera* var. *stenocarpa)*, and one species endemic to the central-western portion of the Amazon Basin: *H. wurdackii*. The remaining ten sampled taxa are known only from their holotype or are rarely collected, geographically restricted taxa to areas of extremely remote

Table 1. List of measured taxa. Scientific names followed by the respective authors and sample numbers.

Scientific name	Author	N
Humiria balsamifera var. *balsamifera* *	Cuatrec.	75
Humiria balsamifera var. *floribunda*	(Mart.) Cuatrec.	99
Humiria balsamifera var. *guianensis*	(Benth.) Cuatrec.	116
Humiria balsamifera var. *stenocarpa*	Cuatrec.	12
Humiria balsamifera var. *laurina*	(Urb.) Cuatrec.	8
Humiria balsamifera var. *parvifolia*	(A. Juss.) Cuatrec.	11
Humiria balsamifera var. *coriacea***	Cuatrec.	4
Humiria balsamifera var. *subsessilis***	(Urb.) Cuatrec.	5
Humiria balsamifera var. *guaiquinimana***	Cuatrec.	1
Humiria balsamifera var. *iluana***	Cuatrec.	1
Humiria balsamifera var. *imbaimadaiensis***	Cuatrec.	1
Humiria balsamifera var. *pilosa***	(Steyerm.) Cuatrec.	1
Humiria balsamifera var. *savannarum***	(Gleason) Cuatrec.	1
Humiria balsamifera var. *minarum***	Cuatrec	1
Humiria wudackii	Cuatrec.	6
*Humiria crassifolia***	Mart. ex Urb	2
*Humiria fruticosa***	Cuatrec.	1
Total		**345**

*including (f. *balsamifera* e f. *attenuata*) forms of *Humiria balsamifera* var. *balsamifera*

**Not included in PCA and UPGMA Cluster analysis.

access. Therefore, these taxa (*H. crassifolia*, *H. fruticosa*, *H. balsamifera* var. *coriacea*, *H. balsamifera* var. *subsessilis*, *H. balsamifera* var. *guaiquinimana*, *H. balsamifera* var. *iluana*, *H. balsamifera* var. *imbaimadaiensis*, *H. balsamifera* var. *pilosa*, *H. balsamifera* var. *savannarum* and *H. balsamifera* var. *minarum*) were left out of PCA and UPGMA analysis. However, all taxa were included in the Non-metric Multidimensional Scaling (NMDS) analysis.

Geographic distribution

The map of the geographical distribution of all sampled taxa was made using QGIS 3.6.3 (QGIS 2019) by entering the coordinates given on herbarium labels, or based on the location records available for the species and varieties of *Humiria* analyzed in this study. Complementary information on the global geographical distribution of the specimens was obtained from databases such as GBIF and Specieslink, and subsequently filtered to the level of identification (only identifications to species and varieties were used) in consideration of the identifier of this material. For GBIF and Specieslink coordinates were validated using Google Earth. Our sampling included Colombia, Venezuela, Suriname, French Guiana and all major Brazilian domains (Amazon Forest, Cerrado, Restinga, Savannah and Atlantic Forest) within the global distribution of *H. balsamifera*. When coordinates were not reported, they were manually georeferenced based on information provided by the collector (*e.g.*, local rivers, roads, hills, mountains, communities, and/or towns).

Data preparation

We quantified the morphological variation of the following vegetative characters: petiole length (PL), petiole width (PW), leaf blade length (BL), leaf blade width (BW), number of secondary veins (VN), and distance among secondary veins (VD). To this end, we used a mean calculated from at least three mature leaves from each specimen. For the remaining characters, we utilized one measurement from each sample. All characters were transformed into millimeters, and log-transformed for statistical analyses. Boxplots were constructed using absolute values to visualize the raw infra-specific differences in phenotypic traits among taxa. Categorical (qualitative) data were utilized in cluster analysis after transformation into presence-absence data.

Individuals with missing data for any measurement were removed from all analyses. Due to the paucity of data, characters such as inflorescence length (IL), fruit length (FL) and fruit diameter (FD) were removed from multivariate analysis; however, they were included in univariate analyses.

Morphological analysis

In order to summarize phenotypic variation among sampled individuals, we performed a Non-metric Multidimensional Scaling (NMDS) using Gower's dissimilarity index (Gower 1971) with both vegetative and reproductive data. Principal Component Analysis (PCA) was also executed in R v.3.6.1 (R Development Core Team 2019) in order to visualize possible clustering among vegetative and reproductive characters by evaluating PCA scores. In order to compare the relative contribution of vegetative and floral characters from the PCA results, analyses were conducted separately for each group as well. UPGMA cluster analyses were also used by applying unweighted pairwise comparisons with arithmetic means (Michener & Sokal 1957), and Gower's index with both continuous and categorical data. In order to test for geographic structure in phenotypic variation, we performed a Mantel test (Manly 1997: 10,000 permutations) to test the relationship among morphological and geographic distance.

Results

Geographical variation

Of the 345 *Humiria* individuals sampled, 327 represented seven of the 19 taxa recognized by Cuatrecasas (1961), the remaining 18 individuals represented 10 rare taxa, each with less than six sampled individuals. The distribution map (Fig. 1) revealed overlapping regional ranges of most varieties and few examples of geographically restricted taxa. The map also substantially expands the global distribution of *Humiria* as originally described by Cuatrecasas (1961) and, more recently, amended by Medeiros *et al.* (2015). The distribution of *H. balsamifera* var. *laurina* is extended herein to include Amazonas and Roraima states, in savannahs and campinaranas; *H. balsamifera* var. *floribunda* to Mato Grosso State; *H. balsamifera* var. *balsamifera* to Rondônia State; *H. balsamifera* var. *subsessilis* to Pará State; and *H. balsamifera* var. *parvifolia* to Amazonas State.

Morphological variation

There is substantial overlap among vegetative and reproductive traits of all taxa sampled (Fig. S1 and S2 in supplementary material). However some characters suggest discontinuities within *H. balsamifera*: *i*) petiole lengths are larger in *H. balsamifera* var. *guianensis*, however, this group includes another species *H. wurdackii* within its range of phenotypic variation (Fig. S1A in supplementary material); ii) blade widths are wider in varieties *H. balsamifera* var. *balsamifera*, *H. balsamifera* var. *floribunda,* and *H. balsamifera* var. *guianensis* (Fig. S1D in supplementary material) which share also more ovate leaf forms when compared to *H. balsamifera* var. *laurina,* and *H. wurdackii*. Overlap among reproductive characters was greater than that for vegetative characters. *H. balsamifera* var. *stenocarpa* showed an average petal length greater than that of *H. balsamifera* var. *laurina,* and *H. balsamifera* var. *parvifolia*; however, remains within the range of variation in the length of petals of the varieties

H. balsamifera var. *balsamifera*, *H. balsamifera* var. *floribunda*, *H. balsamifera* var. *guianensis,* and *H. wurdackii* (Fig. S2E in supplementary material).

Morphometric analyses

The NMDS result (Fig. 2) showed a clear segregation of *H. balsamifera* var. *guianensis* (blue squares) from other varieties, considerable overlap between *H. balsamifera* var. *balsamifera* (red circle), as well as *H. balsamifera* var. *floribunda* (green circle). Relative separation of *H. balsamifera* var. *parvifolia* (orange triangle), and *H. wudackii* (magenta square) in the top right was also observed.

Of the taxa excluded in the PCAs due to the lack of enough sampling, most of them were scattered among the most sampled taxa. Two notable exceptions, however, were *H. crassifolia* and *H. fruticosa*, both taxa distant from other varieties in the morphospace, thus supporting Cuatrecasas (1961) to erect them as separate species.

When combining vegetative and reproductive data, PCA revealed a wide dispersion among all taxa as the percentage of variation explained by the two principal axes was relatively low: 23 % (PCA1) and 14 % (PCA2) (Fig. 3).

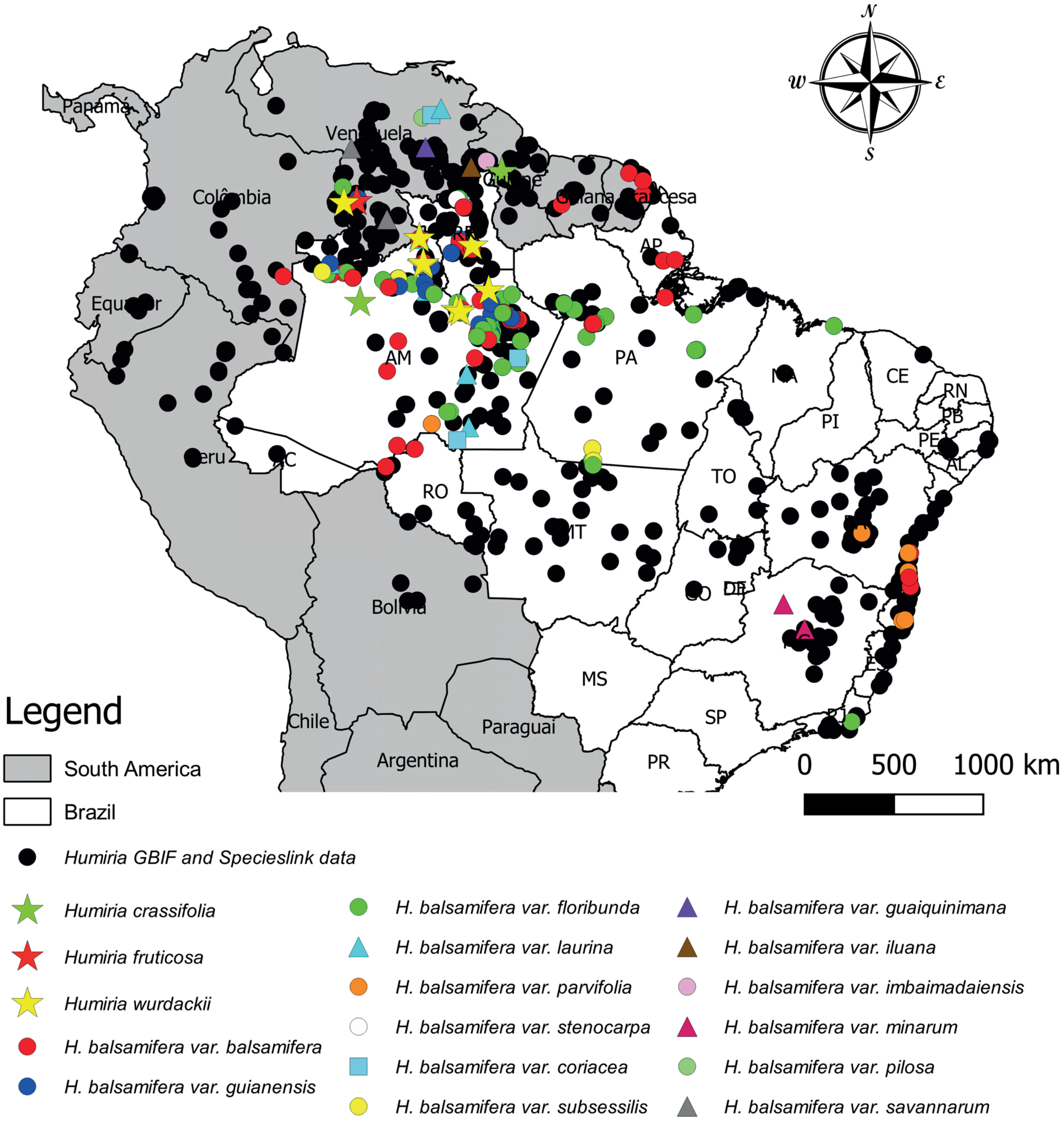

Figure 1. Regional distribution map of the *Humiria*, based on museum collection data. Data source: Specieslink, GBIF and field collections.

When vegetative and reproductive characters were analyzed separately, another pattern in morphospace was revealed. In the PCA generated with vegetative data (Fig. 4), *H. balsamifera* var. *guianensis* (blue circles) appeared grouped and relatively isolated from other taxa, whereas *H. balsamifera* var. *balsamifera* (red circles), and *H. balsamifera* var. *floribunda* (green circles) were grouped together, however *H. balsamifera* var. *balsamifera* is more widely distributed than *H. balsamifera* var. *floribunda*. Another group can be observed among *H. balsamifera* var. *stenocarpa* (black circle), *H. balsamifera* var. *parvifolia* (orange circle), *H. balsamifera* var. *laurina* (cyan circle) and *H. wurdackii* (yellow circle) in the top right on graph, which remained relatively isolated from the other varieties.

Vegetative data explained more variation in both PCA1 (34 %) and PCA2 (21 %) than when combined with reproductive data. In this analysis the characters with the greatest influence on variation of *H. balsamifera*

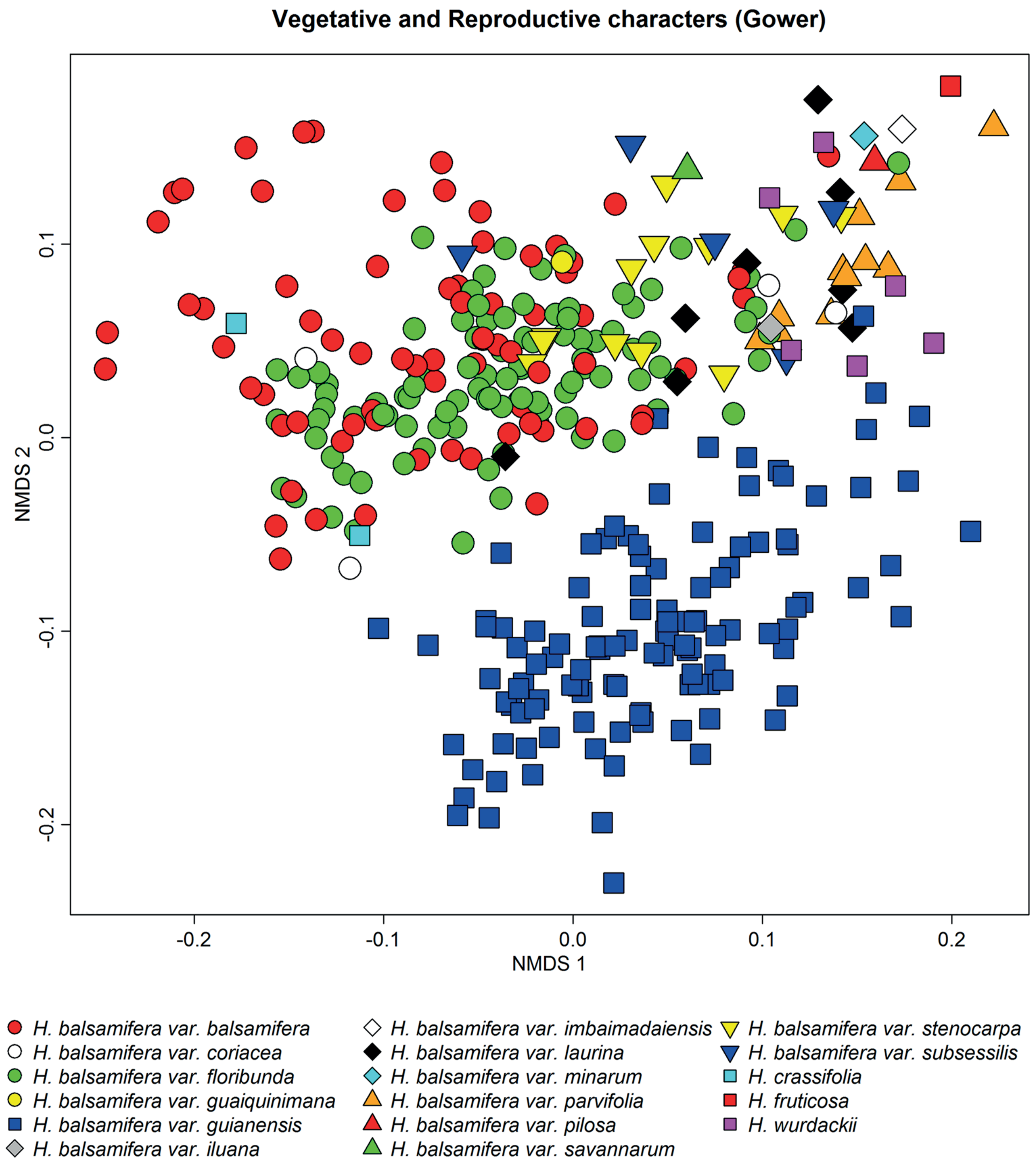

Figure 2. Non-metric Multidimensional Scaling (NMDS) using Gower's dissimilarity index of all varieties of *Humiria balsamifera* and additional species in genus *Humiria*.

var. *guianensis* in relation to other varieties were length and width of the petiole, followed by stipule width. The characters which most influenced grouping among *H. balsamifera* var. *balsamifera* and *H. balsamifera* var. *floribunda* were leaf blade length, distance among secondary veins and vein number. The grouping between the varieties *H. balsamifera* var. *stenocarpa*, *H. balsamifera* var. *parvifolia*, and *H. balsamifera* var. *laurina* and species *H. wurdackii* was influenced by leaf length and width, which are smaller than the other taxa.

The PCA with reproductive characters (Fig. 5) shows substantial overlap among taxa, precluding any clear visualization of grouping tendencies among varieties. On the other hand, vegetative characters alone, revealed a clustering of *H. balsamifera* var. *guianensis* and strong distinction of *H. balsamifera* var. *floribunda* and *H. balsamifera* var. *balsamifera* (Fig. 4).

The dendrogram resulting from the UPGMA cluster analysis using either combined vegetative and reproductive characters (Fig. S3 in supplementary material) did not reveal any consistent grouping among the varieties. However, the dendrogram with the vegetative characters (Fig.S4 in supplementary material) revealed a relatively consistent grouping with individuals of the *H. balsamifera*

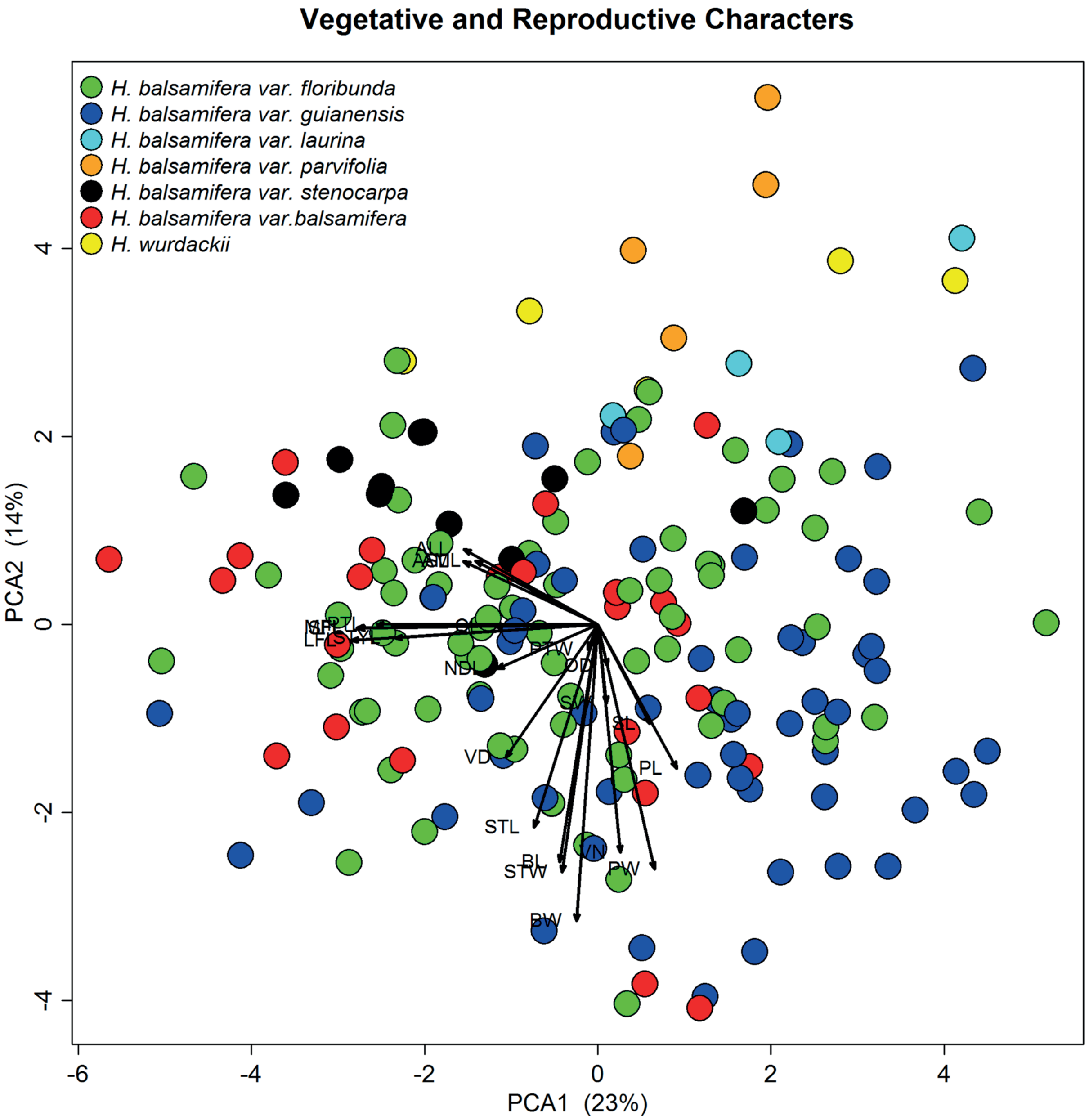

Figure 3. Principal Component Analysis (PCA) for vegetative and reproductive characters combined (**A**), vegetative alone (**B**) and reproductive alone (**C**). Color coding is as in Figure 3: *H. balsamifera* var. *balsamifera*: red, *H. balsamifera* var. *floribunda*: green, *H. balsamifera* var. *guianensis*: blue, *H. balsamifera* var. *laurina*: cyan, *H. balsamifera* var. *parvifolia*: orange, and *H. balsamifera* var. *stenocarpa*: black, *Humiria wurdackii* is represented in yellow.

var. *guianensis* nearly entirely grouped on the right side of the dendrogram, except for five individuals mixed with *H. wurdackii*, and one from *H. balsamifera* var. *laurina* within the *H. balsamifera* var. *guianensis* group. The dendrogram with reproductive data did not reveal any consistent grouping among the varieties (Fig. S5 in supplementary material)

This result was also observed when categorical data was utilized in these analyses (Fig. S6 in supplementary material). The dendrogram for vegetative data showed a similar pattern to that found in the PCA, where varieties *H. balsamifera* var. *balsamifera* and *floribunda* all show a diffuse pattern of clustering with variation most conserved in *H. balsamifera* var. *guianensis* as illustrated by two clusters nearly entirely dominated by this taxon. UPGMA clustering revealed strongest character fidelity with petiole length character (PL) (Fig. 6) by forming a cluster which includes nearly all individuals in variety *H. balsamifera* var. *guianensis*, one of *H. balsamifera* var. *laurina* and four individuals of *H. wurdackii*.

Finally, Mantel tests failed to reveal any significant relation among morphological and geographic distance (Tab. 2) suggesting an absence of geographic structure to the phenotypic variation in the *H. balsamifera* complex.

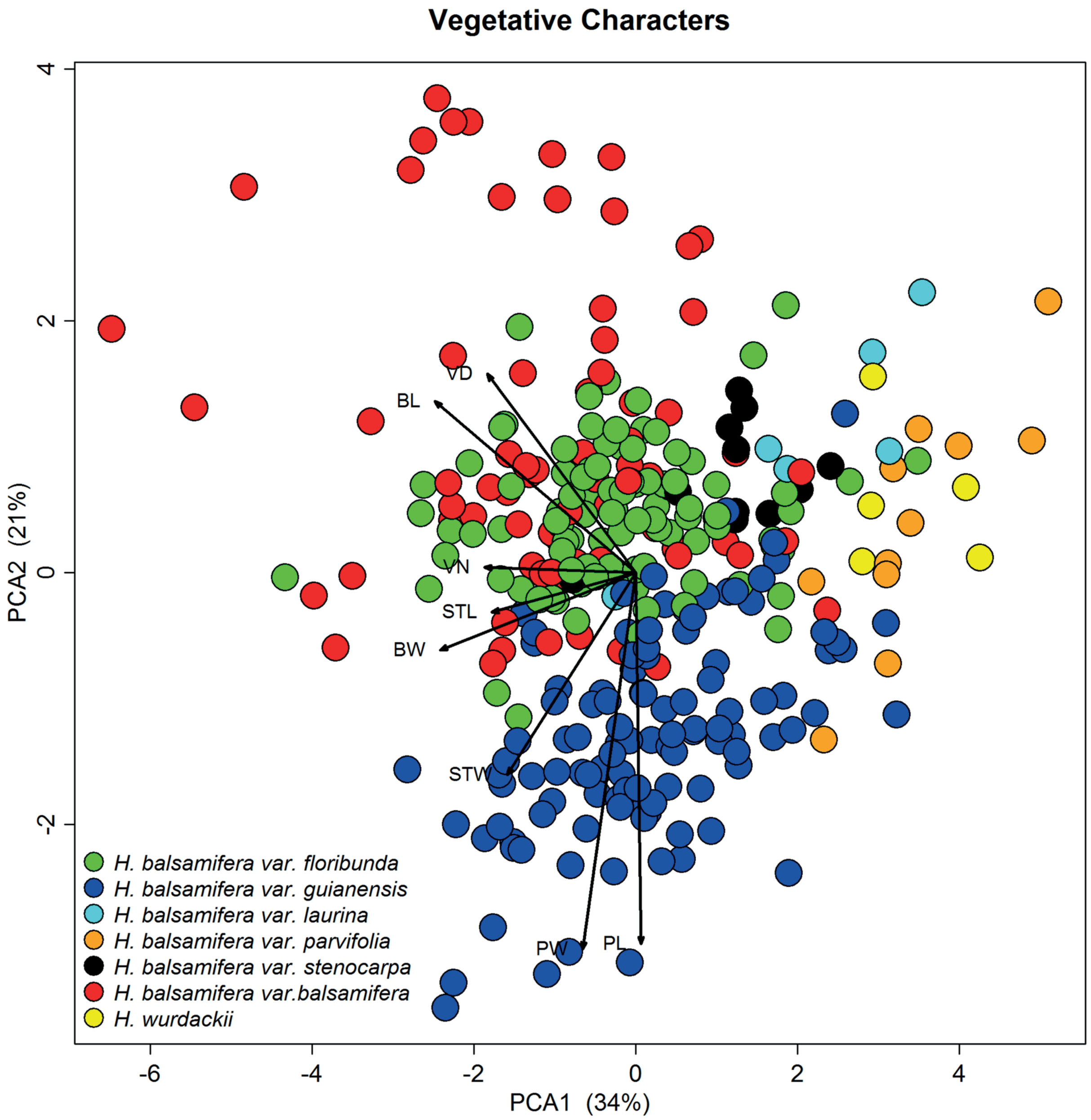

Figure 4. Principal Component Analysis (PCA) for vegetative characters. Color coding: *H. balsamifera* var. *balsamifera*: red, *H. balsamifera* var. *floribunda*: green, *H. balsamifera* var. *guianensis*: blue, *H. balsamifera* var. *laurina*: cyan, *H. balsamifera* var. *parvifolia*: orange, and *H. balsamifera* var. *stenocarpa*: black, *Humiria wurdackii* is represented in yellow.

Discussion

In this study we examined the morphology of two species of *Humiria* and six varieties of *Humiria balsamifera* based on vegetative characters considered taxonomically relevant in the most recent treatment for the genus (Cuatrecasas 1961). We found a large overlap of character traits considered diagnostic in distinguishing varieties in the keys presented by Cuatrecasas (1961). The exceptional morphological stasis among flowers and fruits of sub-specific entities is emphasized when only reproductive characters were used in the analyses. Herein, we show that phenotypic variation alone does not predict traditional taxonomic classification, nor is it geographically structured.

Our results suggest that phenotypic variation in *H. balsamifera* is also unrelated to continental scale historical events, or geological factors. For example, continental scale orogenic barriers such as the Andean Cordillera do not explain patterns of phenotypic variation as suggested by other authors such as Antonelli *et al* (2009) and Hoorn *et al.* (2010). Whether local environmental selective pressures or low genetic correlations suggest exceptional phenotypic plasticity within the complex waits to be tested. Nonetheless, studies have shown sharp phenotypic discontinuities between two of the most abundant varieties (Holanda *et*

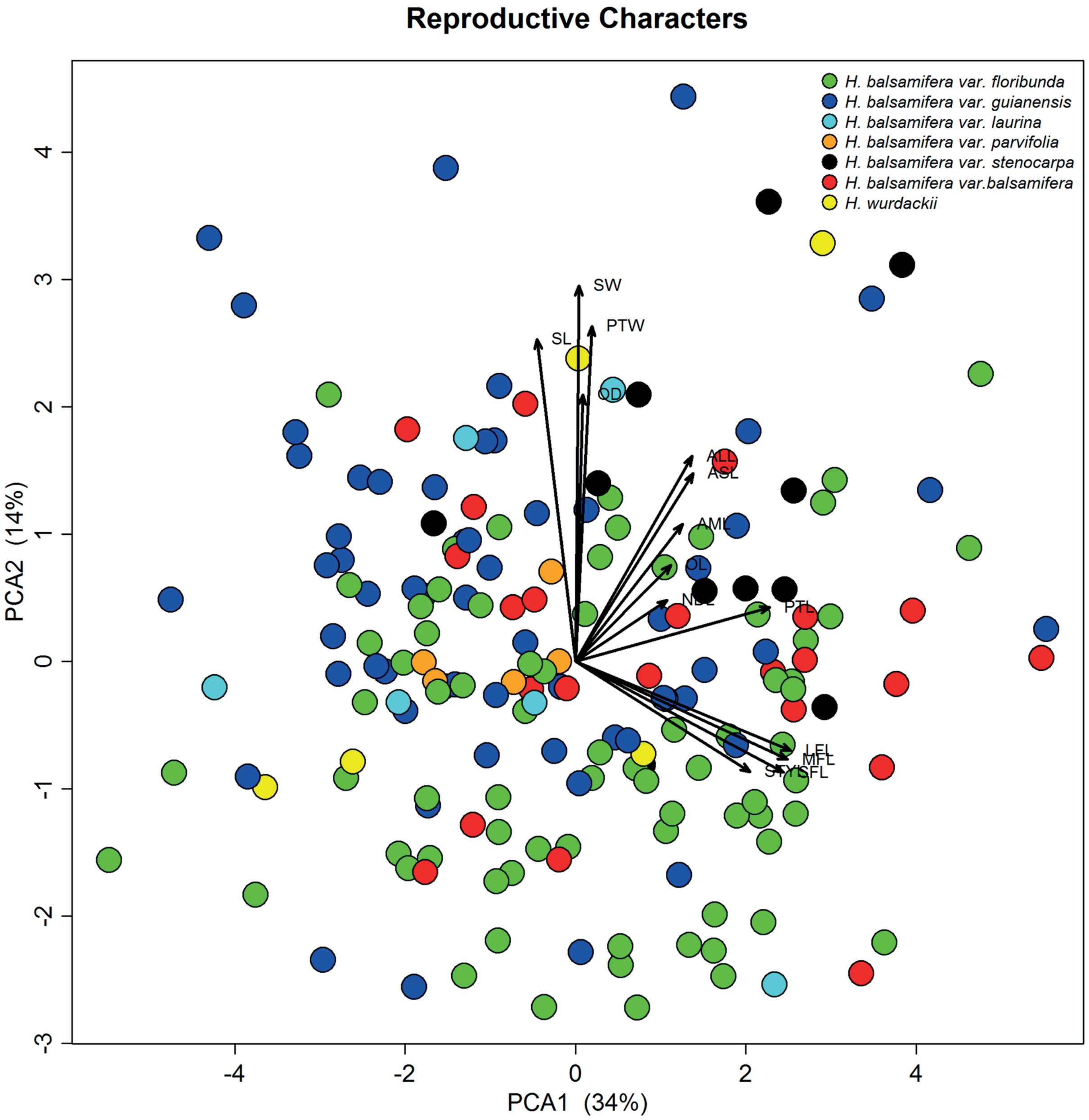

Figure 5. Principal Component Analysis (PCA) for reproductive characters. Color coding: *H. balsamifera* var. *balsamifera*: red, *H. balsamifera* var. *floribunda*: green, *H. balsamifera* var. *guianensis*: blue, *H. balsamifera* var. *laurina*: cyan, *H. balsamifera* var. *parvifolia*: orange, and *H. balsamifera* var. *stenocarpa*: black, *Humiria wurdackii* is represented in yellow.

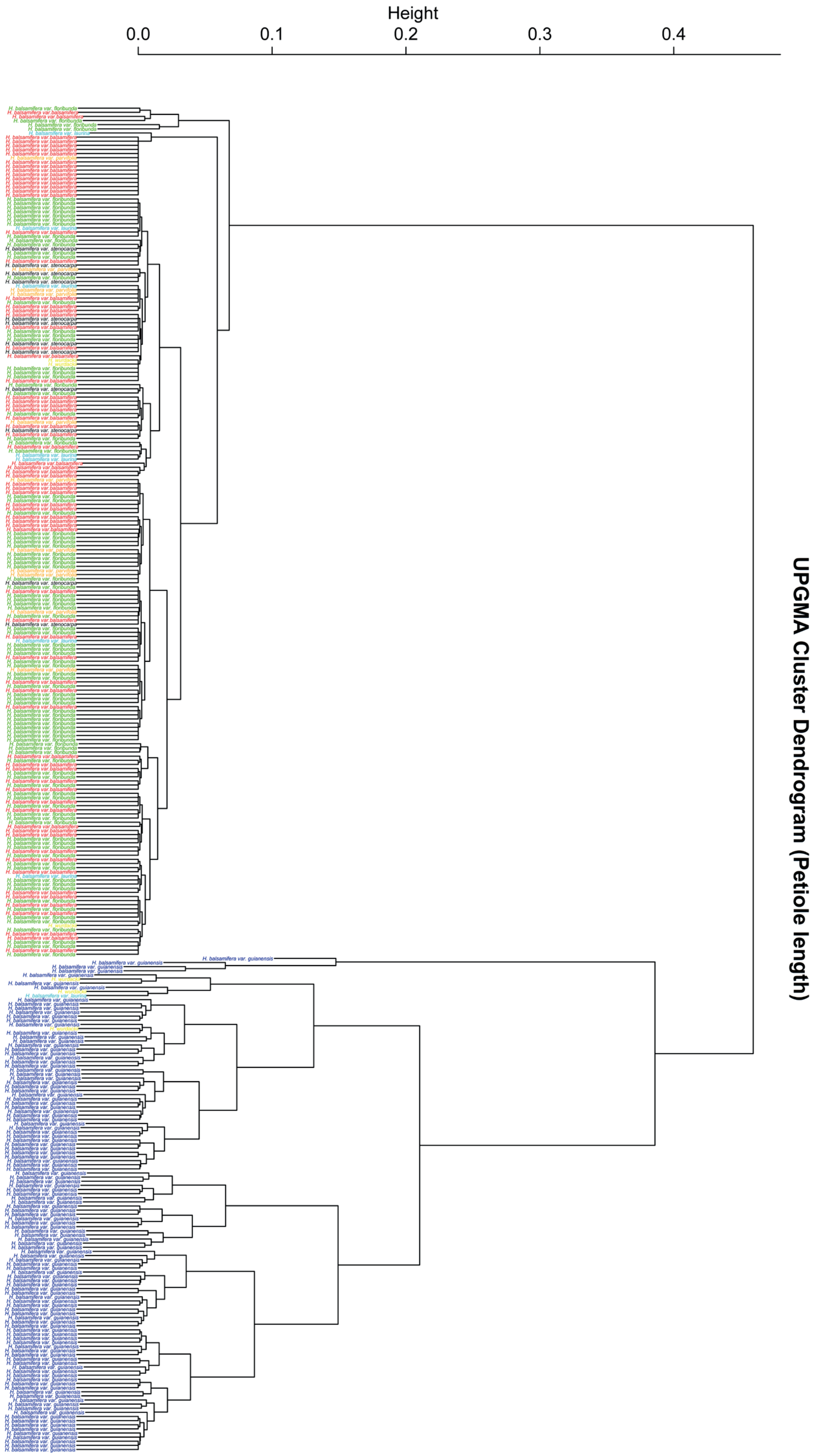

Figure 6. UPGMA Cluster Dendrogram of petiole length. Color coding is as in Figure 3: *H. balsamifera* var. *balsamifera*: red, var. *floribunda*: green, *H. balsamifera* var. *guianensis*: blue, *H. balsamifera* var. *laurina*: cyan, *H. balsamifera* var. *parvifolia*: orange, and *H. balsamifera* var. *stenocarpa*: black, *Humiria wurdackii* is represented in yellow.

Table 2. Mantel test of the comparison between the morphological distance matrices for vegetative data (MV) and for reproductive data (MR) versus Geographic Distance (DG) for the *Humiria balsamifera* complex.

Pairwise Comparison	r statistic	Significance	90 % CL	95 % CL	97.5 % CL	99 % CL
MV x GD	0.1754	0.099	0.0509	0.0672	0.0814	0.0974
MR x GD	0.1562	0.060	0.123	0.163	0.194	0.225

al. 2015) at fine scales, but on larger spatial scales, as shown here, these discontinuities dissipate suggesting that the detectability of phenotypic variation is scale dependent.

Phenotypic plasticity indeed potentially confounds results presented in this study. High levels of phenotypic plasticity could be potentially problematic when identifying and describing species. For example, Rutherford (2020) showed, with nine closely related species of eucalyptus (*Eucalyptus* subgenus *Eucalyptus*) (Myrtaceae), that plasticity in leaf morphology may have confused species boundaries in the group, and the author suggests an interdisciplinary approach to provided greater insights into patterns of speciation and divergence in this group.

Our study also extends the distributional ranges of some of the most common varieties of *H. balsamifera* as considered by Cuatrecasas (1961). Here we register geographic expansions of five varieties (*H. balsamifera* var. *balsamifera*, *H. balsamifera* var. *floribunda*, *H. balsamifera* var. *laurina*, *H. balsamifera* var. *subsessilis* and *H. balsamifera* var. *parvifolia*) for five Brazilian states thus reflecting that *Humiria* and Humiriaceae, in general, despite its wide distribution and high global abundance, remains a relatively neglected group in need of more detailed study.

Inconsistencies still exist in relation to the geographical distribution of *Humiria balsamifera*. According to the Flora do Brasil 2020 (2020), *H. balsamifera* var. *floribunda* does not occur in the state of Acre, however we identified collections of this taxa from Acre deposited at US (collected by Ferreira, CAC 10936; US 3598460). According to Medeiros *et al.* (2015), *H. balsamifera* var. *guianensis* does not occur in the state of Pará, however, Cuatrecasas (1961) identified both of these taxa as represented by material studied from US (Cavalcante, PB.2550; US 2951546) and NY (Cavalcante, PB, 2549; NY 02398908). Such new distributional information emphasizes the necessity for a taxonomic revision of *Humiria*, starting with a more detailed taxonomic and phylogenetic understanding of the *H. balsamifera* complex.

The low correlation for both vegetative and reproductive characters with geographical structure (Tab. 2) may be explained by the fact that geographical variation has little influence on morphology, either because environmental gradients are too weak to select characters, or because past selection has since been blurred by secondary contact or possibly temporal displacement of clines. Alternatively, the lack of geographic structure may be related to imbalanced sample size in this study, given that most analyzed samples were disproportionately made from the Brazilian Amazon Basin) leaving Northeastern Brazil and other countries of the Amazonian Biome under sampled.

This morphological overlap is also confirmed by multivariate analyzes (NMDS and PCA) and Cluster UPGMA. Despite verifying trends in the grouping and separation between some groups of taxa, only variety *H. balsamifera* var. *guianensis* exhibited clear morphological discontinuity for vegetative characters. The character most faithful to traditional concepts was the length of the petiole, which separated *H. balsamifera* var. *guianensis* from the other varieties. Indeed, this same character was also useful to segregate two varieties of *Humiria balsamifera* at fine scales (Holanda *et al.* 2015). Although the UPGMA shows overlap among *H. balsamifera* var. *guianensis* and *H. wurdackii* these two taxa are well distinguished groups in NMDS and PCA analyses. Future evidence from molecular data should contribute to our understanding of the phylogenetic relationships among these two entities as well as the entire complex itself. Varietal status remains a common albeit nebulous area in plant taxonomy as demonstrated, for example, in *Brasiliorchis picta* (Orchidaceae) (Pinheiro & Barros 2009), *Eugenia involucrata* group (Myrtaceae) (Bünger *et al.* 2015), the *Castilleja pilosa* species complex (Orobanchaceae) (Jacobs *et al.* 2019), the *Calamus javensis* complex (Arecaceae) (Atria *et al.* 2017).

The use of morphology in conjunction with ecological (including habitat preference, and reproductive biology) and genetic data, has aided in disentangling groups notoriously difficult in terms of taxonomic delimitation (Padial *et al.* 2010; Schlick-Steiner *et al.* 2010). Recently, integrative approaches have been used to refine sub-specific taxonomic status among species complexes (Pessoa *et al.* 2012; Esteves & Vicentini 2013; Rabelo 2016; Damasco *et al.* 2019; Li *et al.* 2019). For example, Prata *et al.* (2018) used genomic, ecological, morphological, and spectral data to resolve the taxonomy of the *Pagamea guianensis* complex. As a result, 15 well-supported taxa were described, thus illustrating the importance of multiple lines of evidence as an effective approach to better understanding the evolutionary history and taxonomy of widespread species complexes in the Neotropics.

Given the results of this study, integrative taxonomy fundamentally contributed to a better understanding of a taxonomic delimitation with regards to the traditional morphological concepts of sub-specific taxa in *Humiria balsamifera*. Molecular data is needed to provide a more comprehensive understanding of the phenotypic variation, and of the phylogenetic underpinnings within this group. Other sources of information such as anatomical, chemical, ecological and spectral data would also contribute valuable information to understanding the *H. balsamifera* complex as has been shown with great potential to resolve species delimitation problems in other large Amazonian species complexes (Durgante *et al*. 2013; Lang *et al*. 2015; 2017).

Conclusion

The varieties of *Humiria balsamifera* investigated in this study show high phenotypic overlap both vegetative and reproductive characters. An integrative approach including molecular, ecological, spectral as well as morphological data is suggested herein to most likely yield more definitive advances in the delimitation of species, and sub-species within this complex as well as for understanding the evolutionary interpretations of the divergences among lineages that comprise these taxa.

Acknowledgements

We thank the Instituto Nacional de Pesquisas da Amazônia (INPA), the INPA Herbarium, LABCRIP and NAPPA/LBA for support with material, equipment and excursions, the Conselho Nacional de Desenvolvimento Científico e Tecnológico- CNPq through the Taxonomy Training Program (PROTAX Nº 001/2015), Fundação de Amparo à Pesquisa do Estado do Amazonas- FAPEAM (Grants: 062.00806/2017; 018/2015; POSGRAD 2015/INPA). J.C. acknowledges funding by "Investissement d'Avenir" grants (CEBA, ANR-10-LABX-25-01 and TULIP, ANR-10-LABX-0041). C.E.Z acknowledges funding by grants CNPq (No. 441590/2016-0), and FAPEAM (No.015/2016) Project PELD/MAUA.

References

Anderson AB. 1981. White sand vegetation of Brazilian Amazonia. Biotropica 13: 199-210.

Antonelli A, Nylander JAA, Persson C, Sanmartín I. 2009. Tracing the impact of the Andean uplift on Neotropical plant evolution. Proceedings of the National Academy of Sciences-PNAS 106: 9749-9754.

Atria M, Mil H, Baker WJ, Dransfield J, Welzen P. 2017. Morphometric Analysis of the Rattan Calamus javensis Complex (Arecaceae: Calamoideae). Systematic Botany 42: 494-506.

Bacon CD, Bailey CD. 2006. Taxonomy and conservation: A case study from Chamaedorea alternans. Annals of Botany 98: 755-763.

Bacon CD, McKenna MJ, Simmons MP, Wagner WL. 2012. Evaluating multiple criteria for species delimitation: an empirical example using Hawaiian palms (Arecaceae). Evolutionary Biology 12: 23-40.

Barbosa RI, Ferreira CAC. 2004. Biomassa acima do solo de um ecossistema de 21 "campina" em Roraima, norte da Amazônia Brasileira. Acta Amazonica 34: 577-586.

Biye EH, Cron GV, Balkwill K. 2016. Morphometric delimitation of Gnetum species in Africa. Plant Systematics and Evolution 302: 1067-1082.

Boratynski A, Jasinska AK, Marcysiak K, *et al*. 2013. Morphological differentiation supports the genetic pattern of the geographic structure of Juniperus thurifera (Cupressaceae). Plant Systematics and Evolution 299: 773-784.

Bove CP. 1997. Phylogenetic analysis of Humiriaceae with notes on the Monophyly of Ixonanthaceae. Journal Computational Biology 2: 19-24.

Bünger MO, Einsehlor P, Figueiredo MLN, Stehmann JR. 2015. Resolving Species Delimitations in the *Eugenia involucrata* Group (Eugenia sect. Phyllocalyx - Myrtaceae) with Morphometric Analysis. Systematic Botany 40: 995-1002.

Cardoso D, Särkinen T, Alexander S, *et al*. 2017. Amazon plant diversity revealed by a taxonomically verified species list. Proceedings of the National Academy of Sciences 114: 10695-10700.

Costa FM, Terra-Araujo MH, Zartman CE, *et al*. 2020. Islands in a green ocean: Spatially structured endemism in Amazonian white-sand vegetation. Biotropica 52: 34-45.

Coyne JA, Orr HA. 2004. Speciation. Sunderland, MA, Sinauer.

Cuatrecasas J. 1961. A taxonomic revision of the Humiriaceae. In: Killip EP, Cuatrecasas J, Smith LB, Ewan J. 1968. Systematic plant studies. Contributions from the United States National Herbarium 35: 25-214.

Damasco G, Daly DC, Vicentini A, Fine PVA. 2019. Reestablishment of Protium cordatum (Burseraceae) based on integrative taxonomy. Taxon 68: 34-46.

Dexter KG, Lavin M, Torke BM, *et al*. 2017. Dispersal assembly of rain forest tree communities across the Amazon basin. Proceedings of the National Academy Sciences of the United States of America-PNASS 114: 2645-2650.

Duminil J, Kenfack D, Viscosi V, Grumiau L, Hardy O J. 2012. Testing species delimitation in sympatric species complexes: The case of an African tropical tree, *Carapa* spp. (Meliaceae). Molecular Phylogenetics and Evolution 62: 275-285.

Durgante FM, Higuchi N, Almeida A, Vicentini A. 2013. Species spectral signature: Discriminating closely related plant species in the Amazon with near-infrared leaf-spectroscopy. Forest Ecology and Management 291: 240-248.

Endara MJ, Coley PD, Wiggins NL, *et al*. 2018. Chemocoding as an identification tool where morphological-and DNA-based methods fall short: Inga as a case study. New Phytologist 218: 847-858.

Esteves SM, Vicentini A. 2013. Cryptic species in Pagamea coriacea sensu lato (Rubiaceae): evidence from morphology, ecology and reproductive behavior in a sympatric context. Acta Amazonica 43: 415-428.

Fernández M, Ezcurra C, Calviño CI. 2017. Species limits and morphometric and environmental variation within the South Andean and Patagonian Mulinum spinosum species-group (Apiaceae-Azorelloideae). Systematics and Biodiversity 15: 1-17.

Ferreira CAC. 2009. Análise comparativa de vegetação lenhosa do ecossistema campina na Amazônia Brasileira. PhD Thesis, Instituto Nacional de Pesquisas da Amazônia - INPA, Manaus.

Flora do Brasil 2020. 2020. Instituto de Pesquisas Jardim Botânico do Rio de Janeiro. 17 Aug. 2019.

Gower JC. 1971. A general coefficient of similarity and some of its properties. Biometrics 27: 857-871.

Grube M, Kroken S. 2000. Molecular approaches and the concept of species and species complexes in lichenized fungi. Mycology Research 104: 1284-1294.

Henderson A, Martins R. 2002. Classification of specimens in the *Geonoma stricta* (Palmae) complex: The problem of leaf size and shape. Brittonia 54: 202-212.

Henderson A. 2006. Traditional morphometrics in plant systematics and its role in palm systematics. Botanical Journal of the Linnean Society 151: 103-111.

Henderson A. 2011. A revision of Geonoma (Arecaceae). Phytotaxa 17: 1-271.

Holanda ASS, Vicentini A, Hopkins MJG, Zartman CE. 2015. Phenotypic differences are not explained by pre-zygotic reproductive barriers in sympatric varieties of the *Humiria balsamifera* Aubl. (Humiriaceae) complex. Plant Systematics and Evolution 301: 1-13.

Hoorn C, Wesselingh FP, Steege H, *et al*. 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. Science 330: 927-931.

Hopkins MJG. 2007. Modelling the known and unknown plant biodiversity of the Amazon Basin. Journal of Biogeography 34: 1400-1411.

Hopkins MJG. 2019. Are we close to knowing the plant diversity of the Amazon? Anais da Academia Brasileira de Ciências 9. doi: 10.1590/0001-3765201920190396

Jacobs SJ, Herzog S, Tank DC. 2019. Quantifying morphological variation in the Castilleja pilosa species complex (Orobanchaceae). PeerJ 7: e7090. doi: 10.7717/peerj.7090

Knudsen JT. 2002. Variation in floral scent composition within and between populations of Geonoma macrostachys (Arecaceae) in the Western Amazon. American Journal of Botany 89: 1772-1778.

Lang C, Costa FRC, Camargo JLC, Durgante FM, Vicentini A. 2015. Near Infrared Spectroscopy facilitates rapid identification of both young and mature Amazonian tree species. PLOS ONE 10: e0134521. doi: 10.1371/journal-.pone.0134521

Lang C, Almeida DR, Costa FR. 2017. Discrimination of taxonomic identity at species, genus and family levels using Fourier Transformed Near-Infrared Spectroscopy (FT-NIR). Forest Ecology and Management 406: 219-227.

Li Yuan-Cong, Wen J, Ren Y, Jian-Qiang Z. 2019. From seven to three: Integrative species delimitation supports major reduction in species number in Rhodiola section Trifida (Crassulaceae) on the Qinghai-Tibetan Plateau. Taxon 68: 268-279.

Mace GM. 2004. The role taxonomy in species conservation. Philosophical Transactions Royal Society London B 359: 711-719.

Manly BFJ. 1997. Randomization, Bootstrap and Monte Carlo methods in biology. London, Chapman & Hall.

Mayr E. 1982. The growth of biological thought: Diversity, evolution and inheritance. Cambridge, MA, Harvard University Press.

McDade LA. 1995. Species Concepts and Problems in Practice: Insight from Botanical Monographs. Systematic Botany 20: 606-22.

Michener CD, Sokal RR. 1957: A quantitative approach to a problem in classification. Evolution 11: 130-162.

Medeiros H, Holanda ASS, Amorim AMA. 2015. Humiriaceae in: Lista de Espécies da Flora do Brasil. Jardim Botânico do Rio de Janeiro. http://floradobrasil.jbrj.gov.br/jabot/floradobrasil/FB129. 7 Feb. 2020

Padial JM, Miralles A, De la Riva I, Vences M. 2010. The integrative future of taxonomy. Frontiers in Zoology 7: 16-30.

Pessoa EM, Alves M, Alves-Araújo A, Palma-Silva C, Pinheiro F. 2012. Integrating different tools to disentangle species complexes: A case study in Epidendrum (Orchidaceae). Taxon 61: 721-734.

Pinheiro F, Barros F. 2007. *Epidendrum secundum* Jacq. and *E. denticulatum* Barb. Rodr. (Orchidaceae): useful characters for their recognition. Hoehnea 34: 563-570.

Pinheiro F, Barros F. 2009. Morphometric analysis of the Brasiliorchis picta complex (Orchidaceae). Revista Brasileira de Botânica 32: 11-21.

Pinheiro F, Dantas-Queiroz MV, Palma-Silva C. 2018. Plant species complexes as models to understand speciation and evolution: a review of South American studies. Critical Reviews in Plant Sciences 37: 54-80.

Prance GT. 1996. Islands in Amazonia. Phil. Trans. Of the Royal Society of London B Biological Science 351: 823-833.

Prance GT, Daly D. 1989. Brazilian Amazon. In: Campbell DC, Hammond HD. (eds.) Floristic inventory of tropical countries. Bronx, NY, New York Botanical Garden. p. 523-533.

Prata BEM, Sass C, Rodrigues DP, *et al*. 2018. Towards integrative taxonomy in Neotropical botany: disentangling the Pagamea guianensis species complex (Rubiaceae). Botanical Journal of the Linnean Society 188: 213-231.

QGIS - Quantum GIS Development Team. 2019. QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org. 19 Mar. 2019

Rabelo PM. 2016. Complexo Epidendrum secundum como modelo de estudo multidisciplinar na delimitação de espécies. MSc Thesis, Universidade Estadual Paulista "Júlio de Mesquita Filho",, Rio Claro.

Raxworthy CJ, Ingram CM, Rabibisoa Ni, Pearson RG. 2007. Applications of Ecological Niche Modeling for Species Delimitation: A Review and Empirical Evaluation Using Day Geckos (Phelsuma) from Madagascar. Systematic Biology 56: 907-923.

R Development Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/. 8 Jun. 2019

Rutherford S. 2020. Insights into speciation and species delimitation of closely related eucalypts using an interdisciplinary approach. Australian Systematic Botany 33: 110-127.

Saraiva DP, Mantovani A, Forzza RC. 2015. Insights into the Evolution of Pitcairnia (Pitcairnioideae-Bromeliaceae), based on Morphological Evidence. Systematic Botany 40: 726-36.

Schlick-Steiner BC, Steiner FM, Seifert B, Stauffer C, Christian E, Crozier RH. 2010. Integrative taxonomy: A multisource approach to exploring biodiversity. Annual Review of Entomology 55: 421-438.

Sites JW, Marshall JC. 2003. Delimiting species: a Renaissance issue in systematic biology. Trends in Ecology and Evolution 18: 462-470.

Steege H, Vaessen RW, Cárdenas-López D, *et al.* 2016. The Discovery of the Amazonian tree flora with an updated checklist of all known tree taxa. Scientific Reports 6: 29549. doi: 10.1038/srep29549

Steege H, Oliveira SM, Pitman NCA, *et al.* 2019. Towards a dynamic list of Amazonian tree species. Scientific Reports 9: 3501. doi:10.1038/s41598-019-40101-y

Thorne RF. 1972. Major Disjunctions in the Geographic Ranges of Seed Plants. The Quarterly Review of Biology 47: 365-411.

Trofimov D, Rudolph B, Rohwer JG. 2016. Phylogenetic study of the genus Nectandra (Lauraceae), and reinstatement of Damburneya. Taxon 65: 980-996.

Urban I. 1877. Humiriaceae. Flora Brasiliensis, Martius XII 2: 425-454.

Wang X, Edwards RL, Auler AS, *et al*. 2017. Hydroclimate changes across the Amazon lowlands over the past 45,000 years. Nature 541: 204-207.

White F. 1962. Geographic variation and speciation in Africa with particular reference to Diospyros, Ebenaceae. Systematics Association 4: 71-103.

White F. 1998. The vegetative structure of African Ebenaceae and the evolution of rheophytes and ring species. In: Fortune-Hopkins HE, Huxley CR, Pannell M, Prance GT, White F. (eds.) The Biological Monograph: The importance of field studies and functional syndromes for taxonomy and evolution of tropical plants. London, Kew, Royal Botanic Gardens. p. 95-113.

Yang M, Chen X, Huo W, Wei C. 2014. Morphological variation versus genetic divergence: a taxonomic implication for Mogannia species (Cicadidae: Cicadinae). Systematics and Biodiversity 12: 456-472.

Zapata F, Jiménez I. 2011. Species Delimitation: Inferring Gaps in Morphology across Geography. Systematic Biology 61: 179-94.