Eduardo Koerich Nery

AN INTEGRATED TAXONOMIC APPROACH TO THE *Hydrocotyle stella* Pohl ex DC. (ARALIACEAE) COMPLEX FROM THE BRAZILIAN ATLANTIC FOREST

Dissertação submetida ao Programa de Pós-Graduação de Biologia de Fungos, Algas e Plantas da Universidade Federal de Santa Catarina para a obtenção do Grau de Mestre em Biologia de Fungos, Algas e Plantas

Orientador: Prof. Dr. Pedro Fiaschi
Coorientador: Profa. Dra. Mayara Krasinski Caddah

Florianópolis
2019

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Nery, Eduardo
An integrated taxonomic approach to the Hydrocotyle stella Pohl ex DC. (Araliaceae) complex from the Brazilian Atlantic forest / Eduardo Nery ; orientador, Pedro Fiaschi, coorientador, Mayara Caddah, 2019.
50 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro de Ciências Biológicas, Programa de Pós-Graduação em Biologia de Fungos, Algas e Plantas, Florianópolis, 2019.

Inclui referências.

1. Biologia de Fungos, Algas e Plantas. 2. Filogenia. 3. Genética populacional. 4. Morfometria. 5. Taxonomia. I. Fiaschi, Pedro. II. Caddah, Mayara. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Biologia de Fungos, Algas e Plantas. IV. Título.

Eduardo Koerich Nery

AN INTEGRATED TAXONOMIC APPROACH TO THE *Hydrocotyle stella* Pohl ex DC. (ARALIACEAE) COMPLEX FROM THE BRAZILIAN ATLANTIC FOREST

Esta Dissertação foi julgada adequada para obtenção do Título de "Mestre em Biologia de Fungos, Algas e Plantas" e aprovada em sua forma final pelo Programa de Pós-Graduação em Biologia de Fungos, Algas e Plantas

Florianópolis, 22 de março de 2019

_____
Prof. Dr. Pedro Fiaschi
Coordenador do Curso

Banca Examinadora:

_____
Prof. Pedro Fiaschi, Dr.
Orientador
Universidade Federal de Santa Catarina

_____
Prof. Carlos Frederico Deluqui Gurgel, Dr.
Universidade Federal de Santa Catarina

_____
Profa. Samantha Koehler, Dra.
Universidade Estadual de Campinas

Este trabalho é dedicado aos amigos e família, pois são eles que dão sentido à vida.

RESUMO

As espécies são cruciais para a investigação biológica e outros campos além da pesquisa científica, mas os diferentes conceitos de espécies e seus respectivos critérios tornam a delimitação de espécies uma tarefa desafiadora. O conceito de espécie como linhagem geral (GLS) pode superar os problemas de delimitação de espécies ao aceitar os diferentes critérios de delimitação como linhas de evidência para o reconhecimento de espécies. *Hydrocotyle* compreende ervas anuais e perenes que têm chamado atenção devido ao seu impacto ecológico e potencial aplicação farmacológica, mas a delimitação de espécies de *Hydrocotyle* tem sido problemática por meios taxonômicos tradicionais, resultando em diferentes complexos de espécies. *Hydrocotyle stella* apresenta seis morfotipos que ocorrem ao longo de cadeias montanhosas dentro do domínio da Mata Atlântica brasileira, de modo que esses morfotipos poderiam ser linhagens ainda não reconhecidas. Adotando o GLS, nós avaliamos se os morfotipos de *H. stella* seriam ou não espécies, considerando o monofiletismo, a estruturação genética e a descontinuidade morfológica como evidências para o reconhecimento de espécies. Para tal, aplicamos inferência filogenética, genética populacional e morfometria a 12 populações de *H. stella* representando os seis morfotipos e as quatro cadeias de montanhas de principal ocorrência do complexo. A inferência filogenética baseada em ITS retornou um único grupo monofilético, composto por dois morfotipos diferentes, mas a topologia das árvores apresentou resolução baixa, tornando o monofiletismo inconclusivo. A genética populacional baseada em marcadores ISSR indicou que a maioria dos morfotipos não era geneticamente estruturada, mas estes compunham três grupos que apresentavam estrutura genética associada a regiões geográficas. A morfometria descartou a maioria dos morfotipos como morfologicamente descontínuos, mas indicou limites morfológicos consistentes entre os três grupos reconhecidos na genética populacional. Com base nisso, sugerimos uma reavaliação geral dos táxons infraespecíficos dentro do *Hydrocotyle*. Acima de tudo, sugerimos o reconhecimento de três espécies com base nos grupos sustentados pela genética populacional e morfometria, reduzindo a circunscrição de *H. Stella*.

Palavras-chave: Delimitação de espécies. Filogenia. Genética populacional. *Hydrocotyle*. Morfometria.

RESUMO EXPANDIDO

Introdução

As espécies são a base da pesquisa biológica, e além disso elas são uma referência para a exploração de recursos, gestão ambiental e desenvolvimento tecnológico. Contudo, o reconhecimento e a circunscrição das espécies (i.e. delimitação de espécies) ainda é um tarefa desafiadora visto que a mesma pode seguir seguir diferentes critérios que atendem a uma ampla gama de conceitos de espécies. O conceito de espécie como linhagem geral (GLS) pode trazer maior estabilidade à delimitação das espécies ao conciliar os diferentes conceitos de espécie. O GLS considera que as espécies seriam linhagens de metapopulações ou de segmentos de metapopulações e que os diferentes critérios de delimitação seriam linhas de evidência abordando diferentes propriedades que podem ou não aparecer durante a especiação. Entre as propriedades que podem surgir durante a especiação estão o monofiletismo, a estruturação genética e a descontinuidade morfológica. O gênero *Hydrocotyle* L. compreende ervas que frequentemente ocupam ambientes úmidos, sendo potenciais invasoras nestes mesmos ambientes, e algumas espécies deste gênero são reconhecidos componentes da medicina popular. Todavia, a diversidade do gênero é possivelmente subestimada devido a diferentes problemas de delimitação de espécie. O complexo *Hydrocotyle stella* Pohl ex DC. apresenta seis morfotipos foliares que estão distribuídos ao longo de cadeias montanhosas dentro do domínio da Mata Atlântica brasileira. Formas intermediárias entre morfotipos sugerem que *H. stella* seja uma única espécie polimórfica, mas a distribuição ao longo de cadeias montanhosas sugere uma possível interrupção do fluxo gênico que pode esconder espécies ainda não reconhecidas.

Objetivos

Este trabalho objetivou avaliar se *H. stella* compreenderia diferentes espécies ou não. Para tal, foi adotado o GLS como fundamento teórico, considerando o monofiletismo, a estruturação genética e a descontinuidade morfológica como evidências para o delimitação de espécies.

Material e Métodos

A amostra incluiu 12 populações (186 espécimes) de *H. stella*, representando todos os morfotipos foliares do complexo e as principais regiões montanhosas ocupadas pelo mesmo. O monofiletismo foi

avaliado por inferência filogenética sobre a região nuclear ITS, sendo que cada população de *H. stella* foi representada por um indivíduo. A estruturação genética foi avaliada com base em 94 *loci* de ISSR que amplificados a partir dos espécimes de *H. stella*. Os *loci* de ISSR foram submetidos a análises baeysianas de agrupamento (STRUCTURE), a uma análise de componentes principais (PCA), e a uma análise de variância molecular (AMOVA). A descontinudiade morfológica foi avaliada por morfometria de dez caracteres medidos nos espécimes de *H. stella*. As variáveis morfométricas foram submetidas a uma análise de variáveis canônicas (CVA) e a uma análise de permutação sobre a distância da Mahalanobis (D^2).

Resultados
Ambos métodos de inferência filogenética retornaram árvores com baixa resoluação, tornando o monofiletismo de *H. stella* e seus morfotipos inconclusivo. As análises de STRUCUTRE indicaram considerável compartilhamento genético entre a maioria dos morfotipos de *H. stella*. A PCA indicou três grupos geneticamente estruturados dentro de *H. stella*, cada um associado a diferentes regiões geográficas. A AMOVA indicou ausência de estruturação genéticas entre a maioria dos morfotipos, mas esta indicou alta e siginificativa estruturação entre os grupos identificados pela PCA. A CVA e a permutação de D^2 indicaram que a maioria dos morfotipos não estavam separados por descontinuidades morfológicas, mas os mesmos indicaram descontinuidades entre os três grupos indentificados pelas análises genéticas.

Discussão
Os morfotipos de *H. stella* não são espécies, visto que a maioria deles não apresentou monofiletismo, estruturação genética ou descontinuidade morfológica. Contudo, o complexo *H. stella* compreende três grupos geneticamente estruturados e morfologicamente descontínuos que ocupam diferentes regiões geográficas, logo estes grupos são espécies.

Considerações finais
Um tratamento taxonômico é necessário para reduzir a circunscrição de *H. stella* e reconhecer outras duas espécies de *Hydrocotyle* na Mata Atlântica brasileira.

Palavras-chave: Delimitação de espécies. Filogenia. Genética populacional. *Hydrocotyle*. Morfometria.

ABSTRACT

Species are crucial to biological inquiry and other fields beyond scientific research, but the different species concepts and their respective criteria turn species delimitation a challenging task. The general lineage species concept (GLS) can overcome problems of species delimitation by accepting the different delimitation criteria as lines of evidence for species recognition. *Hydrocotyle* comprises annual and perennial herbs that have attained attention due to their ecological impact and potential pharmacological application, but delimitation of *Hydrocotyle* species has been problematic by traditional taxonomic means, resulting in different species complexes. *Hydrocotyle stella* displays six morphotypes that occur along mountain ranges within the Brazilian Atlantic forest domain, so such morphotypes could be underrated lineages. We adopted the GLS to assess whether *H. stella* morphotypes would be species or not, considering monophyly, genetic structure and morphological discontinuity as evidences for species recognition. For such, we applied phylogenetic inference, population genetics, and morphometrics to 12 populations of *H. stella* representing the six morphotypes and four mountain ranges of main occurrence. Phylogenetic inference based on ITS returned a single monophyletic group, which comprised two different morphotypes, but trees had overall low resolution, making monophyly inconclusive. Population genetics based on ISSR markers indicated that most morphotypes were not genetically structured, but they were rather nested within three groups displaying genetic structure, which was associated to geographic regions. Morphometrics dismissed most morphotypes as morphologically discontinuous, but it indicated consistent morphological boundaries among the three groups recognized in population genetics. Based on that, we suggested an overall revaluation of infraspecific taxa within *Hydrocotyle.* Foremost, we suggested the recognition of three species based on groups supported by population genetics and morphometrics, reducing *H. stella* circumscription.

Keywords: *Hydrocotyle*. Morphometrics. Population genetics. Phylogeny. Species delimitation.

SUMMARY

1 INTRODUCTION

1.1 THEORETICAL BACKGROUD

Species are at the core of evolutionary theory (Hull, 1977), and not surprisingly they are the ground upon which biological inquiry lays. Beyond that, species are a main reference for resource exploitation, environmental management, and technological development. Therefore, the recognition and circumscription of species (i.e. species delimitation) is a major goal within Systematics (Wiens, 2007). However, species delimitation can be a challenging task since it can follow different criteria that comply with a wide range of species concepts (Mayden, 1997). Each species concept bears either theoretical or operational limitations (Luckow, 1995; Balakrishnan, 2005), so their respective delimitation criteria can lead to different taxonomic proposals for a same group of organisms (Peterson & Navarro-Sigüenza, 1999; Rheindt & Eaton, 2009). Such inconsistency among delimitation criteria has casted doubts on species as the basis for conservation efforts (Hey *et al.*, 2003; Agapow *et al.*, 2004; Isaac, Mallet, & Mace, 2004; Zachos, 2018), which are a pressing need in the current world-wide scenario of declining diversity. Moreover, due to their potential impact on the number of endangered species, species delimitation has been subject of harsh critics that demand judicialization of species recognition (Garnett & Christidis, 2017). Hence, overcoming the inconsistency among delimitation criteria is a necessary step for the maintenance of species in biological research and other fields.

The general lineage species concept (de Queiroz, 2007) can conciliate species concepts and their delimitation criteria. In short, de Queiroz (2007) proposes that species concepts (the theoretical aspect) are in agreement by acknowledging species as separately evolving metapopulation lineages, but their respective delimitation criteria (the operational aspect) are in disagreement by assessing different properties that lineages may or may not evolve along speciation. Hence, each delimitation criterion can provide an evidence of lineage separation, and the amount of evidence, rather than a particular sort of evidence, would define whether a species should be recognized or not (de Queiroz, 2007). Under this comprehensive perspective, integrative taxonomy by congruence is a sounding approach of species delimitation (Padial *et al.*, 2010). This approach consists of an initial application of different methods and a posterior search for concordant patterns across their results (Padial *et al.*, 2010). By doing so, species delimitation becomes more accurate (i.e. reduced systematic error) (Carstens *et al.*, 2013), as single-

method approaches are inherently biased (Miralles & Vences, 2013; Luo *et al.*, 2018). Foremost, like other integrative approaches, integration by congruence can cover different moments of the speciation continuum (de Queiroz, 2007).

Phylogenetic inference based on coalescence of alleles can recognize species by monophyly. Current practice of phylogenetic inference assumes that hypotheses on coalescence of alleles (gene trees) are an approximation to hypotheses on evolutionary relationships among species (species trees) (Goodman, Czelusniak, & Moore, 1979). Under this assumption, if alleles coalesce more recently within a group of organisms rather than outside of such a group, organisms within the group are more closely related to each other than to organisms outside the group (Baum & Shaw, 1995). Since relatedness is thought to be greater within lineages rather than among lineages, species would be groups of organisms that are more closely related among each other than to organism of other groups and that bear no other such groups within them (genealogy-based phylogenetic species [Baum & Shaw, 1995]). In practical terms, species would display recent common ancestry (i.e. monophyly) at different genetic loci (Hudson, Coyne, & Url, 2002). However, gene trees may not always approximate species trees as alleles may undergo processes that misrepresent (or even hide) lineage separation (e.g. incomplete lineage sorting, horizontal gene transfer, gene duplication/extinction, and recombination) (Degnan & Rosenberg, 2009). On the other hand, species may undergo reticulate evolution (e.g. introgression and hybridization) that do not fit in the strict hierarchical topology of gene trees (Linder & Rieseberg, 2004). Such processes, as well as other deviations from classic allopatric speciation, can render species non-monophyletic (Rieseberg & Brouillet, 1994; Funk & Omland, 2003).

Inter-simple sequence repeat (ISSR) is a dominant, multi-locus molecular marker that can recognize species by genetic structure. ISSR derives from DNA amplifications of genomic regions between near, inversely oriented, and same-motif microsatellites (SSR) (Zietkiewicz, Rafalski, & Labuda, 1994), and each ISSR amplicon is assumed as an allele of an independent neutral locus whose homology among samples is inferred by molecular weight (Bussell, Waycott, & Chappill, 2005). Therefore, ISSR data provide means to characterize genotypes and, consequently, to estimate breeding among samples (Sanz *et al.*, 2009). Assuming that outbreeding reduces or ceases with lineage separation, species would be groups of organisms displaying few or none heterozygotes among each other (genetic-cluster species [Mallet, 1995]).

This lack of heterozygotes makes allele (or genotype) frequencies of species to deviate from expectations for a single panmitic group (i.e. genetic structure) (Wright, 1949), and such genetic structure becomes detectable by ISSR loci via population genetics analysis (Hausdorf & Hennig, 2010). Nonetheless, same-weight ISSR amplicons may not always represent a same allele since homology among samples holds mostly at lower taxonomic levels (Wolfe, Randle, & Wilson, 2001). Additionally, due to its dominant inheritance, ISSR may produce biased estimates of allele frequencies if recessive homozygotes are rare within a sample (Zhivotovsky, 1999). Aside of issues regarding markers, genetic structure also takes place among populations that are not undergoing lineage separation, and such populations are quantitatively undistinguishable from species by different fixation indexes (Hey & Pinho, 2012).

Morphometrics analyzes morphological variation and covariation and is an alternative to validate species by morphological discontinuity. Morphometric methods rely on variables that describe morphological variation, and such variables can result from either linear and angular measurements (traditional morphometrics), or landmarks and outlines of structures (geometric morphometrics) (Rohlf & Marcus, 1993; Adams, Rohlf, & Slice, 2004, 2013). Once obtained, morphometric variables can be subject to different statistical procedures (e.g. clustering techniques, ordinations, and statistical tests) that describe morphological variation (Henderson, 2005). Assuming that morphology diverges after genetic isolation, species would be groups of individuals that are separated from each other by different discontinuities (morphological species [Du Rietz, 1930]). Such discontinuities produce gaps on ordination plots and statistically supported differences in classic parametric tests, so they become detectable by morphometric methods (Henderson, 2005). However, usual analyses applied to morphometric variables may not always be sensible to non-overlapping variation (i.e. discontinuity) (Cadena, Zapata, & Jiménez, 2018). Besides analytical issues, morphological discontinuities may not reflect lineage separation as selection-driven processes can decouple them. Local adaptations can increase morphological variation among populations that are genetically connected (Sambatti, Rice, & Ambatti, 2006; Gonzalo-Turpin & Hazard, 2009). On the other hand, stabilizing selection can maintain low morphological variation between sister-lineages (Grundt *et al.*, 2006).

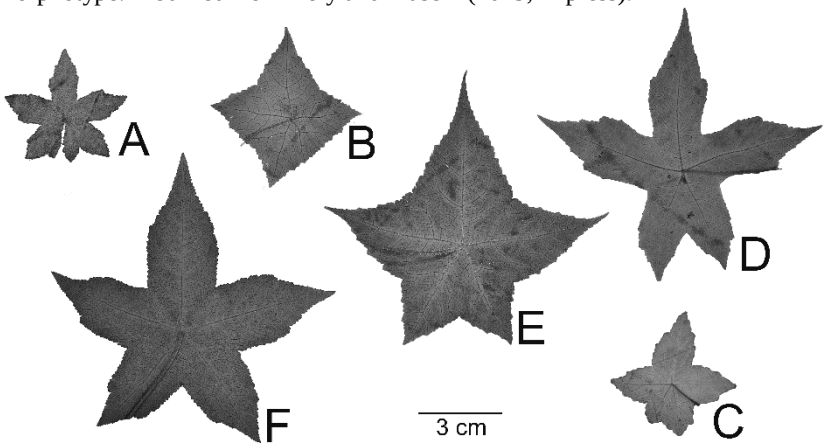
Species delimitation within *Hydrocotyle* L. is troublesome and possibly underestimates species diversity. *Hydrocotyle* composes the sister group to all other araliaceous lineage (Nicolas & Plunkett, 2009),

and it comprises annual and perennial herbs with stoloniferous or prostate stems, simple to compound leaves, lateral stipules, umbellate or verticillate inflorescences, bisexual flowers, inferior bicarpellate ovaries, and laterally flattened fruits. These plants are often associated to mesic and aquatic environments, where they may exhibit an invasive behavior (Ruiz-Avila & Klemm, 1996; Liu *et al.*, 2016), and some of them are known components of folk medicine (Rocha *et al.*, 2011; Huang *et al.*, 2013). *Hydrocotyle* diversity is thought to be greater in South America and Australia (Nicolas & Plunkett, 2014), but accurate estimates are still unavailable as species delimitation within the genus possibly underestimates species diversity. Traditionally, *Hydrocotyle* species are recognized based on leaf morphology (Constance & Dillon, 1990; Mendoza, 2010; Henwood, 2014), but overlapping variation at leaf characters can make species delimitation challenging by traditional taxonomic means. As a consequence, taxonomic treatments to date have favored the recognition of infraspecific taxa over species (Johnson & Webb, 1982), a practice that has led to circumscription of polymorphic species with dubious status (i.e. species complexes). Most of these complexes remain unverified (Eichler, 1987a,b,c), as the last comprehensive revision of the genus dates back to 19th century (Richard, 1820). Hence, world-wide estimates of *Hydrocotyle* diversity are uncertain (Mathias & Constance, 1962; Pimenov & Leonov, 1993), while different species complexes require clearer internal delimitations.

Hydrocotyle stella Pohl ex DC. sensu Nery and Fiaschi (2019, in press) is a species complex with a wide distribution along the Brazilian Atlantic forest. De Candolle (1830) first proposed *H. stella* to designate plants collected in Brazil that had villous to hirsute stems, five-lobed leaves, unequal and widely lanceolate lobes, and double serrated margins. In his treatment for Flora Brasiliensis, Urban (1879) synonymized *H. stella* under *Hydrocotyle quinqueloba* var. *stella* (Pohl ex DC.) Urb., one of the ten infraspecific taxa proposed for *Hydrocotyle quinqueloba* Ruiz & Pav. Based on morphometric analyses, Nery and Fiaschi (2019, in press) reduced *H. quinqueloba* circumscription to plants collected in Peru and elevated four groups of plants collected in Brazil to species. One such group included *H. stella* and five infraspecific taxa of the former *H. quinqueloba* sensu Urban (1879), forming the *H. stella* complex. As a consequence of this treatment, *H. stella* came to comprise six morphologies (Fig. 1): the palacea morphotype (PALAC, Fig. 1A), with villous body, five-lobed leaves, lanceolate lobes, and a deeper hear sinus; the quadrata morphotype (QADRA, Fig. 1B), with glabrous body, four-lobed leaves, and widely ovate-triangular lobes; the quadriloba

morphotype (QADRI, Fig. 1C), with glabrescent body, four-lobed leaves, and ovate-lanceolate lobes; the quinqueradiata morphotype (QINQE, Fig. 1D), with pubescent to villous body, five-lobed leaves, and lanceolate lobes; the stella morphotype (STELL Fig. 1E), with villous to hirsute body, five-lobed leaves, and widely triangular lobes; and the subglabra morphotype (SUBG, Fig. 1F), with glabrous to glabrescent body, five-lobed leaves, and ovate-lanceolate lobes. These plants grow under forest understories and over grasslands within the Brazilian Atlantic Forest domain, and their occurrence localities are mainly at the Espinhaço, Serra da Mantiqueira, Serra do Mar, and Serra Geral mountain ranges. This wide geographic distribution along mountain ranges suggests a possible lack of genetic connectivity within the complex, which may foster different lineages (Nery and Fiaschi 2019, in press). On the other hand, intermediate morphologies among some morphotypes suggest some genetic exchange (Nery and Fiaschi 2019, in press), leaving morphological variation to be mainly due to local conditions.

Figure 1 - *Hydroctyle stella* complex and its morphotypes: **A.** the palacea morphotype; **B.** the quadrata morphotype; **C.** the quadriloba morphotype; **D.** the quinqueradiata morphotype; **E.** the stella morphotype; and **F.** the subglabra morphotype. Modified from Nery and Fiaschi (2019, in press).



This study aimed to assess whether the *H. stella* complex comprises different species or not. For such, we adopted the general lineage species concept (de Queiroz, 2007), and we considered monophyly, genetic structure, and morphological discontinuity as evidences of lineage separation.

1.2 AIM AND OBJECTIVES

1.2.1 Aim

To set species boundaries within the *H. stella* complex from the Brazilian Atlantic forest domain.

1.2.2 Objetives

- To assess monophyly within the *H. stella* complex via phylogenetic inference based on DNA sequencing;
- To assess genetic structure within the *H. stella* complex via population genetics based on ISSR molecular markers;
- To test morphological discontinuity within the *H. stella* complex via morphometrics;
- To provide taxonomic advice regarding species boundaries within the *H. stella* complex based on phylogenetic inference, population genetics, and morphometrics.

2 MATERIAL AND METHODS

2.1 SAMPLING AND DNA EXTRACTION

We sampled 12 populations (186 specimens) of the *H. stella* complex within the Brazilian Atlantic forest domain (Fig. 2), including the six morphotypes and the four mountain ranges where the complex is mostly found (Table 1). Each morphotype had at least two populations sampled, except for PALAC and QADRI. PALAC's records are from two localities, the Dois Irmãos and the Itambé peaks, and we succeed to find this morphotype only at the second locality. QADRI's records are scarse in herbarium collection, and we succeed to find this morphotype in a small forest fragment at the Bom Sucesso do Itararé municipality.

Each population comprised 11 to 20 specimens, stored under a same voucher number (Table 1). We defined as different specimens those samples that were at least ten meters apart in order to minimize resampling of clones. We stored leaf tissue of all specimens in silica gel for DNA extraction and molecular analyses, and we preserved all specimens for morphometrics. DNA extraction followed a modified CTAB protocol (Doyle & Doyle, 1987).

Figure 2 – Sampled populations of *H. stella* within the Brazilian Atlantic forest domain in South America. Gray scale indicates altitude in meters.

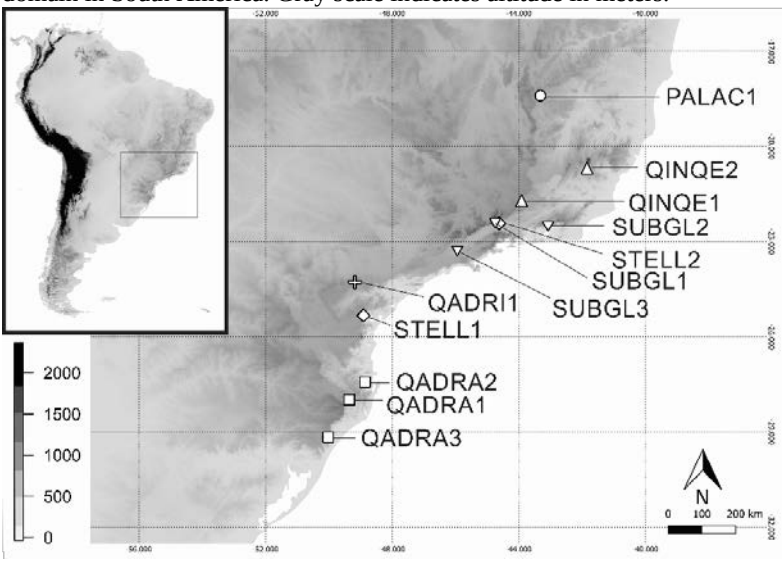


Table 1 - Sampled populations and their respective mountain range of occurrence, geographic coordinates, associated herbarium voucher, and number of collected specimens. Within parenthesis, the number of fertile specimens analyzed in morphometrics.

Population	Mountain range	Geographic coordinates	Voucher	n
PALAC1	Espinhaço	18°24'08.0"S; 43°18'53.1"W	Nery 75	16 (11)
QADRA1	Serra Geral	28°01'11.2"S; 49°17'56.0"W	Nery 57	20 (0)
QADRA2	Serra Geral	27°25'42.6"S; 48°50'54.9"W	Nery 92	18 (9)
QADRA3	Serra Geral	29°10'63.2"S; 50°01'64,3"W	Nery 100	11 (0)
QADRI1	Serra do Mar	24°17'03.0"S; 49°10'17.2"W	Nery 90	13 (8)
QINQE1	Mantiqueira	21°42'41.3"S; 43°54'33.9"W	Nery 60	12 (4)
QINQE2	Mantiqueira	20°41'12.0"S; 41°50'28.0" W	Nery 81	19 (3)
STELL1	Serra do Mar	25°19'54.0"S; 48°54'01.0"W	Nery 29	18 (13)
STELL2	Mantiqueira	22°26'07.9"S 44°36'47.2"W	Nery 42	17 (10)
SUBGL1	Mantiqueira	22°22'24.6"S; 44°45'20.7"W	Nery 36	15 (5)
SUBGL2	Serra do Mar	22°31'47.7"S; 43°03'59.9"W	Nery 44	15 (8)
SUBGL3	Serra do Mar	23°14'47.0"S; 45°56'17.4"W	Nery 89	12 (6)
			TOTAL:	**186 (77)**

2.2 SEQUENCE OBTENTION AND EDITION

We chose the internal transcribed spacer (ITS) of nuclear ribosomal DNA for phylogenetic inference as this locus has displayed enough polymorphism to infer infra-generic relationships within other araliaceous lineages (Plunkett *et al.*, 2005; Tronchet *et al.*, 2005; Fiaschi & Plunkett, 2011; Plunkett & Lowry II, 2012; Li & Wen, 2013).

Our sample included 36 ITS sequences of *Hydrocotyle* and three sequences of *Trachymene* Rudge (Table 2). We obtained one ITS sequence for each sampled population of *H. stella* and of other 11 *Hydrocotyle* spp. occurring in the Brazilian Atlantic forest. We also obtained one ITS sequence for eight other *Hydrocotyle* spp. that did not occur within this domain. Our alignment included 11 out 19 *Hydrocotyle*

spp. recognized within the Brazilian Atlantic forest (Flora do Brasil 2020). We chose *Trachymene* as outgroup because this Australasian genus also composes the sister group to all other araliaceous lineages, like *Hydrocotyle* (Nicolas & Plunkett, 2009).

We obtained most sequences by amplification and sequencing at this study, and some sequences from GenBank (Benson *et al.*, 2005). Using the ITSu1 and ITSu4 primers described by Cheng et al. (2016), amplification reactions had a total volume of 20 ul, containing 200 uM of each dNTP, 1.5 mM of $MgCl_2$, 1.25 U of *Taq* DNA polymerase (TopTaq Master Mix®, Qiagen), 0,25 uM of each primer, and 20-50 ng of genomic DNA. Thermocycling had an initial denaturation of 94 °C for 4 min, 34 cycles of 95 °C for 30 s, 55 °C for 40 s, 72 °C for 1 min, and a final extension of 72°C for 10 min. We purified amplicons by precipitation in a saline solution (2.5 M of NaCl) of polyethylene glycol (PEG) at 20% (Cheng, Jia, & Ran, 2015). Automated DNA sequencing (both strands) was outsourced to Myleus Facility® (Belo Horizonte, BR). We generated consensus sequences and manually solved their ambiguities with Sequencher 5.4.6® (Gene Codes Corporation, Michigan, USA). Sequences were aligned with MEGA X (Kumar *et al.*, 2018).

2.3 ISSR AMPLIFICATION AND GENOTYPING

We amplified ISSR of all *H. stella* specimens (n=186) for population genetics analyses (Table1). We tested 14 SSR-targeting primers and selected only four due to their band polymorphism, $(AC)_7RG$, $(AG)_8TG$, $(CA)_6RY$, and $(CTC)_6T$. To ensure band reproducibility, we performed three amplification essays with each primer on a subsample containing one specimen from each population. Moreover, we adopted high annealing temperatures (see below) to minimize amplification of random and non-repeatable fragments.

Amplification reactions had a total volume of 10 ul, containing 200 uM of each dNTP, 1.5 mM of $MgCl_2$, 1.25 U of *Taq* DNA polymerase (TopTaq Master Mix®, Qiagen), 1 x Coral Load (TopTaq Master Mix®, Quiagen), 5 uM of primer, and 20-50 ng of genomic DNA. Thermocycling had an initial denaturation of 95 °C for 3 min, 32 cycles of 95 °C for 15 s, primer-specific annealing temperature for 40 s, 72 °C for 2 min, and a final extension of 72°C for 7 min. Annealing temperatures were 52 °C for $(AC)_7RG$ and 48 °C for other primers. We stained amplicons with 2 x Gel Loading Dye Blue (Sinapse Inc) and 1 x Gel Red (Biotium), then we separated products along a 100 bp molecular ladder via electrophoresis in 1.5% agarose gels and 1 x TBE buffer, at

100 V, for 3 hours. We photographed electrophoretic gels under UV light. Based on photos, we scored the presence (1) and absence (0) of each band (allele), assuming homology among same-weight bands (Wolfe *et al.*, 2001).

Table 2 – ITS sequences of *Hydrocotyle* spp. and *Trachymene* spp, their respective GenBank accession numbers and nucleotide lengths. Asterisk (*) indicates species occurring within the Brazilian Atlantic forest domain.

Species (population)	GenBank acession	Length (bp)
*H. barbarossa**	not available yet	571
*H. bonariensis**	AF077894.1	619
*H. bradei**	not available yet	623
H. conferta	GU447310.1	613
*H. exigua**	not available yet	624
*H. itatiaiensis**	not available yet	630
H. javanica	KY438928.1	619
*H. langsdorffii**	not available yet	623
*H. leucocephala**	not available yet	710
*H. macrophylla**	not available yet	627
H. maritima	JQ247227.1	611
H. mexicana	AF077893.1	616
H. montana	KX640968.1	683
*H. pusilla**	not available yet	622
H. ramiflora	JQ247228.1	610
H. sibthorpioides	JQ247229.1	610
H. stella (PALAC1)*	not available yet	625
H. stella (QADRA1)*	not available yet	683
H. stella (QADRA2)*	not available yet	628
H. stella (QADRA3)*	not available yet	680
H. stella (QADRI1)*	not available yet	684
H. stella (QINQE1)*	not available yet	682
H. stella (QINQE2)*	not available yet	681
H. stella (STELL1)*	not available yet	620
H. stella (STELL2)*	not available yet	686
H. stella (SUBGL1)*	not available yet	683
H. stella (SUBGL2)*	not available yet	680
H. stella (SUBGL3)*	not available yet	680
*H. verticillata**	AY389025.1	621
H. vulgaris	KY968850.1	744
H. yabei	JQ425410.1	611
Trachymene incisa	AF272355.1	619
Trachymene sp. 1548	AF272353.1	608
Trachymene sp. 1552	AF272354.1	594

2.4 PHYLOGENETIC INFERENCE

We assessed monophyly at ITS by maximum likelihood (ML) and Bayesian inference (BI) criteria. We performed ML and BI analyses with RAxML-HPC2 and MrBayes on XSEDE 3.2.6, respectively, available in the CIPRES Science Gateway v. 3.3 (Miller, Pfeiffer, & Schwartz, 2010). ML analysis adopted the GTR+G model of base substitution and other parameters by default, using 1000 restarts. Node support was estimated by 1000 bootstrap replicates. BI analysis had two independent runs, each run with four chains of 20,000,000 Markov Chain Monte Carlo (MCMC) searches, with a burn-in period of 2,500,000. Chains sampled one tree every 1000 searches, and we built a 50% majority-rule consensus tree based on sampled trees. Node support was based on Bayesian posterior probability. For both ML and BI trees, we disconsidered nodes with less than 75% of support. We visualized and edited trees with the package FigTree v.1.3.1 from BEAST (Drummond & Rambaut, 2007).

2.5 POPULATION GENETICS

We assessed genetic structure within ISSR data via STRUCTURE 2.3.4 (Pritchard, Stephens, & Donnelly, 2000). Adopting a Bayesian framework, STRUCTURE infers the probability of a number of genetic clusters (K) and the relative contribution of clusters to each genotype (ancestry coefficient, Q) (Pritchard *et al.*, 2000). For such, STRUCTURE assigns specimens to clusters in order to maximize Hardy-Weinberg equilibrium (HWE) within clusters and to maximize linkage equilibrium among loci (Pritchard *et al.*, 2000). Among STRUCTURE models, we chose the admixture since intermediate morphologies suggest a possible genetic flow among morphotypes, and we chose the correlated allele frequency to avoid treating linked loci as independent variables. We implemented Falush et al (2007) parameter (X*) to handle the uncertainty of inferring heterozygotes via dominant markers. We inferred probabilities for K values ranging from 1 to 12, each value had 10 iterations with 50,000 MCMC searches after a 10,000 burn-in period. We chose the optimal K value according to delta K index (Evanno, Regnaut, & Goudet, 2005), and we produced a consensus clustering scenario for the optimal K with CLUMPP (Jakobsson & Rosenberg, 2007). We assigned specimens to the cluster with the highest Q value.

We also assessed genetic structure within ISSR data by non-parametric analyses with the R packages adegenet (Jombart, Lyon, & Biome, 2008; Jombart & Ahmed, 2011), and poppr (Kamvar, Tabima, &

Grünwald, 2014; Kamvar, Brooks, & Grünwald, 2015). We performed a principal component analysis (PCA) (Pearson, 1901; Hotelling, 1933) to visually assess genetic structure by non-overlapping variation, retaining PCs by the broken-stick model (Frontier, 1976). Although PCA can detect genetic structure despite deviations from classic model assumptions (e.g. HWE) (Patterson, Price, & Reich, 2006), it does not provide a quantification and formal testing of genetic structure. Hence, we performed an analysis of molecular variance (AMOVA) (Excoffier, Smouse, & Quattro, 1992) to measure a fixation index (Φ_{ST}) between each pair of morphotype, considering population as a nested factor within morphotype. A permutation procedure with 10,000 iterations tested whether measured Φ_{ST} would rise by chance alone, applying the Bonferroni correction for multiple comparisons. Such test adopted a significance level (α) of 0.05. Based on congruencies among parametric and non-parametric analyses, we recognized genetically structured groups.

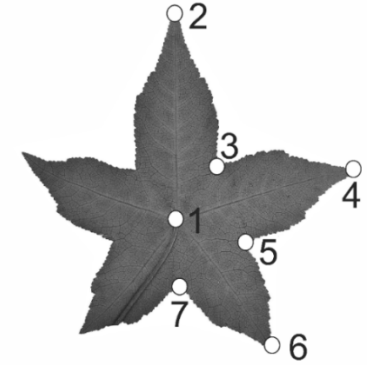
2.6 MORPHOMETRICS

We assessed morphological discontinuity by statistical analyses of morphometric variables. Leaf blade underwent geometric morphometrics to attain separate measures for blade shape and size, and other structures underwent traditional morphometrics since their variation were mostly size-related.

Geometric morphometrics included all sampled specimens (n=186) (Table1), and each specimen had three leaves sampled. Leaves were sampled before the second most distal node in order to avoid measuring variation due to growth. We photographed the abaxial surface of leaf blades next a ruler for scale with a Nikon 5100®, and we converted photos to the .TPS extension with tpsUtil (Rohlf, 2015). Using the tpsDig2 (Rohlf, 2015), we obtained raw coordinates of leaf blade by positioning seven landmarks (Fig. 3): (1) petiole insertion on leaf blade; (2) middle lobe apex; (3) middle lobe sinus; (4) lateral lobe apex; (5) lateral lobe sinus; (6) rear lobe apex; (7) rear lobe sinus. Landmarks covered only the right side of leaves to avoid redundancy due to bilateral symmetry. We analyzed raw coordinates with the R package geomorph (Adams & Otarola-Castillo, 2013). We performed a Procrustes superimposition (Rohlf & Slice, 1990) of the raw coordinates to separate leaf blade shape (aligned coordinates) from leaf blade size (centroid size), and we averaged such variables to specimens to avoid pseudoreplication and to minimize residual variance in further statistical analyses. Finally,

we performed a PCA of the aligned coordinates to summarize leaf shape variation, retaining PCs by the broken-stick model. Wireframe diagrams depicted changes of leaf shape by contrasting the consensus configuration (the great mean) with target configurations (Klingenberg, 2013).

Figure 3 – Landmarks on the abaxial surface of leaf blades of *H. stella*.



Traditional morphometrics included only fertile specimens (n=77) (Table 1), and each specimen had three measurements per character when feasible. We measured seven characters: internode length, petiole length, stipule length, stipule width, peduncle length, pedicel length, and number of flowers per umbel. Measurements occurred only on structures before the second most distal node. We averaged values to specimen to avoid pseudoreplication and to minimize residual variation.

Statistical analyses of morphometric variables utilized the R package Morpho (Schlager, 2017), and they considered two alternative circumscription scenarios: morphotypes and genetically structured groups from population genetics (see Discussion). First, we standardized variables to avoid character overweighting due to measurement scale. We performed a canonical variate analysis (CVA) (Campbell & Atchley, 1981) to assess which characters were discriminant and whether circumscriptions displayed non-overlapping variation. We measured morphological divergence with Mahalanobis distance (D^2) (Mahalanobis, 1936) and tested pairwise differences by a permutation test with 10,000 iterations, applying the Bonferroni correction for multiple comparisons. The test adopted a significance level (α) of 0.05. Based on congruencies, we recognized morphologically discontinuous groups.

3 RESULTS

3.1 PHYLOGENETIC INFERENCE

ITS alignment comprised 634 sites, of which 266 (41%) were variables and 189 (29%) were parsimonious informative. Among *H. stella* populations, 20 sites were variables, and nine sites were parsimonious informative.

ML and BI trees of ITS displayed overall low resolution (Fig. 4). Both trees indicated that *H. stella* morphotypes are scattered within a clade of *Hydrocotyle* spp. that occur within the Brazilian Atlantic forest. A single clade, which included QINQE1 + QINQE2 + SUBGL2, attained considerable support on both trees. Relationships among other *H. stella* morphotypes, as well as their monophyly, were unresolved.

Figure 4 – Maximum likelihood (ML) and Bayesian inference (BI) trees of the *H. stella* complex based on ITS. Numbers near nodes indicate bootstrap support and Bayesian posterior probability for ML and BI trees, respectively. Nodes with support below 75% are collapsed.

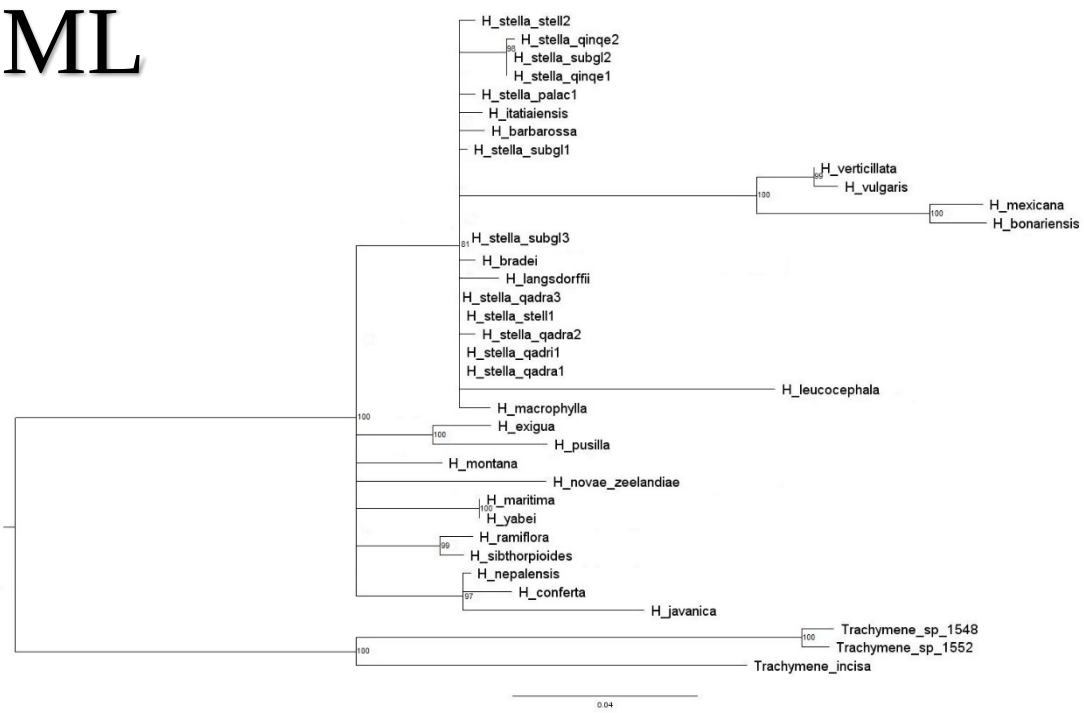
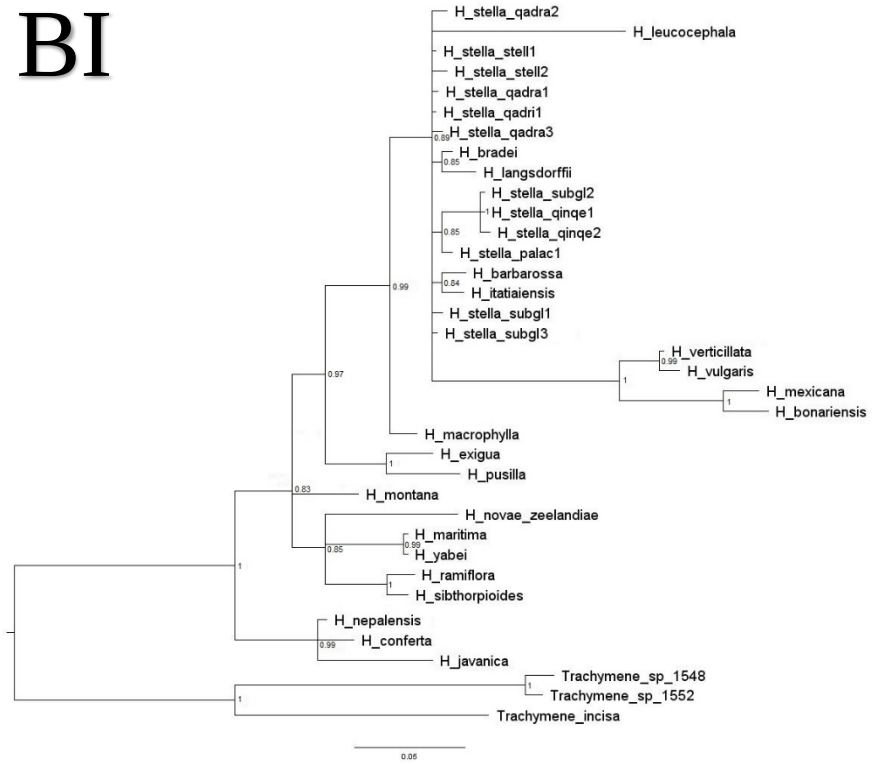


Figure 4 – Continued



3.2 POPULATION GENETICS

The four SSR-targeting primers amplified 94 loci, of which 84 (92%) were polymorphic (Table 3).

Table 3 – SSR-targeting primers and their respective number of amplified loci, polymorphic loci, and percentage (%) of polymorphic loci.

Primer	Amplified	Polymorphic	% polymorphic
$(AC)_7RG$	34	34	100
$(AG)_8TG$	20	19	95
$(CA)_6RY$	16	15	93
$(CTC)_6T$	24	19	79
TOTAL	**94**	**87**	**92**

STRUCTURE analyses of ISSR data indicated increasing mean likelihood and low standard deviation for K values from 1 to 8 (Fig. 5A). Delta K index had its mode at K=8 (Fig. 5B), indicating this value as the optimal. Adopting K=8, ancestry coefficient indicated considerable admixture among most *H. stella* morphotypes since they shared influence of genetic clusters (k) (Fig. 5C). Overall, PALAC included only k1; QINQE and SUBGL shared k2 and k4; STELL included k5 and k6; QADRA and QADRI shared k7 and k8, but QADRA received some contributions from k1, k3 and k6 while QADRI received contributions from k3.

PCA of ISSR data summarized 33% of the genetic variation with its first two axes (Fig. 6). PC1 (17.3% of variation) clearly separated PALAC and roughly separated other morphotypes into two groups: 1) QINQE + SUBGL; 2) QADRA + QADRI + STELL. PC2 (16.1% of variation) separated morphotypes similarly to PC1.

AMOVA of ISSR data indicated significant (p<0.01) fixation values among some morphotypes (Table 4). PALAC displayed high and significant fixation values regarding all other morphotypes. QINQE and SUBGL displayed a low and non-significant fixation value between each other. QADRA did not differ from QADRI and STELL. QADRI displayed a high but non-significant fixation value regarding QINQE.

Figure 5 – STRUCTURE analyses of ISSR data from *H. stella* morphotypes. **A.** Mean likelihood and standard deviation for each K value; **B.** Delta K index for K values; **C.** Ancestry coefficient (Q) of *H. stella* specimens considering a consensus clustering scenario of K=8. Specimens organized by their population. Each color represents a genetic cluster (k).

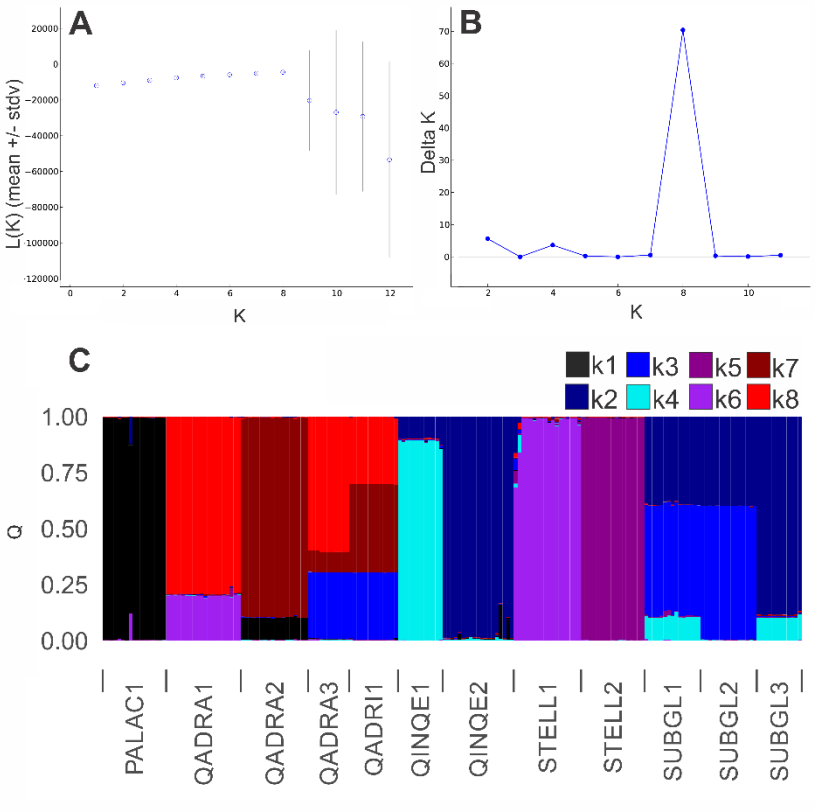
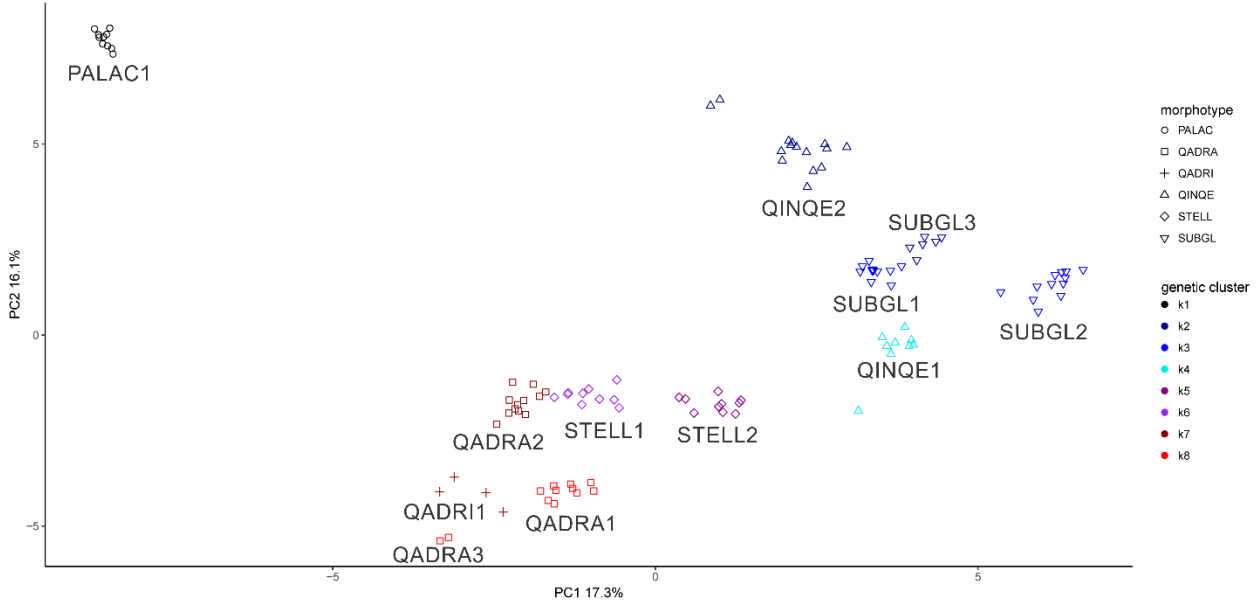


Table 4 – Pairwise AMOVA of ISSR data from *H. stella* morphotypes. Lower semi-matrix displays fixation index (Φ_{ST}) values, and upper semi-matrix displays *p*-values from permutation procedure. Significant values of Φ_{ST} in bold.

	PALAC	QADRA	QADRI	QINQE	STELL	SUBGL
PALAC		0.001	0.001	0.001	0.001	0.001
QADRA	**0.22**		1.000	0.001	1.000	0.001
QADRI	**0.93**	0.00		1.000	0.001	0.001
QINQE	**0.31**	**0.14**	0.28		0.001	1.000
STELL	**0.33**	0.00	**0.26**	**0.20**		0.001
SUBGL	**0.39**	**0.24**	**0.30**	0.03	**0.13**	

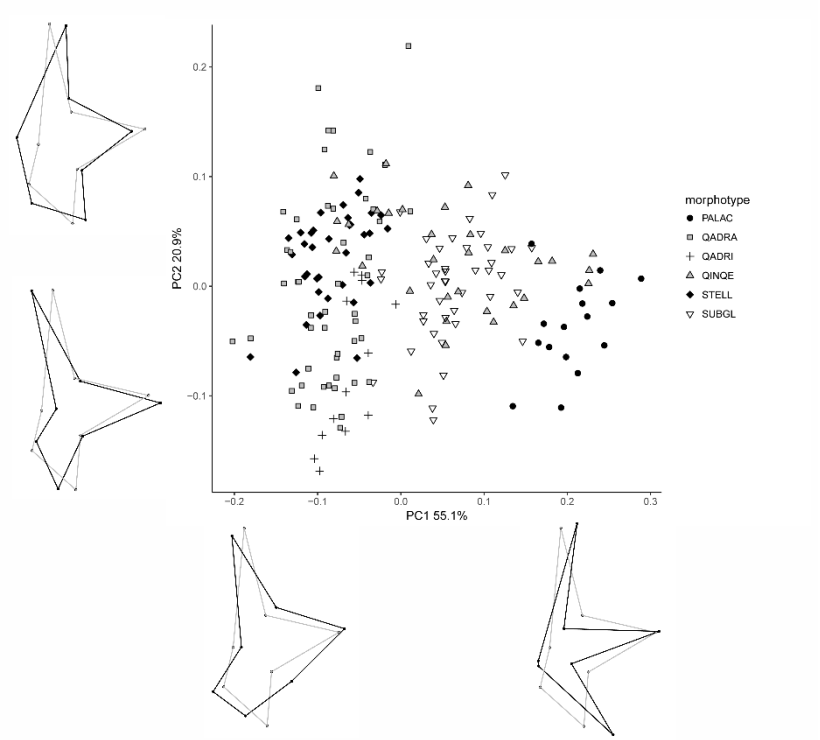
Figure 6 – PCA of ISSR data from *H. stella* morphotypes. Each principal component (PC) displays its percentage of variation. Symbols represent morphotypes, and colors represent genetic clusters inferred by STRUCUTRE. Labels indicate populations.



3.3 MORPHOMETRICS

Geometric morphometrics returned 14 aligned coordinates and the centroid size. PCA summarized 76% of leaf shape variation with its first two axes (Fig. 7). PC1 (55.1% of variation) displayed an approximation among sinuses, an approximation between sinuses and the petiole insertion, and an elongation of middle and rear lobes. PC2 (20.9% of variation) displayed a narrowing of the angle among lobes and a reduction of the lateral lobe. We kept these PCs as leaf shape variables, and we kept the centroid size as a leaf size variable.

Figure 7 – PCA of the aligned coordinates from *H. stella* leaf blades. Each principal component (PC) displays its percentage of variation. Symbols represent morphotypes. Wireframe diagrams display leaf shape at the extreme of PCs.



Traditional morphometrics of vegetative and reproductive structures returned six continuous variables and one discrete variable. Mean values and standard deviation of all variables are in Table 5.

Table 5 – Mean values and standard deviations (within parenthesis) of morphometric variables measured on the *H. stella* complex and its morphotypes. Variables followed by their measurement scale within brackets. Statistics based on all sampled specimens (n=186) for vegetative structures and on a subsample of fertile specimens (n=77) for reproductive structures.

Morphometric variable	Overall	PALAC	QADRA	QADRI	QINQE	STELL	SUBGL
Leaf size [centroid size]	6.6 (2.2)	4.3 (0.8)	5.9 (2.3)	4.6 (1.0)	8.4 (1.9)	7.5 (1.3)	6.8 (2.1)
Internode length [cm]	10.7 (4.5)	9.2 (4.3)	9.7 (5.4)	7.8 (2.9)	11.7 (4.9)	10.8 (4.6)	11.4 (4.6)
Petiole length [cm]	8.9 (3.7)	5.2 (1.8)	9.7 (5.0)	5.8 (2.7)	9.7 (4.1)	9.4 (3.9)	9.2 (3.8)
Stipule length [mm]	2.5 (0.7)	2.3 (0.6)	2.2 (0.4)	1.7 (0.3)	3.2 (0.6)	2.8 (0.7)	2.6 (0.7)
Stipule width [mm]	3.0 (0.9)	2.5 (0.5)	2.5 (0.7)	1.8 (0.4)	3.6 (0.8)	3.6 (0.9)	3.2 (0.8)
Peduncle length [cm]	10.0 (3.3)	5.8 (1.7)	10.5 (3.2)	7.1 (2.5)	11.3 (2.6)	10.7 (2.6)	10.9 (3.1)
Pedicel length [mm]	6.3 (1.6)	4.7 (1.3)	7.2 (1.6)	5.0 (0.8)	7.1 (1.2)	7.3 (1.9)	6.1 (1.5)
Flowers per umbel	40 (12)	26 (6)	44 (5)	30 (5)	50 (21)	46 (8)	44 (13)

When considering morphotypes, CVA indicated two variables as the most discriminant (Table 6). CV1 displayed decreasing leaf shape PC1 scores and increasing flowers per umbel, while CV2 displayed decreasing values for all other variables. When considering genetically structured groups, CVA indicated only one variable as the most discriminant (Table 6). CV1 displayed decreasing leaf shape PC1 scores, while CV2 displayed increasing values for all other variables.

Table 6 – Canonical loadings of morphometric variables measured on *H. stella*, considering currently recognized morphotypes and genetically structured groups recognized in this study. The highest loading for each variable is in bold, whereas signs indicate direction of association.

Morphometric	Morphotypes		Genetic groups	
variable	CV1	CV2	CV1	CV2
Leaf shape PC1	**-0.27**	-0.25	**-0.36**	0.31
Leaf shape PC2	0.08	**-0.39**	0.02	**0.41**
Leaf size	0.18	**-0.32**	0.14	**0.52**
Internode length	0.02	**-0.24**	-0.01	**0.39**
Petiole length	0.11	**-0.25**	0.08	**0.47**
Stipule length	0.06	**-0.40**	0.00	**0.38**
Stipule width	0.15	**-0.34**	0.09	**0.30**
Peduncle length	0.15	**-0.23**	0.14	**0.57**
Pedicel length	0.15	**-0.25**	0.12	**0.49**
Flowers per umbel	**0.17**	-0.16	0.16	**0.19**

When considering morphotypes, CVA summarized 94% of between-group variation with its first two axes (Fig. 8A). CV1 (80.4% of between-group variation) clearly separated PALAC and divided the remaining morphotypes into two mildly overlapping groups: 1) QADRI + QINQE + SUBGL; 2) QADRA + STELL. CV2 (14.0% of between-group variation) mildly separated QADRI from other morphotypes. When considering genetically structured groups, CV1 (91.1% of between-group variation) clearly separated Group 1 and separated the other two groups with mild overlapping. CV2 (8.8% of between-group variation) did not separate groups (Fig. 8B).

Figure 8 – CVA of morphometric variables measured on *H. stella*. Each canonical variate (CV) displays its percentage of between-group variation. **A.** CVA considering currently recognized morphotypes. Wireframe diagrams display mean leaf shape of morphotypes. **B.** CVA considering genetically structured groups recognized in this study. Symbols represent morphotypes, and colors represent genetic structure. Wireframe diagrams display mean leaf shape of groups. Convex hulls encompass the morphological variation within morphotypes (A) and genetically structured groups (B).

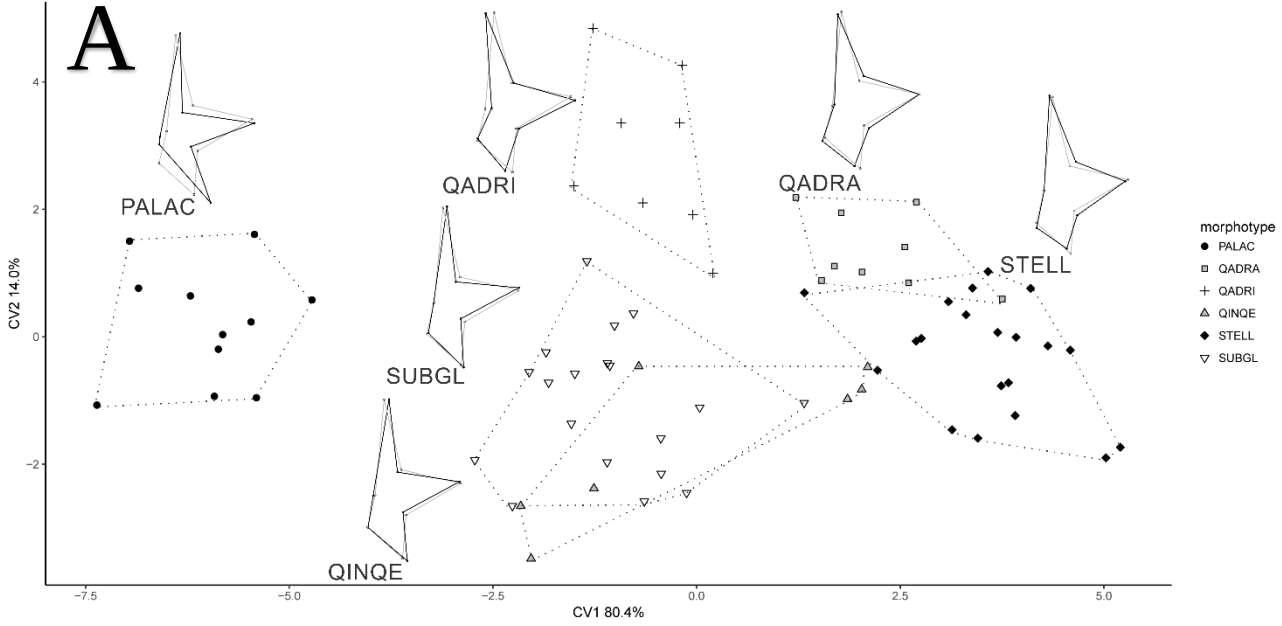
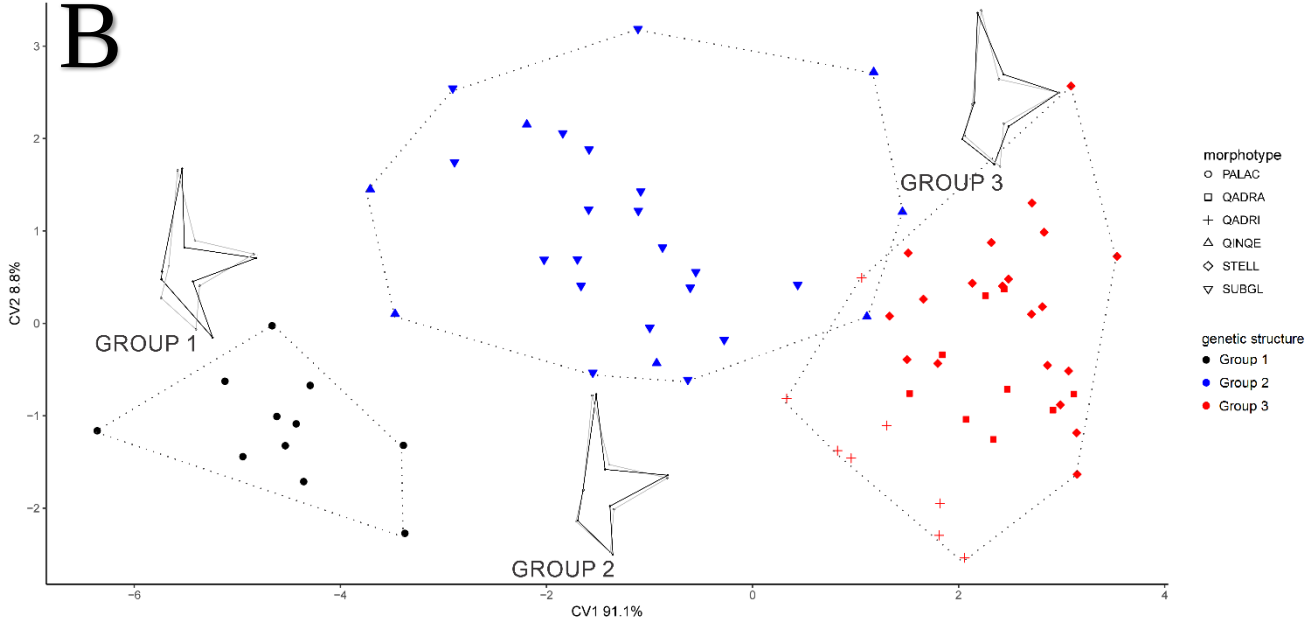


Figure 8 – Continued



When considering morphotypes, permutation test of D^2 indicated few significant (p<0.01) differences (Table 7). PALAC differed from all morphotypes. SUBGL differed only from QADRA and STELL, and QADRI differed only from STELL. When considering genetically structured groups, permutation test of D^2 indicated highly significant (p<0.001) differences among all, with D^2 of 3.9, 6.7, and 3.6 for Group 1-2, Group 1-3, and Group 2-3 comparisons, respectively.

Table 7 - Permutation test of D^2 based on morphometric variables measured on *H. stella* morphotypes. Lower semi-matrix displays D^2, and upper semi-matrix displays *p*-values from permutation procedure. Significant distances in bold.

	PALAC	QADRA	QADRI	QINQE	STELL	SUBGL
PALAC		0.001	0.001	0.001	0.001	0.001
QADRA	**8.3**		0.558	0.424	1.000	0.028
QADRI	**6.2**	3.6		0.121	0.001	0.061
QINQE	**6.4**	3.9	4.7		0.096	1.000
STELL	**9.5**	2.5	**5.3**	4.1		0.001
SUBGL	**5.2**	**4.2**	4.0	2.1	**4.8**	

4 DISCUSSION

Under the general lineage species concept (de Queiroz, 2007), species are separately evolving metapopulation lineages whose delimitation relies on the recognition of properties associated to lineage separation. Here, we assessed species boundaries within the *H. stella* complex, which comprises six morphotypes occurring in the Brazilian Atlantic forest domain. For such, we applied phylogenetic inference, population genetics, and morphometrics to 12 populations of *H stella* in order to asses monophyly, genetic structure, and morphological discontinuity, respectively, among its morphotypes.

Phylogenetic inference based on ITS did not support monophyly of *H. stella* morphotypes. A single monophyletic group, which comprised QINQE and SUBGL2 population, was recovered by both inference criteria (ML and BI) (Fig. 4). Aside this group, neither *H. stella* nor its morphotypes displayed monophyly. The absence of monophyly can be due to either sequence (allele) evolution or lineage evolution (Degnan & Rosenberg, 2009), and both cases are likely to occur within *H. stella*. Considering allele evolution, alleles take different generation times to accumulate substitutions and consequently to reach reciprocal monophyly after lineage separation (Hudson *et al.*, 2002), and our ITS alignment displayed few polymorphic sites among *H. stella* populations, suggesting insufficient substitutions to infer monophyly. Indeed, molecular phylogenies of *Hydrocotyle* spp. from Australia have attained higher resolution and support based on other markers (ETS, *psb*A–*trn*H, and *trn*L–*trn*F) (Perkins, 2019). Hence, the addition of other markers to phylogenetic inference may clarify whether *H. stella* morphotypes are monophyletic or not. On the other hand, considering lineage evolution, reticulate processes (e.g. hybridization and introgression) can break monophyly even though alleles display enough substitutions (Linder & Rieseberg, 2004). Reticulate evolution has been suggested by incongruent gene trees among some *Hydrocotyle* spp. from Australia (Perkins, 2019), so it is likely to occur in other congeneric systems. Nonetheless, addressing reticulation within *H. stella* also would require the addition of other markers to phylogenetic inference, but our results from population genetics (see below) seem to suggest this process.

Population genetics based on ISSR data indicated that most *H. stella* morphotypes were not genetically structured. PALAC derived from a single genetic cluster (k1) (Fig. 5C) and clearly separated from other morphotypes along PCs (Fig. 6), which is supported by its high and significant fixation values regarding all morphotypes (Table 4). QINQE

and SUBGL shared contributions from two genetic clusters (k2 and k4) (Fig. 5C) and greatly overlapped along PCs (Fig. 6), which is supported by their low and non-significant fixation value (Table 4). STELL and QADRA shared one genetic cluster (k6) (Fig. 5C) and overlapped along PCs (Fig. 6), which is supported by their non-significant fixation value (Table 4). QADRA and QADRI shared contributions mainly from two clusters (k7 and k8) and greatly overlapped along PCs (Fig. 6), which is supported by their low and non-significant fixation value (Table 4). On the other hand, results did not consistently support gene flow among some morphotypes. QADRA and QADRI apparently had common genetic background with SUBGL (sharing of k3) (Fig. 5C), but neither PCA nor fixation values support it. Similarly, QADRA also had a common genetic background with PALAC (sharing of k1) (Fig. 5C), but neither PCA nor fixation values supported it. Due their lack of support, we did not consider such inferences of gene flow. Thus, most *H. stella* morphotypes were not genetically structured as congruent results indicated gene flow among them.

Based on congruencies among genetic analyses, we recognized three groups displaying genetic structure: Group 1 = PALAC; Group 2 = QINQE + SUBGL; Group 3 = QADRA + QADRI + STELL. These groups had a clear geographic association. Group 1 was the northernmost group, occupying the Espinhaço range (Fig. 2). Group 2 occupied the Serra da Mantiqueira and northern Serra do Mar (Fig. 2). Group 3 was the southernmost group, occupying the Serra Geral and southern Serra do Mar (Fig. 2). PC2 roughly captured this geographic association by distributing populations in a south-to-north oriented fashion (Fig. 6). Hence, geographic isolation has likely reduced gene flow among populations and consequently led to genetic structure among groups. Moreover, sympatric occurrences suggest that gene flow may have reduced or ceased by other means than geographic isolation. Group 1 and 2 are sympatric at the Serra da Mantiqueira (Serra do Itatiaia locality), where SUBGL1 and STELL2 populations occur (Fig. 2), and they seem genetically isolated, giving the unshared genetic clusters (Fig. 5C) and non-overlapping variation (Fig. 6) between sympatric populations. This pattern suggests a possible ecological speciation (Anacker & Strauss, 2014), which is corroborated by the different altitudes at which SUBGL1 and STELL2 populations occur (2000m and 1200m, respectively). However, further inferences on causes of genetic structure require a characterization of mating systems, which is still unavailable. So far, field and greenhouse observations of *Hydrocotyle* spp. from Australia have indicated autogamy as the predominant strategy of sexual reproduction

(Keighery, 1982). Therefore, genetic structure within *H. stella* is apparently associated to geographic regions, but processes leading to it still require further investigation.

Morphometrics indicated that most *H. stella* morphotypes were not morphologically discontinuous. PALAC clearly separated from other morphotypes (Fig. 8A) and held significant distances regarding all of them (Table 7), suggesting consistent morphological boundaries. On the other hand, remaining morphotypes roughly divided into two groups (Fig. 8A), but such a division was not supported by distances (Table 7), suggesting a morphological continuum. This lack of morphological boundaries among morphotypes is likely due to their original taxonomic circumscription. As morphotypes were based on former infraspecific taxa proposed by Urban (1879), they inherently conform to a typological perspective of taxonomy, which was predominant at the author's time (Williams & Ebach, 2017). Under this perspective, taxon membership relies on the possession of traits present in a reference form, allowing only minor deviations from the reference, so taxonomic circumscriptions could reflect either within-lineage variation (e.g. population polymorphism) or between-lineage variation (Simpson, 1951). Indeed, population genetics indicated that some morphotypes differed from each other as much as populations within a single morphotype (non-significant fixation values) (Table 4), so they would better represent population-level variation. Thus, most *H. stella* morphotypes were not morphologically discontinuous as they might represent within-lineage morphological variation.

By contrast, morphometrics indicated consistent morphological boundaries for genetically structured groups. Group 1 clearly separated from the others, and Group 2 and 3 mildly overlapped (Fig. 8B). Absent and mild overlapping were further corroborated by highly significant ($p<0.001$) distances. Hence, genetically structured groups displayed consistent morphological differences although discontinuity was unclear between the last two groups in CVA. However, the absence of evident gaps on ordinations plots may not indicate a lack of morphological boundaries, as ordination plots may represent samples drawn from distinct distributions as a continuum (Cadena *et al.*, 2018). Based on that, the detection of different modes (peaks) along distributions of morphological variables would be a more reliable criteria to infer morphological boundaries (Cadena *et al.*, 2018). Regarding *H. stela*, CV1 clearly displayed a three-modal distribution (Fig. 6B) (but see Table 1 from Supplementary Material), suggesting three morphologically discontinuous groups. Moreover, such a three-modal distribution corroborates inferred genetic structure, as short-period gene flow can lead

to a unimodal distribution of character states (Wiens & Servedio, 2000). Therefore, genetically structured groups were also morphologically discontinuous when considering the distribution of morphological variables.

By considering the three methods applied in this study, we believe that currently recognized morphotypes of *H. stella* would not represent lineages as most of them did not display monophyly, genetic structure, or morphological discontinuity. These morphotypes would rather represent the aim to classify morphological variation since they derived from former infraspecific taxa (Urban, 1879). Nonetheless, criteria for the recognition of infraspecific taxa are far from consensus (Hamilton & Reichard, 1992), so these taxa bear unclear meaning within a lineage-oriented taxonomy. Indeed, other studies assessing species boundaries have shown that infraspecific taxa often represent underappreciated lineages (Huang & Knowles, 2016), and some have maintained infraspecific status despite consistent evidences in favor of lineage recognition (Turchetto-Zolet *et al.*, 2012). Based on that, infraspecific taxa not only have unclear meaning but also underestimate species diversity. Similarly, our results indicated that morphotypes were nested within putative lineages. Hence, we strongly suggest that infraspecific taxa should be revised in *Hydrocotyle*, and taxonomic proposals of such taxa should be accepted only if followed by clearly stated criteria, allowing further interpretation (and testing) by others.

Population genetics revealed three groups displaying genetic structure, which was associated to geographic regions. Morphometrics further corroborated such groups, which displayed consistent morphological discontinuities. Hence, these groups would be species by two different criteria, genetic structure (Mallet, 1995) and morphological discontinuity (Du Rietz, 1930). Although phylogenetic inference did not indicate these groups as monophyletic, this is not an evidence against their species status. As argued by de Queiroz (2007), properties indicating lineage separation may evolve at different moments of the speciation continuum, so the accumulation of properties, rather than a single type of property, would be necessary to recognize species. Based on that, we suggest that these groups are species that have not attained monophyly, which is not rare (Rieseberg & Brouillet, 1994; Funk & Omland, 2003). Nonetheless, one still could consider such species as phylogenetic. Here, we adopted the genealogy-based phylogenetic species concept (Baum & Shaw, 1995), which requires reciprocal monophyly at gene trees, but we could have also considered the diagnosability-based phylogenetic species concept (Cracraft, 1983; Nixon & Wheeler, 1990), which requires a

unique set of character states or a fixed character state. Under the last concept, we could consider groups as phylogenetic species since leaf shape discriminated them (highest loading at CV1) (Table 6), indicating diagnosability by leaf shape. Indeed, diagnosability may not differ at all from discontinuity, as methods for the inference of the first have set the ground of models for the inference of the second (Zapata & Jiménez, 2012). Thus, we believe that groups recognized in this study should be treated as species, so a taxonomic treatment is necessary to reduce *H. stella* circumscription to Group 3 and to elevate Group 1 and 2 to the species status.

REFERENCES

Adams DC & Otarola-Castillo E. **2013**. Geomorph: An R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution* **4**: 393–399.

Adams DC, Rohlf FJ & Slice DE. **2004**. Geometric morphometrics: Ten years of progress following the 'revolution'. *Italian Journal of Zoology* **71**: 5–16.

Adams DC, Rohlf FJ & Slice DE. **2013**. A field comes of age: Geometric morphometrics in the 21st century. *Hystrix* **24**: 7–14.

Agapow P, Bininda-Emonds ORP, Crandall KA, *et al.* 2004. The impact of species concept on biodiversity studies. *The Quarterly Review of Biology* **79**: 161–179.

Anacker BL & Strauss SY. **2014**. The geography and ecology of plant speciation: Range overlap and niche divergence in sister species. *Proceedings of the Royal Society B: Biological Sciences* **281**: 1–9.

Balakrishnan R. **2005**. Species concepts, species boundaries and species identification: A view from the tropics. *Systematic Biology* **54**: 689–693.

Baum DA & Shaw KL. **1995**. Genealogical perspectives on the species problem. In: Hoch PC, Stephenson AG, eds. *Experimental and molecular approaches to plant biosystematics*. St. Louis: Missouri Botanical Garden, 289–303.

Benson DA, Karsch-Mizrachi I, Lipman DJ, *et al.* 2005. GenBank. *Nucleic Acids Research* **33**: 34–38.

Bussell JD, Waycott M & Chappill JA. **2005**. Arbitrarily amplified DNA markers as characters for phylogenetic inference. *Perspectives in Plant Ecology, Evolution and Systematics* **7**: 3–26.

Cadena CD, Zapata F & Jiménez I. **2018**. Issues and perspectives in species delimitation using phenotypic data: Atlantean evolution in Darwin's finches. *Systematic Biology* **67**: 181–194.

Campbell NA & Atchley WR. **1981**. The geometry of canonical variate analysis. *Systematic Zoology* **30**: 268–280.

De Candolle AP. **1830**. Umbelliferae. In: De Candolle AP, ed. *Prodromus systematis naturalis regni vegetabilis*. Paris: Sumptibus Sociorum Treuttel et Wurtz, 55–249.

Carstens BC, Pelletier TA, Reid NM, *et al.* 2013. How to fail at species delimitation. *Molecular Ecology* **22**: 4369–4383.

Cheng C, Jia JL & Ran SY. **2015**. Polyethylene glycol and divalent salt-induced DNA reentrant condensation revealed by single molecule measurements. *Soft Matter* **11**: 3927–3935.

Cheng T, Xu C, Lei L, *et al.* 2016. Barcoding the kingdom Plantae: New PCR primers for ITS regions of plants with improved universality and specificity. *Molecular Ecology Resources* **16**: 138–149.

Constance L & Dillon MO. **1990**. A new peltate *Hydrocotyle* (Umbelliferae) from Northern Peru. *Brittonia* **42**: 257–259.

Cracraft J. **1983**. Species concepts and speciation analysis. *Current Ornithology* **1**: 159–187.

Degnan JH & Rosenberg NA. **2009**. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology and Evolution* **24**: 332–340.

Doyle JJ & Doyle JL. **1987**. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**: 11–15.

Drummond AJ & Rambaut A. **2007**. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **8**: 1–8.

Eichler H. **1987a**. Nomenclatural and bibliographical survey of *Hydrocotyle* L. (Apiaceae) part II. *Feddes Repertorium* **98**: 145–196.

Eichler H. **1987b**. Nomenclatural and bibliographical survey of *Hydrocotyle* L. (Apiaceae) part I. *Feddes Repertorium* **98**: 1–51.

Eichler H. **1987c**. Nomenclatural and bibliographical survey of *Hydrocotyle* L. (Apiaceae) part III. *Feddes Repertorium* **98**: 273–351.

Evanno G, Regnaut S & Goudet J. **2005**. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611–2620.

Excoffier L, Smouse PE & Quattro JM. **1992**. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.

Falush D, Stephens M & Pritchard JK. **2007**. Inference of population structure using multilocus genotype data: Dominant markers and null alleles. *Molecular Ecology Notes* **7**: 574–578.

Fiaschi P & Plunkett GM. **2011**. Monophyly and phylogenetic relationships of Neotropical *Schefflera* (Araliaceae) based on plastid and nuclear markers. *Systematic Botany* **36**: 806–817.

Flora do Brasil 2020 em construção. *Jardim Botânico do Rio de Janeiro*. http://floradobrasil.jbrj.gov.br/ Acessed January 2019.

Frontier S. **1976**. Étude de la décroissance des valeurs propres dans une analyse en composantes principales: Comparaison avec le modèle du bâton brisé. *Journal of Experimental Marine Biology and Ecology* **25**: 67–75.

Funk DJ & Omland KE. **2003**. Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics* **34**: 397–423.

Garnett ST & Christidis L. **2017**. Taxonomy anarchy hampers conservation. *Nature* **546**: 25–27.

Gonzalo-Turpin H & Hazard L. **2009**. Local adaptation occurs along altitudinal gradient despite the existence of gene flow in the alpine plant species *Festuca eskia*. *Journal of Ecology* **97**: 742–751.

Goodman M, Czelusniak J & Moore GW. **1979**. Fitting the gene lineage into its species lineage, a parsimony strategy illustrated by cladograms constructed from globin sequences. *Systematic Zoology* **28**: 132–163.

Grundt HH, Kjolner S, Borgen L, *et al.* 2006. High biological species diversity in the arctic flora. *Proceedings of the National Academy of Sciences* **103**: 972–975.

Hamilton CW & Reichard SH. **1992**. Current practice in the use of subspecies, variety, and forma in the classification of wild plants. *Taxon* **41**: 485–498.

Hausdorf B & Hennig C. **2010**. Species delimitation using dominant and codominant multilocus markers. *Systematic Biology* **59**: 491–503.

Henderson A. **2005**. The methods of herbarium taxonomy. *Systematic Botany* **30**: 456–459.

Henwood MJ. **2014**. *Hydrocotyle rivularis*: a new trifoliolate species from south-eastern Australia. *Telopea* **17**: 217–221.

Hey J & Pinho C. **2012**. Population genetics and objectivity in species diagnosis. *Evolution* **66**: 1413–1429.

Hey J, Waples RS, Arnold ML, *et al.* 2003. Understanding and confronting species uncertainty in biology and conservation. *Trends in Ecology and Evolution* **18**: 597–603.

Hotelling H. **1933**. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* **24**: 417–441.

Huang JP & Knowles LL. **2016**. The species versus subspecies conundrum: Quantitative delimitation from integrating multiple data types within a single Bayesian approach in hercules beetles. *Systematic Biology* **65**: 685–699.

Huang Q, Zhang S, Huang R, *et al.* 2013. Isolation and identification of an anti-hepatitis B virus compound from *Hydrocotyle sibthorpioides* Lam. *Journal of Ethnopharmacology* **150**: 568–575.

Hudson RR, Coyne JA & Url S. **2002**. Mathematical consequences of the genealogical species concept. *Evolution* **56**: 1557–1565.

Hull DL. **1977**. The ontological status of species as evolutionary units. In: Butts RE, Hintikka J, eds. *Foundational problems in the special sciences*. Dordrecht: R. Reidel Publishing Company, 91–102.

Isaac NJB, Mallet J & Mace GM. **2004**. Taxonomic inflation: Its influence on macroecology and conservation. *Trends in Ecology and Evolution* **19**: 464–469.

Jakobsson M & Rosenberg NA. **2007**. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**: 1801–1806.

Johnson PN & Webb CJ. **1982**. *Hydrocotyle* (Umbelliferae) in New Zealand: the 3-foliolate species. *New Zealand Journal of Botany* **20**: 163–168.

Jombart T & Ahmed I. **2011**. Adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**: 3070–3071.

Jombart T, Lyon D & Biome L De. **2008**. Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**: 1403–1405.

Kamvar ZN, Brooks JC & Grünwald NJ. **2015**. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics* **6**: 208.

Kamvar ZN, Tabima JF & Grünwald NJ. **2014**. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2**: e281.

Keighery GJ. **1982**. Reproductive strategies of western australian Apiaceae. *Plant Systematics and Evolution* **140**: 243–250.

Klingenberg CP. **2013**. Visualizations in geometric morphometrics: How to read and how to make graphs showing shape changes. *Hystrix* **24**: 1–10.

Kumar S, Stecher G, Li M, *et al.* 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35**: 1547–1549.

Li R & Wen J. **2013**. Phylogeny and biogeography of *Dendropanax* (Araliaceae), an amphi-Pacific disjunct genus between tropical/subtropical Asia and the Neotropics. *Systematic Botany* **38**: 536–551.

Linder CR & Rieseberg LH. **2004**. Reconstructing patterns of reticulate evolution in plants. *American Journal of Botany* **91**: 1700–1708.

Liu L, Quan H, Dong BC, *et al.* 2016. Nutrient enrichment alters impacts of *Hydrocotyle vulgaris* invasion on native plant communities. *Scientific Reports* **6**: 1–10.

Luckow M. **1995**. Species concepts: Assumptions, methods, and applications. *Systematic Botany* **20**: 589–605.

Luo A, Ling C, Ho SYW, *et al.* 2018. Comparison of methods for molecular species delimitation across a range of speciation scenarios. *Systematic biology* **67**: 830–846.

Mahalanobis PC. **1936**. On the generalized distance in statistics. *Journal of the Asiatic Society of Bengal* **2**: 541–588.

Mallet J. **1995**. A species definition for the modern synthesis. *Trends in Ecology & Evolution* **10**: 294–299.

Mathias ME & Constance L. **1962**. Umbelliferae. In: Mathias ME, Constance L, eds. *Flora of Peru*. Chicago: Field Museum of Natural History, 3–97.

Mayden RL. **1997**. A hierarchy of species concepts: The denouement in the saga of the species problem. In: Claridge MF, Dawah HA, R.Wilson M, eds. *Species: the units of biodiversity*. London: Chapman and Hall, 381–423.

McLachlan GJ, Lee SX & Rathnayake SI. **2019**. Finite mixture models. *Annual Reviews of Statistics and Its Application* **6**: 1–24.

Mendoza JM & Fuentes AF. **2010**. *Hydrocotyle apolobambensis* (Apiaceae), una Especie Nueva Andina del Noroeste de Bolivia. *Novon* **20**: 303–306.

Miller MA, Pfeiffer W & Schwartz T. **2010**. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *2010 Gateway Computing Environments Workshop (GCE).*, 1–8.

Miralles A & Vences M. **2013**. New metrics for comparison of taxonomies reveal striking discrepancies among species delimitation methods in Madascincus lizards. *PLoS ONE* **8**.

Nicolas AN & Plunkett GM. **2009**. The demise of subfamily Hydrocotyloideae (Apiaceae) and the re-alignment of its genera across the entire order Apiales. *Molecular Phylogenetics and Evolution* **53**: 134–151.

Nicolas AN & Plunkett GM. **2014**. Diversification times and biogeographic patterns in Apiales. *Botanical Review* **80**: 30–58.

Nixon KC & Wheeler QD. **1990**. An amplification of the phylogenetic species concept. *Cladistics-the International Journal of the Willi Hennig Society* **6**: 211–223.

Padial JM, Miralles A, De la Riva I, *et al.* 2010. The integrative future of taxonomy. *Frontiers in Zoology* **7**: 1–14.

Patterson N, Price AL & Reich D. **2006**. Population structure and eigenanalysis. *PLoS Genetics* **2**: e190.

Pearson K. **1901**. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine* **2**: 559–572.

Perkins AJ. **2019**. Molecular phylogenetics and species delimitation in annual species of *Hydrocotyle* (Araliaceae) from South Western Australia. *Molecular Phylogenetics and Evolution* **134**: 129–141.

Peterson AT & Navarro-Sigüenza AG. **1999**. Alternate species concepts as bases for determining priority conservation areas. *Conservation Biology* **13**: 427–431.

Pimenov MG & Leonov M V. 1993. *The genera of the Umbelliferae: a nomenclator*. London: Royal Botanic Gardens, Kew and Botanical Garden of Moscow University.

Plunkett GM & Lowry II PP. **2012**. Phylogeny and diversification in the Melanesian *Schefflera* clade (Araliaceae) based on evidence from

nuclear rDNA spacers. *Systematic Botany* **37**: 279–291.

Plunkett GM, Lowry II PP, Frodin DG, *et al.* 2005. Phylogeny and geography of *Schefflera*: Pervasive polyphyly in the largest genus of Araliaceae. *Annals of the Missouri Botanical Garden* **92**: 202–224.

Pritchard JK, Stephens M & Donnelly P. **2000**. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

de Queiroz K. **2007**. Species concepts and species delimitation. *Systematic Biology* **56**: 879–886.

Rheindt FE & Eaton JA. **2009**. Species limits in *Pteruthius* (Aves: Corvida) shirke-babblers: A comparison between the biological and phylogenetic species concepts. *Zootaxa* **2301**: 29–54.

Richard A. **1820**. *Monographie du genre hydrocotyle de la famille des ombellifères* (A Richard, Ed.). A Brussells: De l'Imprimerie de Weissenbruch.

Rieseberg LH & Brouillet L. **1994**. Are many plant species paraphyletic? *Taxon* **43**: 21–32.

Du Rietz E. **1930**. The fundamental units of biological taxonomy. *Svensk Botanisk Tidskrift* **24**: 333–428.

Rocha FF, Almeida CS, dos Santos RT, *et al.* 2011. Anxiolytic-like and sedative effects of *Hydrocotyle umbellata* extract in mice. *Brazilian Journal of Pharmacognosy* **21**: 115–120.

Rohlf FJ. **2015**. The tps series of software. *Hystrix* **26**: 1–4.

Rohlf FJ & Marcus LF. **1993**. A revolution in morphometrics. *Trends in Ecology and Evolution* **8**: 129–132.

Rohlf FJ & Slice D. **1990**. Extensions of the procrustes method for the optimal superimposition of landmarks. *Systematic Zoology* **39**: 40–59.

Ruiz-Avila RJ & Klemm V V. 1996. Management of *Hydrocotyle ranunculoides* L.f., an aquatic invasive weed of urban waterways in western Australia. *Hydrobiologia* **340**: 187–190.

Sambatti JBM, Rice KJ & Ambatti JUBMS. **2006**. Local adaptation, patterns of selection, and gene flow in the Californian serpentine sunflower (*Helianthus exilis*). *Evolution* **60**: 696–710.

Sanz N, Araguas RM, Fernández R, *et al.* 2009. Efficiency of markers and methods for detecting hybrids and introgression in stocked populations. *Conservation Genetics* **10**: 225–236.

Schlager S. **2017**. Morpho and Rvcg - Shape Analysis in R. In: Zheng G, Li S, Szekely G, eds. *Statistical Shape and Deformation Analysis*. Academic Press, 217–256.

Scrucca L, Fop M, Murphy TB, *et al.* 2016. mclust 5: Clustering, classification and density estimation using Gaussian finite mixture models. *The R journal* **8**: 289–317.

Simpson GG. **1951**. The species concept. *Evolution* **5**: 285–298.

Tronchet F, Plunkett GM, Jeremie J, *et al.* 2005. Monophyly and

major clades of *Meryta* (Araliaceae). *Systematic Botany* **30**: 657–670.

Turchetto-Zolet AC, Cruz F, Vendramin GG, *et al.* 2012. Large-scale phylogeography of the disjunct Neotropical tree species *Schizolobium parahyba* (Fabaceae-Caesalpinioideae). *Molecular Phylogenetics and Evolution* **65**: 174–182.

Urban I. **1879**. Umbelliferae. In: Martius CFP, Eichler AW, Urban I, eds. *Flora brasiliensis*. Monachii et Lipsiae: Friedrich Fleischer Comm, 261–354.

Wiens JJ. **2007**. Species delimitation: New approaches for discovering diversity. *Systematic Biology* **56**: 875–878.

Wiens JJ & Servedio MR. **2000**. Species delimitation in systematics: Inferring diagnostic differences between species. *Proceedings of the Royal Society B: Biological Sciences* **267**: 631–636.

Williams DM & Ebach MC. **2017**. What is intuitive taxonomic practice? *Systematic Biology* **66**: 637–643.

Wolfe AD, Randle CP & Wilson P. **2001**. Relationships within and among species of the holoparasitic genus *Hyobanche* (Orobanchaceae) inferred from ISSR banding patterns and nucleotide sequences. *Systematic Botany* **26**: 120–130.

Wright S. **1949**. The genetical structure of populations. *Annals of Eugenics* **15**: 323–354.

Zachos FE. **2018**. Mammals and meaningful taxonomic units: the debate about species concepts and conservation. *Mammal Review* **48**: 153–159.

Zapata F & Jiménez I. **2012**. Species delimitation: Inferring gaps in morphology across geography. *Systematic Biology* **61**: 179–194.

Zhivotovsky LA. **1999**. Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology* **8**: 907–913.

Zietkiewicz E, Rafalski A & Labuda D. **1994**. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* **20**: 176–183.

SUPPLEMENTARY MATERIAL

Table 1 – Optimal density distribution of *H. stella* phenotype inferred by normal mixture model analysis (McLachlan, Lee, & Rathnayake, 2019). Phenotype measured by the dimension reduction analysis implemented in the R package mclust (Scrucca *et al.*, 2016). Box-whisker plots at the bottom represent the phenotype range of groups recognized in population genetics, where boxes and whiskers encompass 75% and 95% of observations, respectively.