Evolutionary diversification and historical biogeography of orchidaceae in Central America with emphasis on Costa Rica and Panama

Bogarin Chaves, D.G.

Chapter 4

Anchored Hybrid Enrichment generated nuclear, plastid and mitochondrial markers resolve the *Lepanthes horrida* (Orchidaceae: Pleurothallidinae) species complex

Diego Bogarín, Oscar A. Pérez-Escobar, Dick Groenenberg, Sean D. Holland, Adam P. Karremans, Emily Moriarty Lemmon, Alan R. Lemmon, Franco Pupulin, Erik Smets and Barbara Gravendeel

Molecular Phylogenetics and Evolution 129: 27–47. 2018.

Abstract. Phylogenetic relationships in species complexes and lineages derived from rapid diversifications are often challenging to resolve using morphology or standard DNA barcoding markers. The hyper-diverse genus *Lepanthes* from Neotropical cloud forest includes over 1200 species and many recent, explosive diversifications that have resulted in poorly supported nodes and morphological convergence across clades. Here, we assess the performance of 446 nuclear-plastid-mitochondrial markers derived from an anchored hybrid enrichment approach (AHE) coupled with coalescence and species network-based inferences to resolve phylogenetic relationships and improve species recognition in the *Lepanthes horrida* species group. In addition to using orchid-specific probes to increase enrichment efficiency, we improved gene tree resolution by extending standard angiosperm targets into adjacent exons. We found high topological discordance among individual gene trees, suggesting that hybridization/polyploidy may have promoted speciation in the lineage via formation of new hybrid taxa. In addition, we identified ten loci with the highest phylogenetic informativeness values from these genomes. Most previous phylogenetic sampling in the Pleurothallidinae relies on two regions (ITS and matK), therefore, the evaluation of other markers such as those shown here may be useful in future phylogenetic studies in the orchid family. Coalescent-based species tree estimation methods resolved the phylogenetic relationships of the *L. horrida* species group. The resolution of the phylogenetic estimations was improved with the inclusion of extended anchor targets. This approach produced longer loci with higher discriminative power. These analyses also disclosed two undescribed species, *L. amicitiae* and *L. genetoapophantica*, formally described here, which are also supported by morphology. Our study demonstrates the utility of combined genomic evidence to disentangle phylogenetic relationships at very shallow levels of the tree of life, and in clades showing convergent trait evolution. With a fully resolved phylogeny, is it possible to disentangle traits evolving in parallel or convergently across these orchid lineages such as flower color and size from diagnostic traits such as the shape and orientation of the lobes of the petals and lip.

4.1 Introduction

Identification and inventory of plant species' diversity remains an arduous task in tropical countries. Species identification in many plant groups is still largely based on phenotypic differences and proportionally very little of this has been supported from molecular evidence (Fujita et al., 2012; Granados Mendoza et al., 2013; Yang and Rannala, 2010). Although morphological differences are sometimes sufficient to separate species, precise circumscriptions of closely related species in species complexes with poor morphological differentiation or recently diversified lineages often require the support of additional sources of evidence, including molecular data. Molecular species delimitations are also important in recognizing potential unknown cryptic species or to assess taxon descriptions previously proposed on the basis of morphology only (Fujita et al., 2012; Lahaye et al., 2008).

Orchidaceae are a prime example of a highly diverse plant family with recent explosive diversifications leading to a server's worth of backlogged species to identify in the tropics (Givnish et al., 2015). In this species rich family, molecular information has been largely used to assess generic and supra-generic relationships but not as commonly for assessing species delimitations (Karremans et al., 2015a; Lahaye et al., 2008; Ramos-Castro et al., 2012). This is particularly true for hyperdiverse, young tropical lineages derived from rapid diversifications, for which standard DNA molecular markers are insufficient to resolve phylogenetic relationships at very shallow levels (Pérez-Escobar et al., 2016b). Molecular-based approaches of species delimitations could lead to false inferences when single or few-locus datasets are analyzed because of topological discordances among gene trees and species trees (Edwards, 2009). The main sources of phylogenetic incongruence in single locus datasets and resulting gene trees are caused by systematic or stochastic errors and by biological evolutionary processes such as hybridization, introgression, gene duplication, deep coalescence and branch length heterogeneity (Chan et al., 2017; Maddison, 1997; Mallo and Posada, 2016; Pérez-Escobar et al., 2016a).

To cope with phylogenetic incongruence and resolution, recent studies have focused on the assessment of large multi-locus datasets derived from multiple genomic compartments (i.e. nuclear, plastid and/ or mitochondrial) to achieve species delimitation (Brandley et al., 2015; Granados Mendoza et al., 2013; Hamilton et al., 2016; Peloso et al., 2016; Ruane et al., 2015). Xi, Liu, and Davis (2015) found that genes with low phylogenetic resolution produce unreliable gene trees affecting species tree estimations based on gene tree coalescent methods, a problem that can be solved by sampling more genes. Thereby, inferences based on large multi-locus datasets gathered from Next Generation Sequencing (NGS) or High-throughput-sequencing (HTS) techniques theoretically help improving the accuracy of species-tree based delimitations because of the higher amount of genomic data analyzed and the higher levels of sequence divergence obtained with respect to traditional approaches employing only a few markers (Jeffroy et al., 2006; Wagner et al., 2013). One of the NGS protocols for high-throughput phylogenomics which allows the capture of hundreds of orthologous markers is Anchored Hybrid Enrichment (AHE) (Buddenhagen et al., 2016; Fragoso-Martínez et al., 2016; Lemmon et al., 2012; Wanke et al., 2017). Multi-locus datasets derived from AHE of plants usually contain fragments from plastid, mitochondrial and nuclear genomes. These genomes might have linked but different evolutionary histories because their different modes of inheritance, hence the resulting phylogenetic re-

lationships might be incongruent (Jeffroy et al., 2006; Pérez-Escobar et al., 2016a). Increasing the number of analyzed genes does not guarantee *per se* the inference of an accurate species tree because the detection of topological discordance in species/ gene trees is pervasive in multi-locus based inferences (Jeffroy et al., 2006). However, these incongruences are informative because they provide clues on relevant biological phenomena for speciation such as hybridization and polyploidization (Maddison, 1997; Soltis and Soltis, 2016).

In addition to inconsistencies among multiple genome regions inherent to evolutionary processes, different species tree estimation methods might produce discordant results as well. Methods using concatenation of multi-locus datasets assume that the loci analyzed evolved in a similar way and thus phylogenies or species delimitations are inferred from standard concatenation of hundreds of anonymous markers (Lemmon and Lemmon, 2012; Mirarab and Warnow, 2015; Xi et al., 2015). However, the results produced by concatenation approaches can differ from coalescent species tree reconstructions because multi-species coalescent models recognize that gene trees exhibit different evolutionary histories and thus reduce the influence of incomplete lineage sorting (ILS) and gene duplication. Recently, Sukumaran and Knowles (2017) argued that multi-species coalescent models delimit genetic structure without making any statistical distinction between structure due to population-level processes or due to speciation. Therefore, it is possible that population structure might be misidentified as a putative species boundary. In conclusion, the authors suggested that hypotheses based on multispecies coalescent models require validation with other evidence such as morphological or ecological information (Pyron et al., 2016).

The performance of AHE multi-locus datasets in plants has been assessed in several groups, including Arecaceae, Fabaceae, Lamiaceae, Oxalidaceae, Pinaceae, Proteaceae, Sarraceniaceae and Zingiberales (Fragoso-Martínez et al., 2016; Heyduk et al., 2016; Mitchell et al., 2017; Wanke et al., 2017). However, no studies assessing species delimitations based on concatenation or coalescent species tree estimations from AHE datasets in the Orchidaceae have been published. In this study, we focus on a species complex of the highly diverse Neotropical orchid genus *Lepanthes* Sw. (Pleurothallidinae), which contains more than 1,200 species. Recent studies on the evolutionary diversification of Neotropical orchids revealed that *Lepanthes* is a relatively young group which diverged 5–10 Ma and that shows the highest net diversification rates across the Pleurothallidinae (Pérez-Escobar et al., 2017a). In *Lepanthes*, the percentage of endemism is high, especially in relatively young mountain ranges such as the Cordillera de Talamanca in southern Central America and other Andean regions.

The *Lepanthes horrida* group consists of five taxa endemic to Costa Rica and Panama along the Cordillera Volcánica Central and Cordillera de Talamanca: *L. chameleon* Ames, *L. horrida* Rchb.f., *L. maxonii* Schltr., *L. nymphalis* Luer, and *L. wendlandii* Rchb.f. Although previously known specimens belonging to the *Lepanthes horrida* group could be easily separated morphologically, a clear distinction between recent collections from populations in Cordillera de Talamanca could not be made morphologically nor in combination with conventional phylogenetic analysis of the traditional markers used in Pleurothallidinae, namely nrITS and *mat*K.

To investigate the utility of AHE in resolving species complexes in lineages with rapid diversifications we inferred the phylogenetic relationships of the species in the *Lepanthes horrida* group. We obtained 446 target orthologous loci, generated with AHE by specific probes designed

for Pleurothallidinae orchids across the recently published *Phalaenopsis equestris* (Schauer) Rchb.f. (Cai et al., 2014) and *Dendrobium catenatum* Lindl. (Zhang et al., 2016) genomes. The performance of loci recovered was evaluated with both concatenation and coalescent-based methods (including and excluding missing sequences) and with analyses of phylogenetic informativeness (Townsend, 2007). The loci were identified and classified in three separate groups: plastid, mitochondrial and nuclear; each dataset was evaluated for incongruences inherent to multi-locus based inferences (Jeffroy et al., 2006). This study assesses the evolutionary relationships of the species of the *Lepanthes horrida* group by answering the following questions: (i) can the application of NGS with concatenated and multi-species coalescent models disclose species relationships in recently diverged clades which were unsolved with Sanger sequencing generated nuclear (ITS) and plastid markers (*mat*K)? (ii) are the hypotheses of supermatrix and species coalescent delimitations consistent with morphological evidence? (iii) are there phylogenetic incongruences among inferences based on plastid, mitochondrial and nuclear multi-locus datasets and what are the possible sources of this discordance?.

4.2 Materials and Methods

4.2.1 Taxon sampling

Living plant specimens were collected in the field and cultivated at JBL between 2013 and 2017. We sampled the five taxa belonging to the *L. horrida* group including individuals that do not correspond morphologically to any of the known species. As outgroup, we selected *Gravendeelia chamaelepanthes* (Rchb.f.) Bogarín & Karremans, a closely related species to the genus *Lepanthes* according to the latest phylogenetic studies in the group (Bogarín et al., 2018c) in addition to *Lepanthes elata* Rchb.f., *Lepanthes gargantua* Rchb.f. and *Lepanthopsis prolifera* Garay (Table 1). Vouchers were preserved as herbarium/spirit specimens for future reference at CR, JBL and L.

4.2.2 DNA extraction

Total genomic DNA was extracted from about 100 mg of silica gel dried leaf/flower tissue. Each dried sample was frozen in liquid nitrogen and powdered in a Retsch MM 300 shaker for 5 min. We followed the 2×CTAB (Hexadecyltrimethylammonium bromide) protocol for isolating DNA (Doyle and Doyle, 1987). Resulting total DNA was treated with Ribonuclease A (RNase A, Qiagen) and quantified with a Qubit 3.0 Fluorometer (TermoFischer Scientific®) to ensure 2.0 µg of DNA per sample in 130 µl of buffer. All DNA samples (4 µL of sample DNA and 2 µL of 6× loading dye) were checked on a 2% agarose gel in 1× TAE (Tris-Acetate-EDTA) buffer running for 90 min at 120 V with a size ladder of 100 bp–1000 bp fragments.

4.2.3 Anchored phylogenomics locus selection and probe design

We aimed to collect data for the Angiosperm AHE target loci (Buddenhagen et al., 2016; Léveillé-Bourret et al., 2018; Wanke et al., 2017). Refinement of the target regions and corresponding probes for data collection in orchids was conducted at the Center for Anchored Phylogenomics (www.anchoredphylogeny.com). To improve enrichment efficiency in Orchidace-

ae, we leveraged the published genomes of two orchid species, *P. equestris* (Cai et al., 2015; NCBI Bioproject PRJNA192198), and *D. catenatum* (Zhang et al., 2016; NCBI Bioproject PRJNA192198). Following the approach of (Ruane et al., 2015), we obtained AHE target locus sequences for these two species using three reference sequences from the Angiosperm AHE V1 kit: *Lactuca sativa* L., *Arabidopsis thaliana* (L.) Heynh., and *Oryza sativa* L. In addition to using orchid-specific probes, we extended the standard angiosperm targets into adjacent exons to obtain longer loci. Candidate regions identified using spaced kmers (17 of 20 matches) were verified as a good match if at least 55 of 100 consecutive bases matched between the orchid sequence and one or more of the references. For each locus, a 4,000 bp region centered on the best-matching region was isolated for each species. For each locus, alignments containing the isolated *P. equestris* and *D. catenatum* sequences in addition to the three corresponding reference sequences were then estimated using MAFFT v.7 (Katoh and Standley, 2013). The alignments were inspected in Geneious (R9; Biomatters Ltd., Kearse et al., 2012) and the largest well-aligned region containing the AHE V1 probe region was identified. This procedure produced target loci substantially larger than that found in the V1 design, such that neighboring loci sometimes overlapped. When this occurred, the smaller target locus was removed as a target region. Finally, sequences were profiled following (Hamilton et al., 2016) and repetitive regions were masked. Final alignments used for probe design represented 451 target loci (448 of which contained both species). The loci averaged 887 bp in length (90% were between 261 bp and 1,973 bp) and had pairwise identity values averaging 77.7% (90% of loci had values between 65.5% and 87.5%). Probes of length 120 bp were tiled uniformly at 10x density across the two orchid sequences in each alignment, producing 53,881 probes in total.

4.2.4 Sample processing

Data were collected by the Center for Anchored Phylogenomics. Following DNA extraction, a Covaris E220 Focused ultrasonicator with Covaris microTUBES was used to fragment genomic DNA to a distribution of 300–800 bp. Libraries were prepared and indexed (8 bp) on a Beckman-Coulter Biomek FXp liquid-handling robot that implemented a modified version of Meyer and Kircher (2010) protocol. All libraries were then pooled at equal quantities, and enrichments were performed using an Agilent SureSelect XT probe kit containing the probes described above. Enriched library pools were pooled and sequencing on one half of an Illumina HiSeq2500 lane (23.9 Gb of raw data). Sequencing was performed in the Translational Science Laboratory in the College of Medicine at Florida State University.

4.2.5 Raw data processing

Raw sequence reads were processed using Illumina's CASAVA pipeline (v1.8) and low-quality reads were quality filtered using the high chastity setting. Quality filtered reads were then demultiplexed using 8bp indexes, which differed by at least 2 bases. Reads with corresponding indexes not matching one of the 16 expected indexes were discarded. Read accuracy and length were enhanced through pairedread merging, which was performed following the approach of Rokyta et al. (2012). Reads were assembled using a quasi de novo approach described by Prum et al. (2015) and Hamilton et al. (2016). Reads were mapped to probe region sequences from the orchid design

Table 4.1. Voucher specimens and species analyzed.

Sample code	Species	Voucher	Country
DB01	*Lepanthes nymphalis* Luer	DB11781 (JBL)	Costa Rica
DB02	*Lepanthes elata* Rchb.f.	DB11778 (JBL)	Costa Rica
DB03	*Lepanthes genetoapophantica* Bogarín & Gravend.	DB9745 (JBL)	Costa Rica
DB04	*Lepanthes wendlandii* Rchb.f.	DB11827 (JBL)	Costa Rica
DB05	*Lepanthes amicitiae* Bogarín & Pupulin	DB5911 (JBL)	Panama
DB06	*Lepanthes elata* Rchb.f.	AK6632 (JBL)	Costa Rica
DB07	*Lepanthes genetoapophantica* Bogarín & Gravend.	DB8682 (JBL)	Costa Rica
DB08	*Lepanthes chameleon* Ames	DB8371 (JBL)	Costa Rica
DB09	*Lepanthes wendlandii* Rchb.f.	DB11885 (JBL)	Costa Rica
DB10	*Lepanthes maxonii* Schltr.	DB5914 (JBL)	Panama
DB11	*Lepanthes amicitiae* Bogarín & Pupulin	AK6144 (JBL)	Costa Rica
DB12	*Lepanthes horrida* Rchb.f.	DB11459 (JBL)	Costa Rica
DB13	*Lepanthes wendlandii* Rchb.f.	DB11946 (JBL)	Costa Rica
DB14	*Gravendeelia chamaelepanthes* (Rchb.f.) Bogarín & Karremans	DB11881 (L)	Colombia
DB19	*Lepanthopsis prolifera* Garay	DB12048 (L)	Colombia
DB32	*Lepanthes gargantua* Rchb.f.	DB11868 (L)	Ecuador

described above. The assembly approach involves applying divergent references to initiate the assembly of each locus in the conserved probe region, and then assembling subsequent reads to the initially-mapped reads in a reference-based assembly style. To prevent low-level contamination from being included in downstream analyses, consensus sequences derived from fewer than 742 reads were removed from further analysis. Orthology across consensus sequences was established using pairwise distances following Hamilton et al. (2016). Haplotypes were phased assuming diploidy following (Pyron et al., 2016) and the haplotypes were aligned for each locus using MAFFT v.7 (Katoh and Standley, 2013). Lastly alignments were trimmed using the methods of Hamilton et al. (2016), but requiring 50% of bases in a site to be identical to identify a site as conserved, a minimum of 14 conserved sites in a 20 bp stretch for the stretch to be retained, and allowing only 18% missing data at each site for a site to be retained (see Hamilton et al. (2016) for details). Following the automated alignment trimming/masking procedure, alignments were manually inspected in Geneious R9 (Biomatters Ltd., Kearse et al., 2012) to verify the absence of misaligned regions and obvious paralogs (one locus was removed due to the presence of these issues).

4.2.6 Loci identification and datasets

The 446 loci retrieved were identified and classified in plastid, mitochondrial and nuclear genome datasets by conducting automatic BLAST searches in NCBI GenBank (https://blast.ncbi.nlm.

Table 4.2. Datasets analyzed in this study. All tree inferences contained 32 terminals (16 species with two haplotypes). Datasets derived from the three separate genomes did not have any missing sequences.

Inference	Dataset	Loci	Description
ML-423/ASTRAL-423	Complete matrix	423	All loci with/without missing sequences
ML-305/ ASTRAL-305	Reduced Matrix	305	All loci without missing sequences
ML-n/ ASTRAL-n	Nuclear	254	All loci from the nuclear genome
ML-m/ ASTRAL-m	Mitochondrial	14	All loci from the mitochondrial genome
ML-p/ ASTRAL-p	Plastid	37	All loci from the plastid genome

nih.gov) with an in-house designed script (https://github.com/dickgroenenberg/ Bogarin_Anchored_Phylogenomics) and subsequent further characterization in the TAIR database (https://www.arabidopsis.org). The BLAST hits mostly matched with annotated genes identified in the sequenced genomes of the orchid species *D. catenatum* (Zhang et al., 2016) and *P. equestris* (Cai et al., 2014) and other monocots with fully sequenced genomes such as oil palm, maize and rice. We performed inferences based on five datasets of trimmed loci matrices: (1) using the 305 loci alignments with 100% coverage (without missing sequences), (2) using 423 loci alignments with ~72% coverage (118 with missing sequences for one-two accessions), this because excluding/ including loci with missing sequences from some individuals can affect the outcome derived from both types of datasets (Huang and Lacey Knowles 2016; Mitchell et al. 2017), (3) with 254 loci derived from nuclear, (4) 14 from mitochondrial and (5) 37 from plastid datasets (Table 2). After the initial filtering by number of reads mapped to each locus and checking for orthologs, 446 loci were retained; however, once all data was aligned, if a locus had no sequence data for 3 or more of the 16 taxa in the alignment ($x > 18.75\%$), it was not included in the analyses (a total of 23 alignments, see Results).

4.2.7 Supermatrix (concatenation)

We inferred a maximum likelihood (ML) species tree with the supermatrix approach using the ML-305 and ML-423 datasets. Statistical support was calculated with bootstrap support (BS) and the analysis was performed in RAxML-HPC2 on XSEDE v. 8.2.11 under the GTRGAMMA model for bootstrapping phase and 1000 bootstrap iterations in CIPRES Science Gateway V. 3.3 (Miller et al., 2015; Stamatakis, 2014). We applied the same ML analysis to the concatenated mitochondrial (ML-m), nuclear (ML-n) and plastid (ML-p) datasets.

4.2.8 Gene tree estimation

We generated unrooted ML gene trees for each locus with 100 bootstrap replicates using rapid bootstrapping with RAxML v. 8.2.11 under the GTRGAMMA model (command raxmlHPC-PTHREADS -f a -x 12,345 -p 12,345 -# 100 -k -m GTRGAMMA) as inputs for calculating species-tree estimations and further analysis on concordance and conflict among gene and species trees. The model of evolution for each loci was calculated using the Akaike Information Criterion (AIC) in jModelTest2 v2.1.7 (Darriba et al., 2012).

4.2.9 Super tree estimation

We used the estimated ML gene trees with collapse nodes with >33% bootstrap support in the R package phyloch to infer species trees with four programs developed under the coalescent model: ASTRAL-II v. 5.5.7 (Mirarab and Warnow, 2015; Sayyari and Mirarab, 2016), which is a disagreement reduction method (ASTRAL) and multi-locus bootstrapping (ASTRAL-mlbs) with support calculated with local posterior probability (LPP) and bootstrap support (BS) respectively, MP-EST v1.6 (Liu et al., 2010), NJst (Liu and Yu, 2011) and STAR (L. Liu et al., 2009) which are single evolutionary process methods considering ILS (Liu et al., 2015; Mallo and Posada, 2016; Shaw et al., 2013). NJst and STAR were run with R programming language (R Core Team, 2017) under R Studio (R Studio Team, 2016) using the package phybase (Liu and Yu, 2010) and the STRAW webserver (http://bioinformatics.publichealth.uga.edu/SpeciesTree-Analysis/index.php). Gene trees were rooted online in the STRAW webserver. We inferred a species network with Phylonet v. 3.6.1 (https://bioinfocs.rice.edu/phylonet) using ML and estimated branch lengths of the gene trees for the inference. The network analysis was visualized in Dendroscope v.3 (Huson and Scornavacca, 2012). Phylonet is a multiple evolutionary processes method that reconstructs phylogenetic networks of reticulate evolutionary events (considering ILS and hybridization) (Than et al., 2008; Yu et al., 2014). We also inferred a species network with SplitsTree4 v 4.13.1 (Huson and Bryant, 2006) using the 423 loci dataset and excluding outgroup species with the following settings: Jukes-Cantor for "Characters" and NeighborNet method for "Distances". We tested these methods to evaluate possible discrepancies in the topology of each resulting species tree and the supermatrix approach (Simmons and Gatesy, 2015). We also inferred ASTRAL species trees from mitochondrial (ASTRAL-m), nuclear (ASTRAL-n) and plastid (ASTRAL-p) datasets and for each multi-locus alignment of 305 (ASTRAL-305) and 423 loci (ASTRAL-423). Final trees were manipulated in R using the packages APE, ggtree, phangorn and phytools (Paradis et al., 2004; Revell, 2012; Schliep, 2011; Yu et al., 2017) and later edited in Adobe Illustrator CS6 (Adobe Systems Inc., California, USA).

4.2.10 Gene and species tree concordance/discordance

We evaluated the topological concordance among gene trees, supermatrix approach, species trees and the analysis from the different genomic datasets with the R package TreeSpace v. 1.10.19 (Jombart et al., 2017). We identified clusters of similar trees with Metric Multidimensional Scaling (MDS) based on Robinson-Foulds (RF) symmetric difference (Robinson and Foulds, 1981) (unrooted, topological) and Ward clustering method and topological concordance among gene trees with RF and Kendall-Colijn metric vector (Kendall and Colijn, 2016) in order to test possible differences in unrooted and rooted based tests, respectively. We evaluated the level of concordance among gene trees (without missing data and collapsed nodes with < 33% support) against the mapping reference ASTRAL-305 species tree, ASTRAL-m, ASTRAL-n and ASTRAL-p and the ML supermatrix tree, ML-m, ML-n and ML-p with the program PhyParts (https://bitbucket.org/blackrim/phyparts) (Smith et al., 2015). Trees were previously rooted in R with the package APE and *G. chamaelepanthes* as outgroup. The output obtained with Phyparts was visualized by plotting pie charts on the ASTRAL species tree and ML concatenated tree with the script PhyPartsPieCharts (https://github.com/mossmatters/MJPythonNotebooks) using the ETE3 Python toolkit (Huerta-Cepas et al., 2016).

4.2.11 Phylogenetic informativeness

We evaluated the performance of the AHE datasets by measuring the net phylogenetic informativeness (PI) through an arbitrary time scale (tips assigned to time 0 and root to time 1) as described by Townsend (2007). This method has been used to calculate the power of each locus in resolving a node over time in AHE datasets (Fragoso-Martínez et al., 2016; Pyron et al., 2014; Wanke et al., 2017). To estimate the PI, we converted the rooted consensus ML trees (ML-305 and ML-423) to ultrametric trees with PATHd8, a program for phylogenetic dating without a molecular clock (https://www2.math.su.se/PATHd8/) (Britton et al., 2007; Schoch et al., 2009). The PATHd8 method calculates ultrametric trees with branch lengths proportional to the number of substitutions and these substitutions rates are smoothed locally (Britton *et al*., 2007). The partitioned concatenated matrices were built in SequenceMatrix v100.0 from the trimmed loci alignments (Vaidya et al., 2011). These input files were uploaded in the web application PhyDesign (López-Giráldez and Townsend, 2011; http://phydesign. townsend.yale.edu/) to estimate phylogenetic informativeness profiles with the HyPhy substitution rates algorithm for DNA sequences (Kosakovsky Pond et al., 2005). For the identification of sites with unusually high substitution rates that could cause phylogenetic noise, we followed the R script and filtering method described by Fragoso-Martínez et al. (2016). Sites with rate values higher than five were removed manually from the alignments using Geneious R9 (Biomatters Ltd., Kearse et al., 2012) and these corrected matrices were uploaded again to PhyDesign as described above.

4.3 Results and discussion

4.3.1 New phylogenetic markers generated

A total of 446 innovative loci generated from 16 plant samples and two haplotypes (diploidy as default assumption) for each locus were identified (supplementary material: Table 4.3; https://www.sciencedirect.com/science/article/pii/S1055790318301623#s0310). Of these, 367 (82.10%) loci were nuclear, 19 (4.47%) mitochondrial and 58 (13.42%) plastid derived. Only two fragments could not be assigned to any genome (scored as N/A in Table 4.3); 54 (12.11%) could not be linked to any protein (product scored as uncharacterized in Table 3). To the best of our knowledge, no studies using AHE have characterized and annotated loci recovered under this approach. The average locus length was 1074 bp, the shortest was 114 bp and the longest was 4644 bp. The number of loci recovered is similar to that obtained in *Salvia* (Lamiaceae) (448) and higher than other monocots (i.e. Arecaceae: 133, and Zingiberales: 308) (Fragoso-Martínez et al., 2016). The average locus length is higher than previous AHE datasets in *Aristolochia* (670–687 bp), *Protea* (551 bp) and *Salvia* (704 bp) (Heyduk et al., 2016; Mitchell et al., 2017; Wanke et al., 2017). The optimization of orchid probes to increase enrichment efficiency and the extension of the standard angiosperm targets into adjacent exons successfully improved the discriminative power in terms of orchid gene tree resolution. When working on shallow scales, there is the potential for extending anchor regions to produce longer loci that increases the chances of producing well-resolved gene trees. This approach has been successfully tested in salamanders and is currently under development in other non-model organisms (McCartney-Melstad et al., 2016). From the 446 loci alignments, two concatenated matrices were produced. One matrix

comprised of 31,905 bp and contained 305 loci with complete representation of taxa. The other matrix comprised of 444,631bp and contained 423 loci, including loci with missing sequences for one-two accessions (taxa). The remaining 23 loci alignments were not included in either matrix due to missing three or more accessions. We obtained 305 gene trees based on alignments without missing sequences and 118 gene trees from alignments with missing sequences. From the 305 gene trees, a total of 254 gene trees were derived from nuclear, 14 from mitochondrial and 37 from plastid datasets. The models of evolution calculated from each locus belonged to the General Time Reversible (GTR) family (supplementary material: Table 3).

4.3.2 Super tree estimation and supermatrix (concatenation)

The topology of the ASTRAL and ASTRAL-mlbs based on 305 loci was identicaL. Most of the nodes showed high support values of bootstrap and LPP. A total of 7,166,891 induced quartet trees were retrieved in the ASTRAL-305 species tree accounting for 65.34% of all quartet trees found in the species tree and for the ASTRAL-423, 63.18% of the quartet trees and 8,775,615 induced quartet trees. The MP-EST, NJst and STAR showed identical topologies compared to the ASTRAL analyses (Figs. 4.1 and 4.2A, B). The species tree analyses were congruent and recognized three main clades with high support (ASTRAL: LBS = 100, LPP ≥ 0.91): (1) *L. wendlandii*, (2) *L. horrida* and (3) *L. maxonii*. The *L. wendlandii* clade (1) clustered with the grouping of *L. horrida* (2) + *L. maxonii* (3) clades. Within the *L. horrida* clade (2), *L. horrida* clustered with *L. chameleon* and two accessions of the here described species *L. genetoapophantica* (DB8682 and DB9745). The two samples of *L. genetoapophantica* did not cluster together showing paraphyly; *L. genetoapophantica* (DB8682) was linked to *L. chameleon* and the other sample of *L. genetoapophantica* (DB9745) was linked in a more internal node to the two previous accessions (but with low support, ASTRAL: LPP < 0.56 and BS < 81). ASTRAL-305, MP-EST, NJst and STAR clustered all haplotype sets (1 and 2) from each sample (Fig. 4.2B–D). However, entire missing sequences in individual gene matrices caused discrepancies in the topology of gene trees, lower support (ASTRAL: LPP < 0.41 and BS < 15) and resulting species tree inferences in ASTRAL-423 and ML423 supermatrix approaches (BS < 84) (Fig. 4.2B–D). These incongruences were restricted to the *L. horrida* clade and the main discrepancy was the splitting of the two haplotypes and resulting paraphyly of *L. horrida*. All the inferences with missing and non-missing sequences recognized similar topologies for the *L. wendlandii* and *L. maxonii* clades (Fig. 4.2). The *L. maxonii* clade (3) contained *L. nymphalis* as a species related to a group made up of *L. maxonii* and two accessions of the undescribed species *L. amicitiae* (AK6144 and DB5911). Low LPP values are related to incongruences among gene trees (see further discussion Section 3.4.).

The ML supermatrix approaches retrieved essentially the same clades as the species tree analyses with the highest bootstrap support (BS) of 100% for most of the nodes (Fig. 4.2C and D). High support in inferences derived from NGS datasets are related to the markedly increasing supermatrix size (Wagner et al., 2013). Composition of the *L. wendlandii* clade (1) and *L. maxonii* clade (3) was similar in the topology of the ASTRAL (with 305 and 423 loci), ASTRAL-mlbs, MP-EST, NJst and STAR analyses, however, the most problematic clade for these analyses was again the *L. horrida* clade (2) because of the unexpected separation of the two haplotypes of *L. horrida* (DB11459). This separation was also observed in the ASTRAL-423 analyses (Fig.

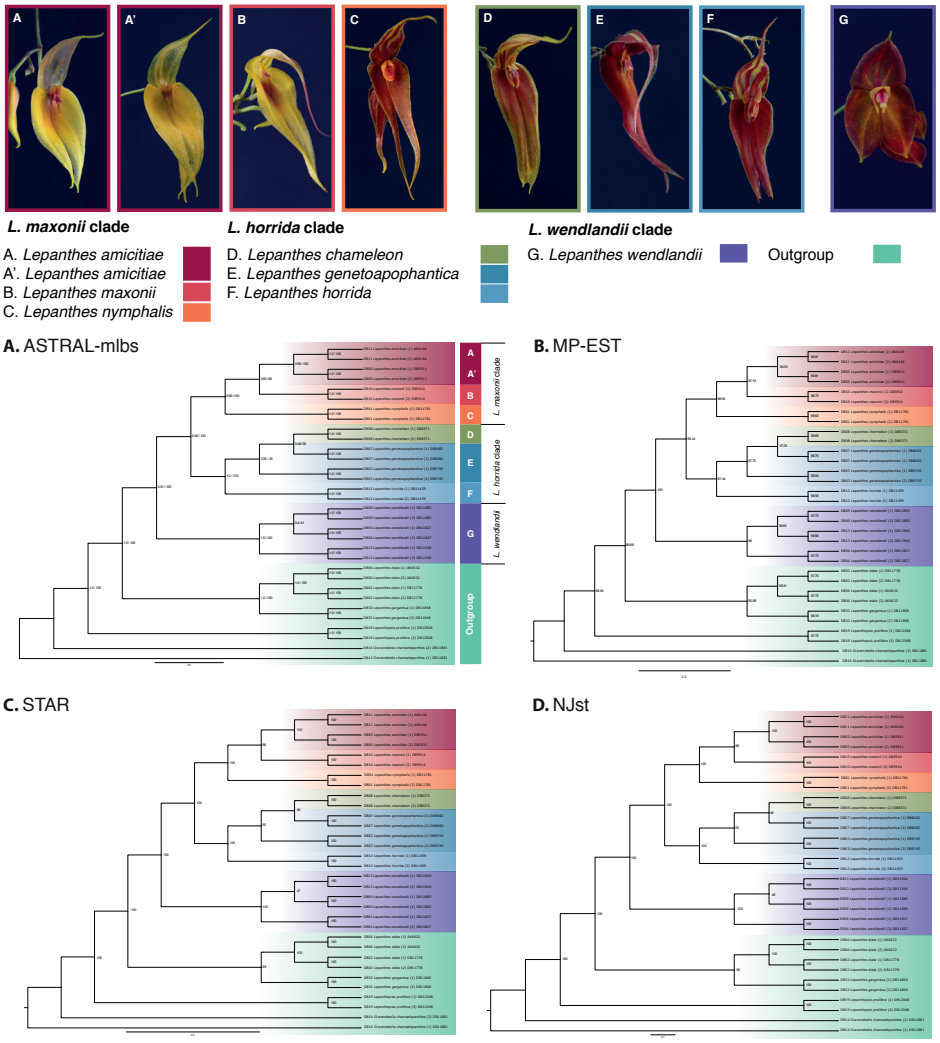

Figure 4.1. Flower morphology of the species of the *Lepanthes horrida* group and inferred species tree topologies from A. ASTRAL-mlbs, B. MP-EST, C. NJst and D. STAR. Local posterior probability/bootstrap support is shown for the nodes of ASTRAL-mlbs. These analyses support the clustering of three main clades: *L. maxonii* (**A.** A', *L. amicitiae*+ **B.** *L. maxonii*+ **C.** *L. nymphalis*) and *L. horrida* (**D.** *L. genetoapophantica*+ **E.** *L. chameleon*+ **F.** *L. horrida*) both sister to *L. wendlandii* (G)**.**

4.2B). Both ML supermatrix approaches including/excluding missing data did not group both haplotypes of *L. horrida* together (Fig. 4.2C-D). In the ML-305 supermatrix, the two samples of *L. genetoapophantica* were clustered but with low support (BS < 51%) and haplotype 1 of *L. horrida* was placed as sister to a clade made up of *L. genetoapophantica*, *L. chameleon* and haplotype 2 of *L. horrida* (Fig. 4.2C). In addition, lower bootstrap support values were observed for the internal nodes of the *L. maxonii* clade, in particular the node linking *L. nymphalis* with

L. maxonii + *L. amicitiae* (LBS = 60%). The support for this node was higher in the ML supermatrix including missing data (LBS = 97%). Also, in the *L. wendlandii* clade, one node linking two samples of *L. wendlandii* showed low bootstrap support (LBS = 61%) (Fig. 4.2C). In the ML-423 supermatrix haplotype 2 of *L. horrida* was grouped with *L. chameleon* (DB8371), BS = 84% and haplotype 1 with *L. genetoapophantica* (DB8682), BS = 93% (Fig. 4.2D). These were the only two nodes with bootstrap values less than 100% in the ML-423 analyses. In addition, the two samples of *L. genetoapophantica* were separated: *L. genetoapophantica* (DB9745) clustered with *L. chameleon* and to haplotype 2 of *L. horrida* whereas the other sample of *L. genetoapophantica* (DB8682) clustered with haplotype 1 of *L. horrida*. Similar to the species tree, the PhyloNet approach grouped together the haplotypes of *L. horrida* (DB11459), however, the two samples of *L. genetoapophantica* were not grouped together: *L. genetoapophantica* (DB9745) clustered with *L. chameleon* (DB8371) and *L. genetoapophantica* (DB8682) with *L. horrida* (DB11459) (Fig. 4.3A). This topology was similar to the tree inferred with ASTRAL-305, MP-EST, NJst and STAR. Similar as in the species tree analyses, the two samples of *L. amicitiae* were grouped together and *L. maxonii* ended up as closely related to this species. One sample of *L. wendlandii* (DB11827) did not cluster with the other two samples of *L. wendlandii* (DB11885 and DB11946). The network derived from SplitsTree did not cluster the two samples of *L. genetoapophantica* together and separated both haplotypes of *L. horrida* (DB11459) (Fig. 4.3B).

4.3.3 Species recognition

Species delimitations based on coalescent methods were consistent with the morphology of the species and agreed with previous species circumscriptions. The results also supported the recognition of two undescribed species and resolved the species relationships that were not previously disclosed using only nrITS and matK. Species tree estimations showed a strong tendency to recognize *L. wendlandii* as sister to *L. maxonii* + *L. horrida*. *Lepanthes wendlandii* is the most divergent species of the group. Unlike *L. maxonii* + *L. horrida*, plants of *L. wendlandii* are all characterized by short-pubescent, blackish ramicauls often longer than 20 cm, reddish flowers with the sepals widely ovate, obtuse and short-caudate and a cylindrical column (Luer, 2003a). Individuals of *L. wendlandii* show little morphological variation. Although in some analyses, one sample of *L. wendlandii* (DB11827) was positioned apart from the other two samples, the morphological evidence presented above suggests that the three samples belong to the same species. As suggested by by Pyron et al. (2016) assessments of species delimitations with computational genetic models should include traditional morphological data for species recognition so we refrain from recognizing additional taxa in this complex. The sister group of *L. wendlandii* comprises all the species related to the *L. maxonii* and *L. horrida* clades. It is characterized by plants with hirsute ramicauls, elongated, acuminate sepals and a column conspicuously flattened at the apex (Luer, 2003a). Within this group, the topology of the *L. maxonii* clade was constant in all the analyses. The clade clustered *L. nymphalis*, an endemic species to the Cordillera Central of Costa Rica with reddish flowers and distinctive for its long ciliated lip blades with two yellow-flowered species, *L. amicitiae* and *L. maxonii*, both endemic to the Cordillera de Talamanca between Costa Rica and Panama. Morphological differences suggest that individuals with yellow flowers correspond to two different species. One distinguished by the rounded shape of the upper lobe of

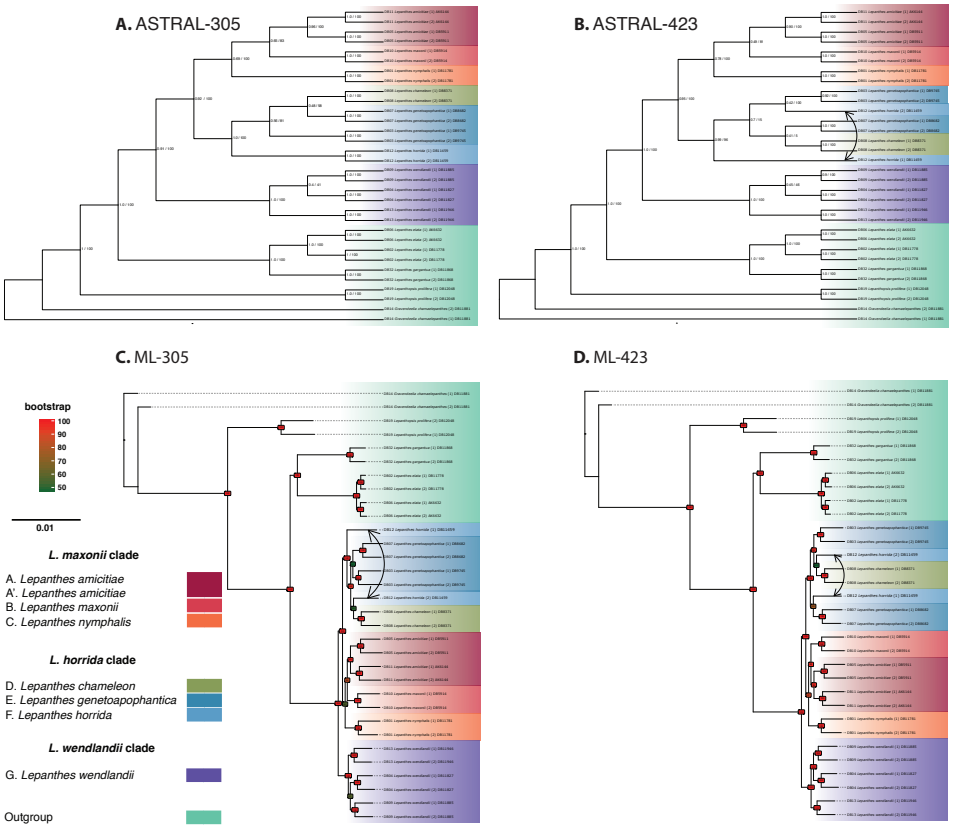

Figure 4.2. ASTRAL and Maximum likelihood (ML) inferences of concatenated datasets: **A.** ASTRAL of 305 loci without missing sequences. **B.** ASTRAL of 423 loci. **C.** ML-p based on 305 loci. **D.** ML-m based on 423 loci. Arrows in C and D show the splitting of the two haplotypes of *L. horrida* (DB11459), that do not cluster together. Note the lower BS support.

the petals (*L. amicitiae*) and other by the elongated lobes of the petals (*L. maxonii*). This hypothesis was supported consistently by ML and species tree analyses. Recognition of the *L. horrida* clade was constant in the grouping of *L. horrida*, *L. chameleon* and the undescribed species *L. genetoapophantica* (DB8682 and DB9745). Morphological evidence supports the recognition of these three species. However, unlike the *L. wendlandii* and *L. maxonii* clades, the topology of the *L. horrida* clade showed discrepancies in the positioning of *L. genetoapophantica* and both haplotypes of *L. horrida* and consequently low support BS in the ML analyses. Theoretically, in absence of ILS and gene flow, mutation is the only possible source of allelic variation. Therefore, haplotypes retrieved from the same sample are expected to be monophyletic. When haplotypes are not monophyletic in concatenated analyses (because concatenation does not take into account other gene evolutionary processes), it is possible that other sources of allelic variation operate such as gene flow or ILS (Pyron et al., 2016).

Hybrid origin may be one of the possible explanations to the nonmonophyly of species. *Lepanthes genetoapophantica* is morphologically similar to *L. horrida*, but it is distinguished by the smaller, divergent blades of the lip in contrast to the larger and elongated blades of *L. horrida*. The paraphyly of the two samples of *L. genetoapophantica* in the ASTRAL-305, MP-EST, NJst, STAR and Phylonet analyses supports a hybrid affinity or shared genetic diversity due to ILS or ancestral polymorphisms. This possible hybrid affinity is likely ancestral and not due to actual spontaneous hybridization because populations of *L. horrida*, endemic to the Cordillera Central of Costa Rica, are geographically isolated from the Cordillera de Talamanca where *L. chameleon* and *L. genetoapophantica* are endemic. In addition, there is no evidence of morphological variation between the characters that distinguish *L. horrida* and *L. genetoapophantica* that could suggest spontaneous hybridization. Thus, ancient hybridization could be a hypothesis for the discordant grouping of these species and might have contributed to speciation through the formation of new hybrid taxa (Abbott et al., 2013). Artificial hybridization in *Lepanthes* is possible but few natural hybrids have been documented probably because of the highly specialized pollination system (sexual mimicry).

An alternative hypothesis explaining the non-monophyly of *L. horrida* is polyploidy. Wanke et al. (2016) found that "assumed" diploidy in phased AHE alignments in *Aristolochia* yielded non-monophyletic allelic groupings from the same sample, however, upon assuming tetraploidy, these allelic groupings could be forced into monophyly in their analyses, suggesting that polyploidization could have occurred during the evolution of these species. Allopolyploidy occurs at high frequency in plants and can create postzygotic reproductive barriers in speciation events mediated by hybridization. Allopolyploids are common in the Orchidaceae and are expected to occur in the Pleurothallidinae as well because of the high intercompatibility among species and the many artificial hybrids created for commercial purposes. In Pleurothallidinae, polyploidy has been recorded in the genera *Octomeria* and *Scaphosepalum* (de Oliveira et al., 2015) but no data are yet available for *Lepanthes*. However, in absence of experimental evidence (i.e. data on genome size, chromosome counts, the relative success of artificial hybrid crossings, selfing and outcrossing), we refrain from favoring any of the current hypotheses that may explain the sources of reticulation in the *L. horrida* complex.

4.3.4 Concordance among gene trees/ASTRAL and ML trees

Even though the backbone nodes showed LBS = 100 in the supermatrix approach, the Phyparts analysis on the ASTRAL species tree and the ML supermatrix tree from the 305 gene trees showed a similar high degree of gene tree conflict (Fig. 4.4). In the ASTRAL-305 and ML-305 supermatrix tree (Fig. 4.4A and B), the only well supported clade in the gene trees was the *L. wendlandii* clade (1), supported by 257 (~84%) of the 305 loci tree topologies. The remaining 16% of the gene trees supported alternative topologies (Fig. 4.4A). In contrast, a very low number of gene trees (3.6%) supported the separation of the *L. horrida* (2) and *L. maxonii* (3) clades. This node showed dominance for other conflicting topologies (indicated in red in the pie charts of Fig. 4.4). The same pattern was shown in the nodes linking the species within the two groups. Low gene tree support and dominance of other conflicting bipartitions was also observed in the *L. maxonii* clade (3). *Lepanthes nymphalis* as sister to *L. maxonii* and *L. amicitiae* was support-

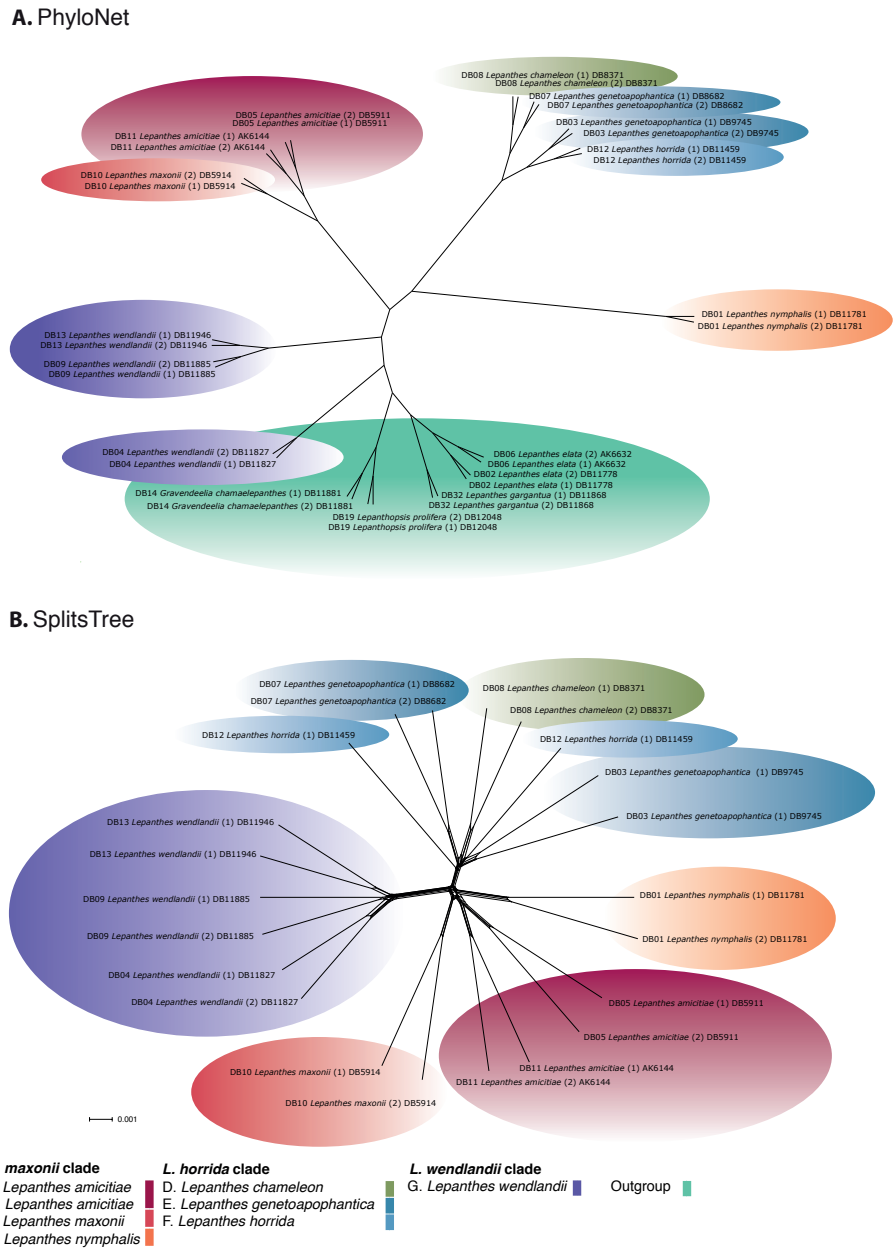

Figure 4.3. Inferred species network analyses of **A.** PhyloNet approach showing a similar clustering of species compared to the species tree analyses but grouping the two samples of *L. genetoapophantica* separately: one grouped with *L. horrida* and the other with *L. chameleon*. The affinity of *L. maxonii* and the two samples of *L. amicitiae* is also evident **B.** SplitsTree network showing the non-monophyly of *L. horrida*, the separate clustering of *L. genetoapophantica* and the networking on the three clades.

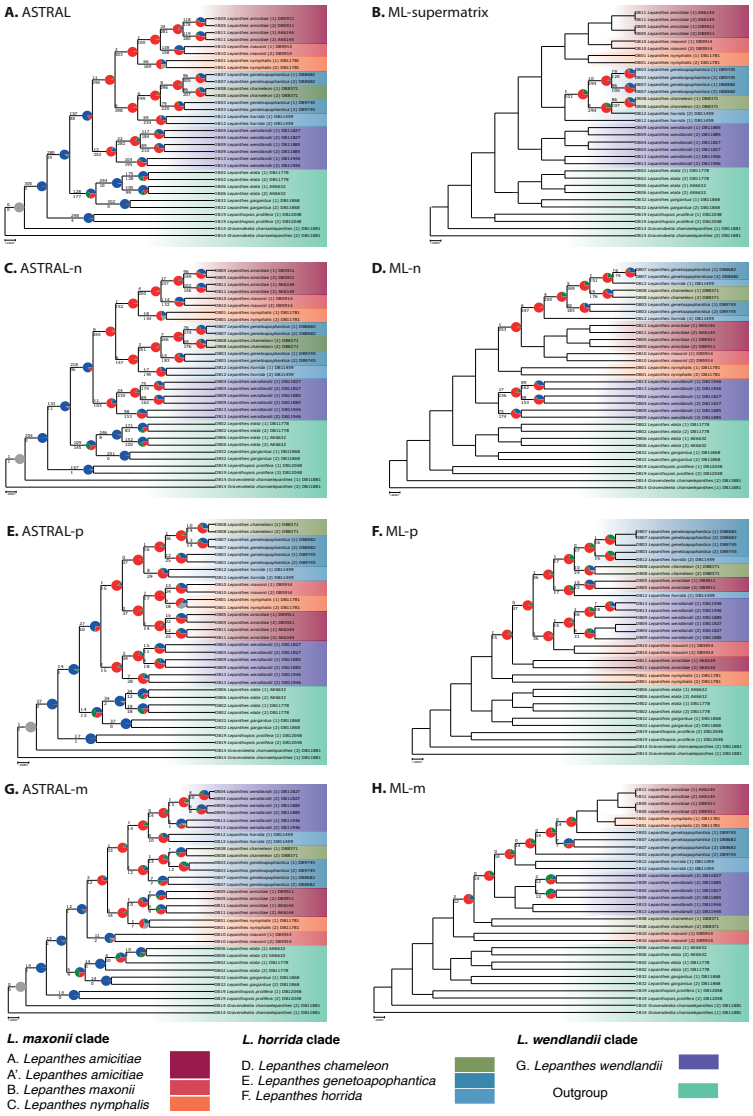

Figure 4.4. ASTRAL and ML species trees inferred from 305 ML gene trees and genomic analyses: **A.** ASTRAL-mlbs. **B.** ML-concatenated. **C.** ASTRAL-nuclear. **D.** ML-nuclear (based on 288 nuclear loci) **E.** ASTRAL-plastid. **F.** ML-plastid (based on 40 plastid loci). **G.** ASTRAL- mitochondrial. **H.** ML-mitochondrial (based on 18 mitochondrial loci). Numbers on branches represent the gene trees supporting each node (top) and the number of gene trees in conflict with the shown topology of the species tree (bottom). Pie charts show the proportion of gene trees concordant with the shown topology (blue), conflict with the shown topology that support the main alternative for that clade (green), other dominant alternative (conflicting) supported bipartitions (red) and unsupported nodes due to conflicting bipartitions with less than 70% bootstrap support (gray). Pie charts are shown only where the topology differs in the ML with respect to the ASTRAL analyses.

ed only by two (~1%) gene trees, *L. maxonii* as sister to *L. amicitiae* by only five (1.6%), and only 24 (7.8%) gene trees supported the grouping of the two samples of *L. amicitiae* in a single clade. The nodes of the *L. maxonii* and *L. horrida* clades mostly showed alternative (conflicting) bipartitions (indicated in red in the pie charts of Fig. 4.4). Increasing support for the shown topologies was observed in the tip nodes grouping all haplotypes together (indicated in blue in the pie charts of Fig. 4.4), however, conflicting bipartitions were still observed (dominance of red and to a lesser extent green in the pie charts in Fig. 4.4). The topology of *L. horrida* clade (2) was the main difference between the ASTRAL and ML supermatrix approach (Fig. 4.4A and B). In the ASTRAL analyses, only six (1.9%) gene trees supported the relationship between *L. horrida* and *L. chameleon* + *L. genetoapophantica* and six (1.9%) the clustering of *L. genetoapophantica* (DB9745) and *L. chameleon* + *L. genetoapophantica* (DB8682). The grouping of the latter species was supported by eight (2.6%) gene tree topologies (Fig. 4.4A). In the ML supermatrix, the separate clustering of haplotype 1 of *L. horrida* was supported by six (1.9%) gene trees and the clustering of haplotype 2 with *L. chameleon* was supported by nine (~3%) gene trees (Fig. 4.4B). In addition, the nodes showed higher dominance of other main topologies (indicated in green in the pie charts in Fig. 4.4) as compared to ASTRAL.

Similar to other studies on animals and plants in which the performance of multi-locus datasets was evaluated, we found that the analyses of multiple gene copies do not necessarily result in concordance or high support of the topologies obtained with coalescent-based methods of species tree estimations and individual gene trees (Jeffroy et al., 2006; Sun et al., 2015; Tang et al., 2015). On the contrary, individual gene trees with divergent topologies are common in many groups, suggesting that hybridization, horizontal gene transfer, gene duplication and ILS are pervasive phenomena and could be important causes of these topological discordances (Jeffroy et al., 2006; Mallo and Posada, 2016; Sun et al., 2015; Yu et al., 2013).

Incomplete lineage sorting might cause discordances in groups of closely related species with rapid diversifications in part because the alleles within a population do not have enough time to coalesce (Degnan and Rosenberg, 2009; Tsutsumi et al., 2016). Recent studies in the evolution of the Pleurothallidinae revealed that *Lepanthes*, with an estimation of over 1200 species, radiated in the last 2.5 million years (Pérez-Escobar et al., 2017a). In addition, Tremblay and Ackerman (2001) found that genetic drift is important in population differentiation due to small population size and restricted gene flow common in *Lepanthes*. Therefore, ILS could be a plausible explanation for the incongruences observed as well, next to hybridization and polyploidy.

Individual gene tree clustering based on RF and Kendall-Colijn distance showed a wide array of topologies and consequently little resolution in the topologies as compared to species tree and concatenated methods (Fig. 4.5A and B). The species tree analyses clusters produced in group six showed similar topologies according to RF based clustering (Fig. 4.5C and D). Only 17 gene trees (16 nuclear and one plastid) clustered in the same group of the supermatrix while the remaining 288 gene trees (94.4%) and the species tree analyses clustered separately, thus showing other topologies (Fig. 4.5C). The topologies of both ML concatenated supermatrix approaches were slightly divergent with respect to ASTRAL, MP-EST, NJst and STAR species trees (Figs. 4.1 and 4.5D). The most notable topological difference among them was the separation of the haplotypes of *Lepanthes horrida* (DB11459) and the two samples of *L. genetoapophantica* as discussed above.

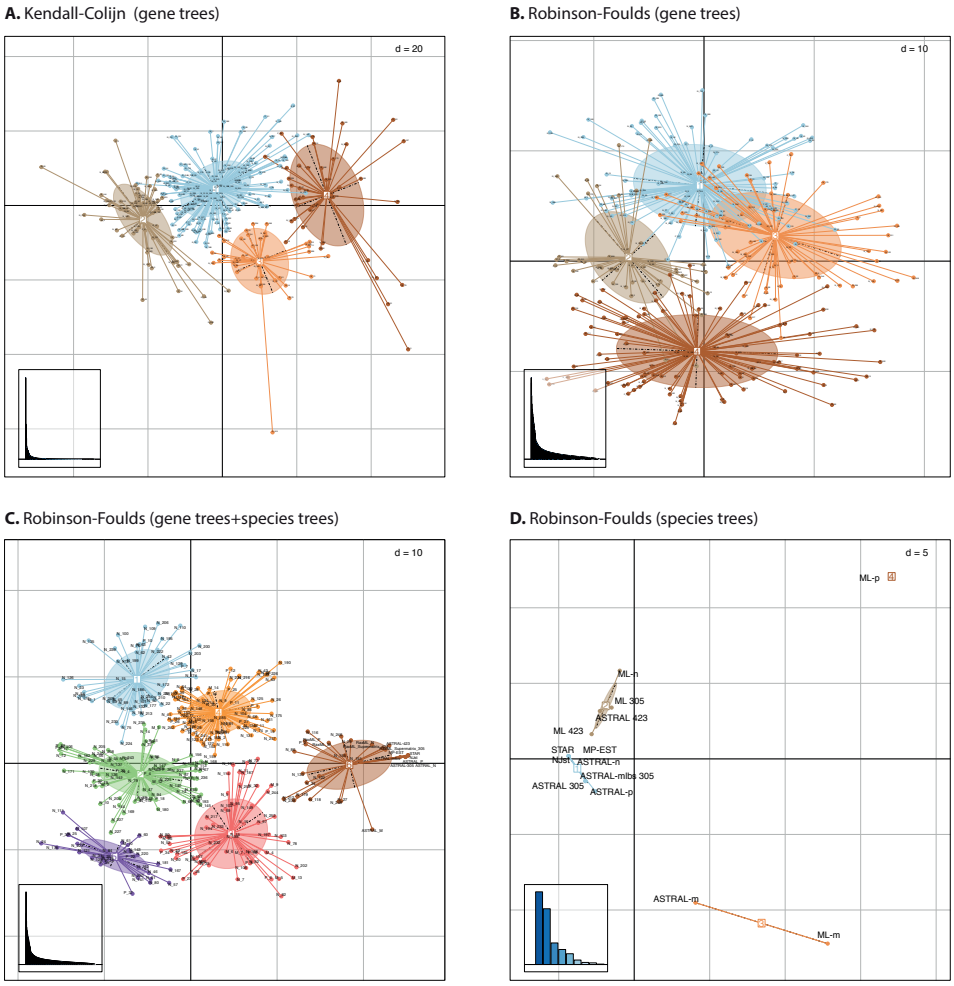

Figure 4.5. Gene trees and species trees cluster analysis of Metric Multidimensional Scaling (MDS) **A.** Clustering of the 305 gene trees (rooted) based on the Kendall-Colijn metric vector and four clusters. M=mitochondrial, N=nuclear, P=plastid. **B.** Clustering of the 305 gene trees (unrooted) based on Robinson-Foulds (RF) symmetric difference. **C.** Clustering of the 305 gene trees, species (ASTRAL, MP-EST, NJst, STAR) trees and ML concatenated datasets based on RF. Cluster 6 contains the species trees and ML trees, (d) MDS of species trees and ML concatenated datasets based on RF.

In addition to ILS and deeper speciation, incongruence could also be the result of estimation errors in gene trees derived from alignments containing missing sequences, long-branch attraction or phylogenetic noise (Mallo and Posada, 2016; Mirarab and Warnow, 2015). Excluding loci alignments with missing sequences indeed produced different topologies. The topology of the ASTRAL-423 (inferred from alignments with missing sequences in less than two samples) was similar to the topology of the ML supermatrix approaches and ML-n (Figs. 4.2 and 4.4D–5) mostly because it recovered a non-monophyletic *L. genetoapophantica* and *L. horrida* (Figs. 4.1

and 2A). In contrast, ASTRAL-305 grouped these haplotypes together similar to the other species tree analyses (Figs. 4.1 and 5D), possibly because according to Huang and Lacey Knowles (2016), information is lost because of reduction of matrix size and biased representation of mutations when missing data are excluded.

4.3.5 Nuclear and organellar datasets

As observed in the species tree and supermatrix approaches, a low number of gene tree topologies supported the topologies of the species trees based on the nuclear/organellar datasets (Fig. 4.4C–H). The gene trees from each genomic dataset did not form specific conglomerates because the topologies were mixed without showing any pattern that could be linked to a unique genomic origin (Fig. 4.5A–C). Plastid and nuclear derived gene trees were observed across the six clusters and the mitochondrial derived gene trees were only absent in cluster six where the species trees and ML analyses were placed (Fig. 4.5C). The topology of ASTRAL-n was similar to ASTRAL-305 and to the other species tree analyses (Fig. 4.5D) but a low number of gene trees supported this topology (Fig. 4.4A, C and 4.5C, D). In contrast, ML-305 differed from ML-n in the separation of the two samples of *L. genetoapophantica* in ML-n, which were grouped together in ML-305 but with low support and a low number of gene tree topologies supporting (BS < 54% in ML-305 and BS < 62% in ML-n) the internal branches of the *L. horrida* cade (Figs. 4.2C and 4.6A). ASTRAL-n and ML-n differed in the placement of both haplotypes of *L. horrida* because ASTRAL-n clustered them together contrary to ML-n. However, both agreed on the topology of the *L. maxonii* clade.

The analyses based on the mitochondrial datasets ASTRAL-m and ML-m (Figs. 4.4G–H and 4.5D) and the plastid ML-p dataset (Figs. 4.4F and 4.5D) were the most divergent with respect to the ML-supermatrix, ML-n, ASTRAL-n, ASTRAL-p and all the species tree analyses (Fig. 4.5D). These trees showed very low bootstrap support for most of the internal branches of the three main clades (BS < 70%) (Fig. 4.6C and D). All ML analyses failed in clustering the two haplotypes of *L. horrida* (DB11459) together. Although the ML-m analyses grouped both haplotypes of *L. horrida* together, BS was low and this analysis also failed in clustering the two haplotypes of *L. genetoapophantica* (DB9745) together (Fig. 4.6C). The ASTRAL-m analyses showed the most divergent species tree topology (Figs. 4.5C, D and 4.6C) as compared to ASTRAL-n and ASTRAL-p (Fig. 4.5C and D), which were more similar to the ASTRAL, MP-EST, NJst and STAR species trees (Fig. 4.1). However, this topology was supported by a very low number of gene trees (< 3 gene trees, 16.6%). In addition, the topology of the ML-m analyses was divergent as well, the main differences being the placement of *L. maxonii* and *L. chameleon* as sister to the rest of the species of the group and the separation of the haplotypes of one sample of *L. genetoapophantica* (DB9745). These topologies were poorly supported by individual gene trees (Fig. 4.4H). In contrast, the ASTRAL-p topology was more similar to the species tree analyses. It recognized the same clades with the same topology for the *L. wendlandii* and *L. horrida* clades, but a different topology for the *L. maxonii* clade (3), because the two *L. amicitiae* samples clustered together and they were grouped with a clade formed by *L. nymphalis* and *L. maxonii*. However, the alternative topology of the ASTRAL-p analyses was supported by a low number of gene trees (< 3 gene trees, 7.5%) (Fig. 4.4E). The paraphyly of *L. horrida* was the main difference between the ASTRAL-p and ML-p analyses.

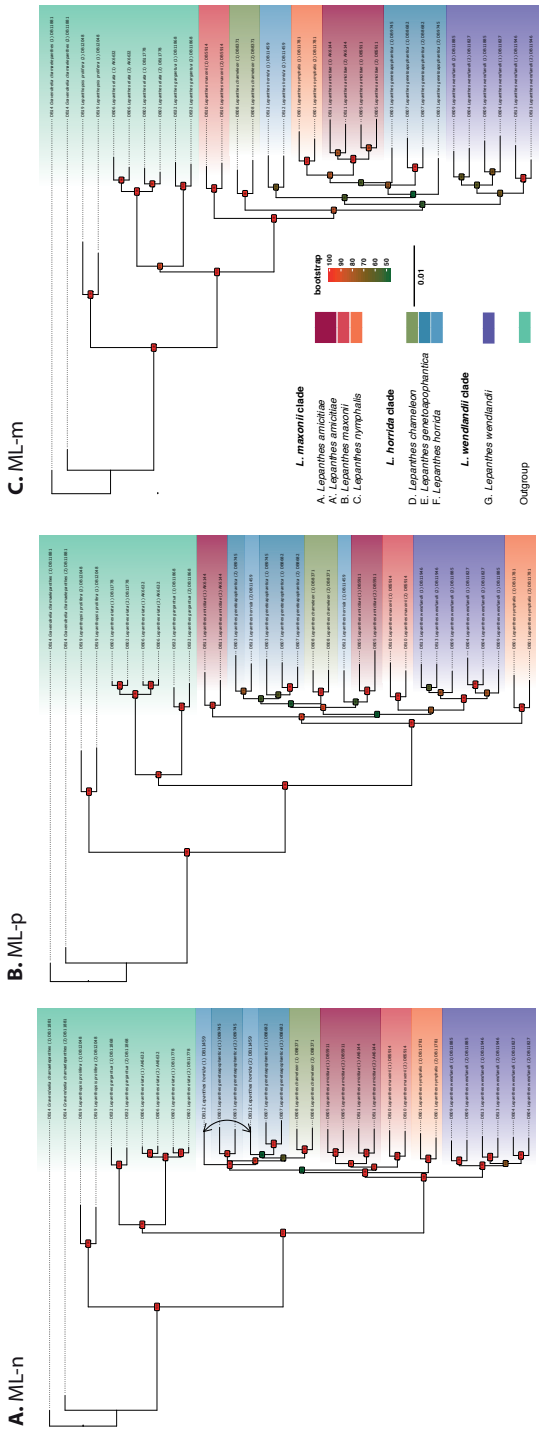

Figure 4.6. Maximum likelihood (ML) inferences of concatenated datasets showing bootstrap support for each node. **A.** ML-n based on 288 nuclear loci. **B.** ML-p based on 40 plastid loci,. **C.** ML-m based on 18 mitochondrial loci.

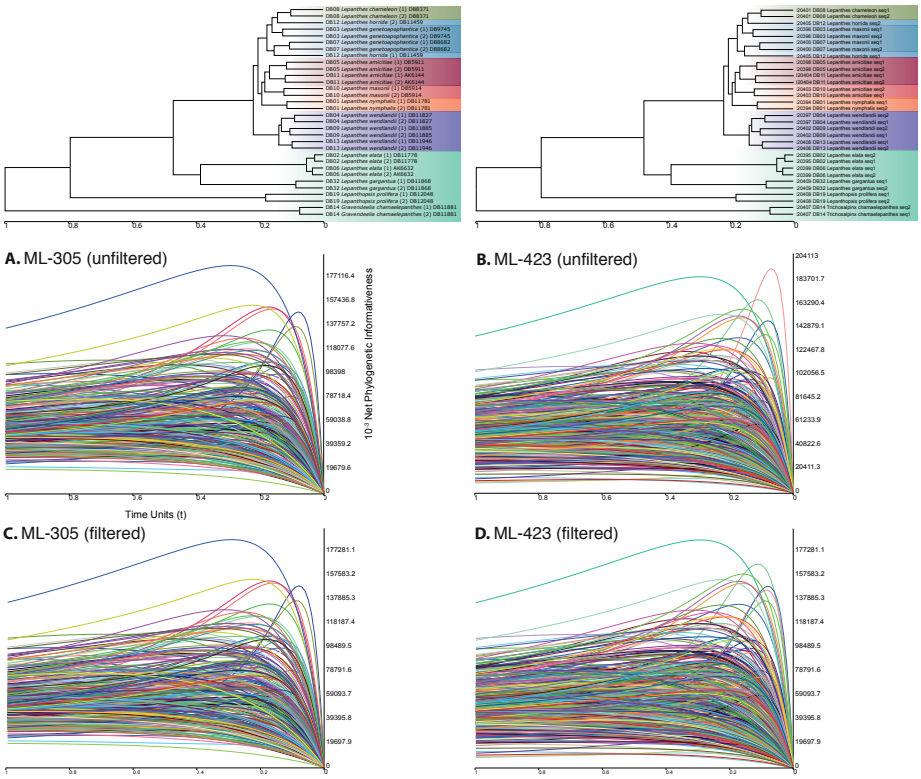

Figure 4.7. Net phylogenetic informativeness profiles per each locus of the ML-305 and ML-426 datasets. Ultrametric trees were obtained with PATHd8 using a relative time scale (0 to 1): **A.** ML-305. **B.**ML-423, **C.** Filtered ML-305. **D.** Filtered ML-423. Curves are smoother in unfiltered and filtered datasets. The analyses recovered the same topology and the BS values slightly increased in the filtered datasets.

The greater number of nuclear genes obtained likely produced a stronger influence in the topology of the concatenated species tree analyses that was almost identical to the inferences based on nuclear genes. Incongruences between nuclear and plastid datasets might also suggest hybridization. Studies in Rosidae showed conflicts in the topology derived from plastid, nuclear and mitochondrial datasets likely produced by ILS and ancient hybridization (Sun et al., 2015). The high incongruence observed in our results among individual nuclear loci (if not related to estimation/stochastic errors) suggests that conflict among nuclear and organellar datasets is due to biological evolutionary events such as ILS and/or ancient hybridization. Mitochondrial derived trees showed disparate topologies that disagree with the morphology of the species. Because our results were based on a few mitochondrial gene trees only, these could be misleading due to undersampling (Parks et al., 2017).

4.3.6 Phylogenetic informativeness

Species net phylogenetic informativeness plots showed slightly increasing, stable curves over time in most of the loci from both ML approaches (Fig. 4.7A and B). A total of 75% of the

loci reached a maximum net PI between 78.03 and 44.33 at a reference time (t) between 0.30 and 0.81. The obtained unfiltered datasets showed plots with smooth curves lacking "phantom" spikes and the filtering method (with rate values > 5) detected only six loci with high substitution rates (Fig. 4.7C and D). In addition, filtered and unfiltered datasets recovered identical topologies and almost the same BS values (but slightly higher in the filtered datasets). These findings are similar to those found in AHE datasets from Aristolochiaceae and Lamiaceae (Fragoso-Martínez et al., 2016; Wanke et al., 2017).

The ten individual loci with the highest PI values for each analysis are shown in Table 4. They were derived from all three separate genomes (despite the dominance of nuclear genes in our datasets), thus highlighting the importance of organellar loci in phylogenetic analyses. The fragment with the highest net PI was the inhibitor of Bruton tyrosine kinase like (IBTK) gene, of which 25,98bp were analyzed. Phylogenetic analyses of *Lepanthes*, (Pleurothallidinae in general), have been based mostly on nrITS and *mat*K only, so the recovery of the new loci published here would be useful for future molecular systematic studies in this group (Pridgeon et al., 2001).

4.4 Conclusions

Anchored hybrid enrichment coupled with coalescence-based methods is a powerful tool to solve complicated phylogenetic relationships in linages derived from recent, rapid diversifications. Despite the high discordance in the topology of the gene trees reconstructed, combined ASTRAL, MP-EST, NJst and STAR analyses could resolve the phylogenetic relationships of the *L. horrida* species group. These analyses also disclosed two undescribed species, *L. amicitiae* and *L. genetoapophantica*. These results could not have been obtained by morphology and standard nrITS and matK analyses. Phenomena such as ILS, hybridization and polyploidy may be common in groups recently diverged such as Pleurothallidinae causing discordance among datasets. The data presented here showed high incongruences in the topologies among individual loci that were probably produced by different biological phenomena. Due to the various sources of incongruence, species delimitations based on multi-locus datasets should be interpreted in conjunction with traditional morphological observations. Only with a large number of innovative phylogenetic markers generated from three different genomes, the phylogeny could be fully resolved and this enabled us to separate traits evolving in parallel or convergently across these orchid lineages, such as flower color and size, from real diagnostic traits such as the shape and orientation of the lobes of the petals and lip.

4.5 Taxonomic treatment

Key to the species of the Lepanthes horrida group

1. Sheaths of the ramicaul shortly pubescent; synsepal broadly ovate-orbicular, the free apices obtuse — ***L. wendlandii***

1”. Sheaths of the ramicaul densely ciliate-hirsute; synsepal narrowly lanceolate, the free apices linear-acuminate — 2

2. Margins of the sepals ciliate-dentate — 3

3. Synsepal glabrous at the base; lateral lobes of the lip long ciliate-hispid along the margins — ***L. nymphalis***

3”. Synsepal hirsute at the base; lateral lobes of the lip glabrous — ***L. chameleon***

2”. Margins of the sepals glabrous — 4

4. Lateral sepals fused almost to the apex, yellow, striped with red; lateral lobes of the lip rose-purple, large, covering the column almost to the apex — ***L. horrida***

4”.Lateral sepals fused just to the middle of their length or less, solid red or yellow, sometimes with a basal reddish blotch, but never striped; the lateral lobes of the lip yellow to orange, small, only covering the basal part of the column — 5

5.Flowers red; upper lobes of the petals long, narrowly linear, acute; lower lobe of the petals as long as the upper lobe; blades of the lip diverging at apex — ***L. genetoapophantica***

5”.Flowers yellow; upper lobes of the petals short, elliptic to rounded, sub truncate; lower lobe of the petals three times longer as the upper lobe; blades of the lip parallel at apex — 6

6. Upper lobes of the petals oblong-elliptic, acute, subtruncate; lower lobe of petals subequal to the upper lobe (elliptic); appendix shorter than the conectives of the lip — ***L. maxonii***

6”. Upper lobes of the petals rounded, lower lobe of the petals three times longer as the upper lobe (filiform); appendix longer or as long as the connectives of the lip — ***L. amicitiae***

Table 4.4. The 10 loci with the best performance (highest PI values) in both ML-305 and ML-423 analyses.

Dataset	Loci ID	Product	Maximum PI value	Max. PI value at time (t)	Length (bp)	Genome
RaxML 423						
	T272_L184	inhibitor of Bruton tyrosine kinase (IBTK)	65.38	0.99	2598	Nuclear
	T272_L45	histidine biosynthesis bifunctional protein hisIE (HISN2)	53.76	0.85e	1595	Plastid
	T272_L277	tRNA:m(4)X modification enzyme TRM13 (TRMT13)	48.78	0.44	1197	Nuclear
	T272_L317	transcription factor bHLH140 (BHLH140)	47.44	0.61	1592	Nuclear
	T272_L217	uracil-DNA glycosylase (UNG)	47.31	0.56	695	Mitochondrial
	T272_L8	Ribosomal-protein-alanine-acetyltransferase (Rps13)	45.20	0.75	873	Nuclear
	T272_L162	chromatid cohesion protein (DCC1)	45.11	0.95	1783	Nuclear
	T272_L55	cytochrome c biogenesis protein CCS1 (Ccs1)	45.00	0.67	1950	Plastid
	T272_L248	Uncharacterized	44.88	0.79	1272	uncharacterized
	T272_L420	yrdC domain-containing protein (YRDC)	43.35	0.73	1296	Mitochondrial
RaxML 305						
	T272_L184	inhibitor of Bruton tyrosine kinase (IBTK)	67.58	0.99	2598	Nuclear
	T272_L45	histidine biosynthesis bifunctional protein hisIE (HISN2)	57.82	0.77	1595	Plastid
	T272_L217	uracil-DNA glycosylase (UNG)	50.20	0.56	695	Mitochondrial
	T272_L55	cytochrome c biogenesis protein CCS1 (Ccs1)	50.14	0.59	1950	Plastid
	T272_L162	chromatid cohesion protein (DCC1)	47.98	0.88	1783	Nuclear
	T272_L248	uncharacterized	47.58	0.74	1272	uncharacterized
	T272_L420	yrdC domain-containing protein (YRDC)	46.10	0.68	1296	Mitochondrial
	T272_L286	PCNA domain-containing protein (PCNA)	45.17	0.77	879	Nuclear
	T272_L365	calcium sensing receptor (CAS)	44.83	0.94	1674	Plastid
	T272_L352	glycosylphosphatidylinositol anchor attachment 1 protein (GPAA1)	44.04	0.94	1607	Nuclear

4.5.1 *Lepanthes amicitiae* Bogarín & Pupulin, Molec. Phylogen. Evol., 129: 40–43. 2018., Figs. 4.8F and 9A.

Type: Costa Rica-Panama. Puntarenas-Bocas del Toro: Coto Brus-Valle del Risco, línea fronteriza sobre la divisoria de aguas ingresando por el camino de la Finca Sandí-Hartmann "El Capricho", 8°57'12.34"N 82°43'32.69"W, 2154 m, bosque pluvial montano bajo, 11 diciembre 2013, *A. P. Karremans 6144, D. Bogarín, M. Fernández & L. Sandoval* (holotype: JBL).
Diagnosis: This species is similar to *Lepanthes maxonii* Schltr. but it differs in the rounded upper lobe of the petals (vs. oblong-elliptic) and linear-acuminate lower lobe (vs. elliptic), subfalcate, convergent lobes of the lip (vs. straight. divergent) and hirsute appendix (vs. pubescent).
Description: Epiphytic, densely caespitose, erect herb, up to 30 cm tall. Roots slender, filiform, glabrous, to 1 mm in diameter. Ramicauls ascending to erect, 10–25 cm long, enclosed by 5–22 lepanthiform sheaths, long-ciliate along the thickened ribs, dilated at apex into a horizontal, ovate, acute ostia with densely ciliate margins. Leaf erect, subcoriaceous, elliptic, acute, emarginated with a short apiculous at apex, 3.0–3.8 cm long, 1.8–2.3 cm wide, the base cuneate into a petiole ca. 2 mm long. Inflorescence a dense, distichous, successively many-flowered (up to 20 or more flowers) raceme to 8 cm long, borne by a filiform peduncle 4.0–4.8 cm long. Floral bract broadly ovate-triangular, cucullate, obtuse, sparsely glandular, ca. 2 mm long. Pedicel terete, glabrous, 6–8 mm long. Ovary terete-subclavate, the intralocular ridges along the veins thickened into low, rounded crests, ca. 4.0 mm long. Flowers relatively large for the genus, the dorsal sepal widespread, yellow-hyaline, the veins solid yellow; the laterals sepals yellow, often with a large rose-red blotch at the base; the petals yellow to orange, sometimes purple along the mid vein and flushed purple along the lower lobe; the lip orange to orange-pink; the column rose-purple. Dorsal sepal ovate, slightly concave, acute, long acuminate, glabrous, 6.0–6.5 × 2.0–2.3 mm, connate to the lateral sepals for about 2 mm. Lateral sepals partially fused at the base into a bifid, lanceolate synsepal, 7.0–7.4 × 2.8–3.0 mm, connate for about 5 mm, the free apices narrowly acute-subacuminate. Petals transversely bilobed, linear-subfalcate, 0.5–0.7 × 3.5–3.7 mm; the upper lobe shorter and broader than the lower lobe (ca. 1mm long), rounded,the lower lobe linear-acuminate, subsigmoid. Lip 3-lobed, the lateral blades elliptic, rounded, subfalcate, 0.6–0.7 mm long, held parallel to the column, the connectives cuneate to subrectangular, connate to the column near the middle, the appendix hirsute, slender, ligulate. Column hemiterete, flattened at apex into elliptic, rounded wings, 1.2 mm long; the anther dorsal, the stigma ventral Pollinia 2, narrowly obpyriform-complanate, on an elliptic, orange brown viscidium.
Distribution and ecology: Endemic to the Cordillera de Talamanca between south-eastern Costa Rica and western Panama in montane cloud forests at 2100-2500 m.
Etymology: From the Latin *amicitia*, friendship, in allusion to La Amistad (The Friendship) International Park, a protected area which spans over southeastern Costa Rica and western Panama, where the type specimen was collected, and alluding to the friendship among the researchers of the University of Costa Rica and the University of Chiriquí, who are linked by a long-term, common floristic project.
Discussion: This species is mostly closely related to *L. maxonii*, another species with similar yellow sepals and a red blotch at the base of the synsepal. Rudolf Schlechter described *L. maxonii* from Cerro de Horqueta, Chiriquí, Panama from a collection by R.W. Maxon in 1911. The type

specimen was destroyed in the herbarium B during the Second World War, however, the drawing based on the holotype specimen shows the oblong-elliptic upper lobe of the petals, which differs from the rounded, suborbicular lobe of *L. amicitiae*.

Additional specimens examined: Costa Rica-Panama. Puntarenas-Bocas del Toro: Coto Brus-Valle del Risco, línea fronteriza sobre la divisoria de aguas ingresando por el camino de la Finca Sandí-Hartmann "El Capricho", 8°57'12.34"N 82°43'32.69"W, 2154 m, bosque pluvial montano bajo, 11 diciembre 2013, fl. 8 Jan. 2014, *D. Bogarín et al. 10751* (JBL). PANAMA. Chiriquí: Bugaba, Las Mirandas, Las Nubes, ca. 3 km al noroeste de Cerro Punta, Parque Internacional La Amistad, 2500 m, colectada por E. Olmos, setiembre 2008, floreció en cultivo en Finca Drácula, Chiriquí, Guadalupe, Panamá, 10 diciembre 2008, *D. Bogarín 5911* (JBL).

4.5.2 *Lepanthes chameleon* Ames, Schedul. Orchid. 4: 28. 1923. Figs. 4.8D and 4.10C.

Type: Costa Rica: near Cartago, *C. H. Lankester s.n*. (holotype: AMES; detail of type: AMES).

Distribution and ecology: Endemic to the Cordillera de Talamanca, Costa Rica in montane cloud forests at 2200-2700 m.

Etymology: From the Latin *chameleon*, "a ground lion", a group of lizards called *chameleon*s (Chamaeleonidae) which are able to change their skin coloration under certain circumstances. Oakes Ames noted that the purple color of young flowers of *L. chameleon* fades away in mature flowers, hence the comparison.

Discussion: This species is closely related to *L. genetoapophantica* and *L. horrida*, but it is easily distinguished by the ciliate-dentate sepals and the hirsute synsepal (mostly at the base) contrasting with the entire, glabrous sepals of *L. genetoapophantica* and *L. horrida*.

4.5.3 *Lepanthes genetoapophantica* Bogarín & Gravend., Molec. Phylogen. Evol. 129: 44. 2018., Figs. 4.8C and 4.10D.

Type: Costa Rica. Puntarenas: Coto Brus, Sabalito, Zona Protectora Las Tablas, 15 km al noreste de Lucha, Sitio Tablas, Finca Sandí-Hartmann "El Capricho", camino a El Surá, 8°57'0.63"N 82°44'59.72"W, 2017 m, bosque pluvial montano bajo, 10 diciembre 2013, *D. Bogarín 10644, A. Karremans, M. Fernández & L. Sandoval* (holotype: JBL).

Diagnosis: This species is similar to *Lepanthes horrida* but it differs in the linear-subfalcate petals, the yellowish, diverging, sub-trapezoid lobes of the lip, the lower apices of the lip not reaching the anther cap, the appendix extending beyond the lower apex of the lip and the truncate apex of the column.

Description: Epiphytic, densely caespitose, erect herb, up to 25 cm tall. Roots slender, filiform, glabrous. Ramicaus ascending to erect, stout, 8–13 cm long, enclosed by 4–12 lepanthiform sheaths, long papillous-ciliate along the thickened ribs, dilated at apex into an oblique, ovate, acuminate ostia with densely ciliate margins. Leaf erect, coriaceous, ovate to elliptic, acute, emarginated with a short apiculous at apex, 3.5–5.5 cm long, 1.5–2.5 cm wide, the base cuneate into a distinct petiole 2–3 mm long. Inflorescence a dense, distichous, successively many-flowered (up to 40or more flowers) raceme to 12 cm long, borne by a filiform peduncle 2–3 cm long. Floral bract broadly ovate-triangular, cucullate, acute, glabrous, ca. 2 mm long. Pedicel terete,

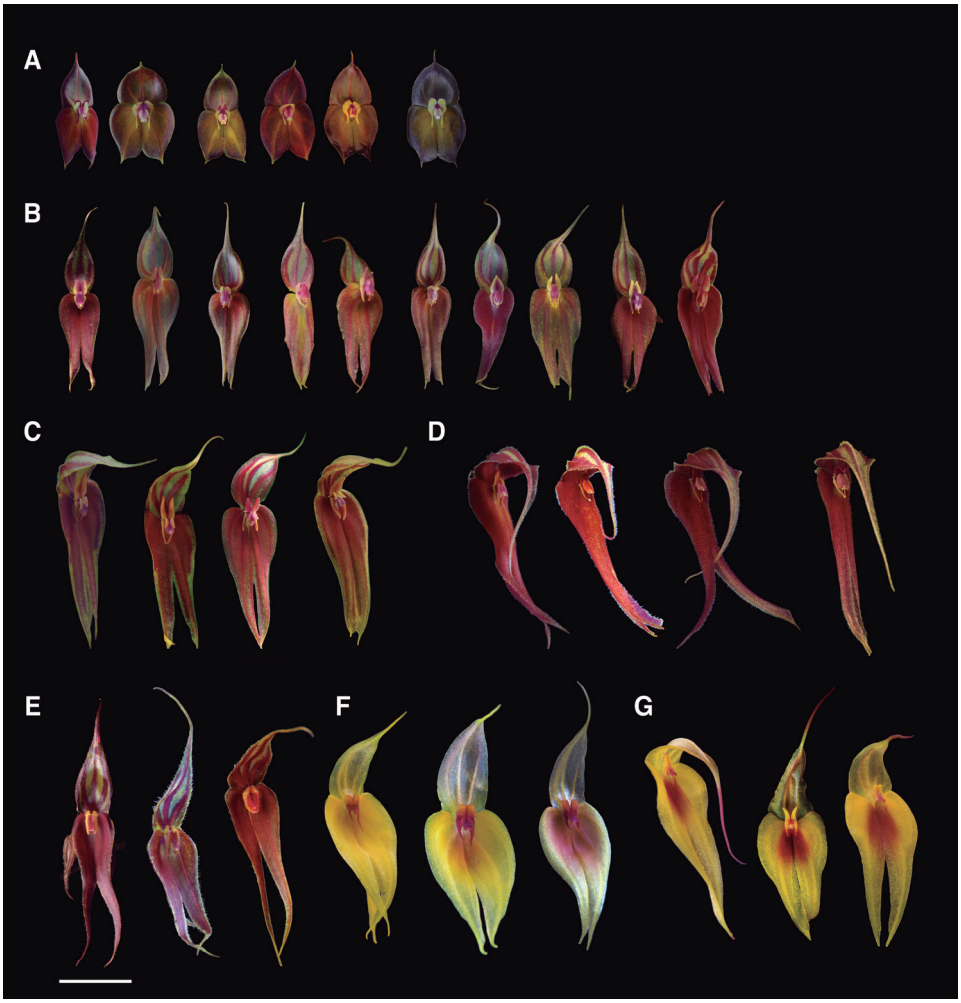

Figure 4.8. Flower morphology and variation among individuals of **A.** *Lepanthes wendlandii*, **B.** *L. horrida*, **C.** *L. genetoapophantica*. **D.** *L. chameleon*. **E.** *L. nymphalis*. **F.** *L. amicitiae*. **G.** *L. maxonii*.

glabrous, 5 mm long. Ovary terete-subclavate, the intralocular ridges along the veins provided with low, semi hyaline-cartilaginous crests, 3.5–4.0 mm long. Flowers relatively large for the genus, the dorsal sepal widespread, yellow, suffused with purple along the veins and the margins; the laterals red, edged in yellow on the external margin toward the apex; the petals yellow, suffused with orange-red at the base; the lip yellow; the column rose-purple. Dorsal sepal elliptic, concave, acute, long acuminate, glabrous, 1.2–1.4 × 0.4–0.6 cm, connate to the lateral sepals for about 1 mm. Lateral sepals fused at the base into a bifid, lanceolate synsepal, 1.2–15 × 0.5–6 cm, connate for about 8 mm, the free apices narrowly acute. Petals transversely bilobed, linear-subfalcate, 0.8–1.0 × 3.7–4.5 mm, the outer margin between the lobes provided with a minute, rounded apiculum; the upper lobe shorter and broader than the lower lobe, linear, rounded, the

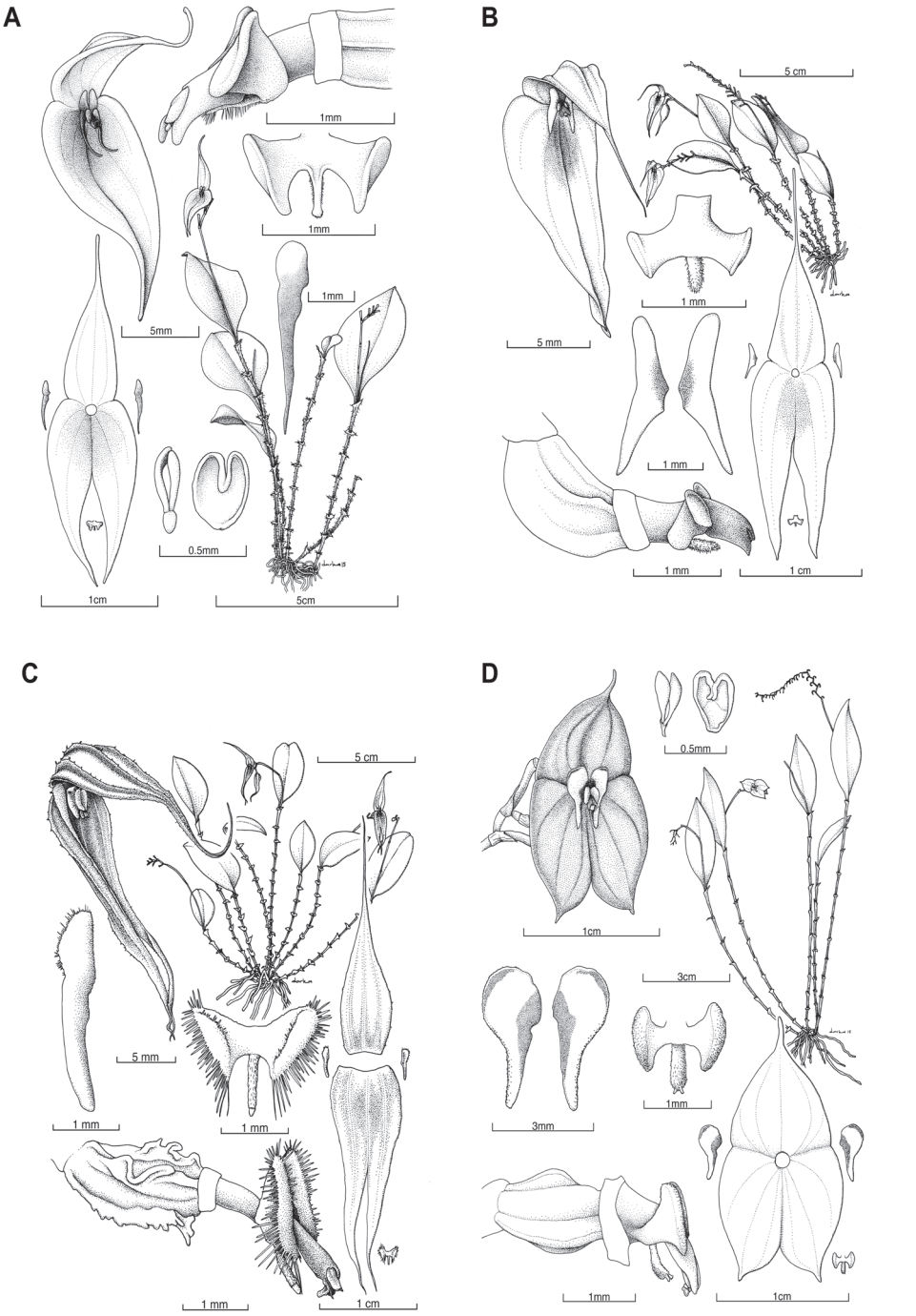

Figure 4.9. Composite-line drawings of **A.** *L. amicitiae* (Bogarín 10751). **B.** *L. maxonii* (Bogarín 5914). **C.** *L. nymphalis* (Bogarín 8307). **D.** *L. wendlandii* (Pupulin 6711).

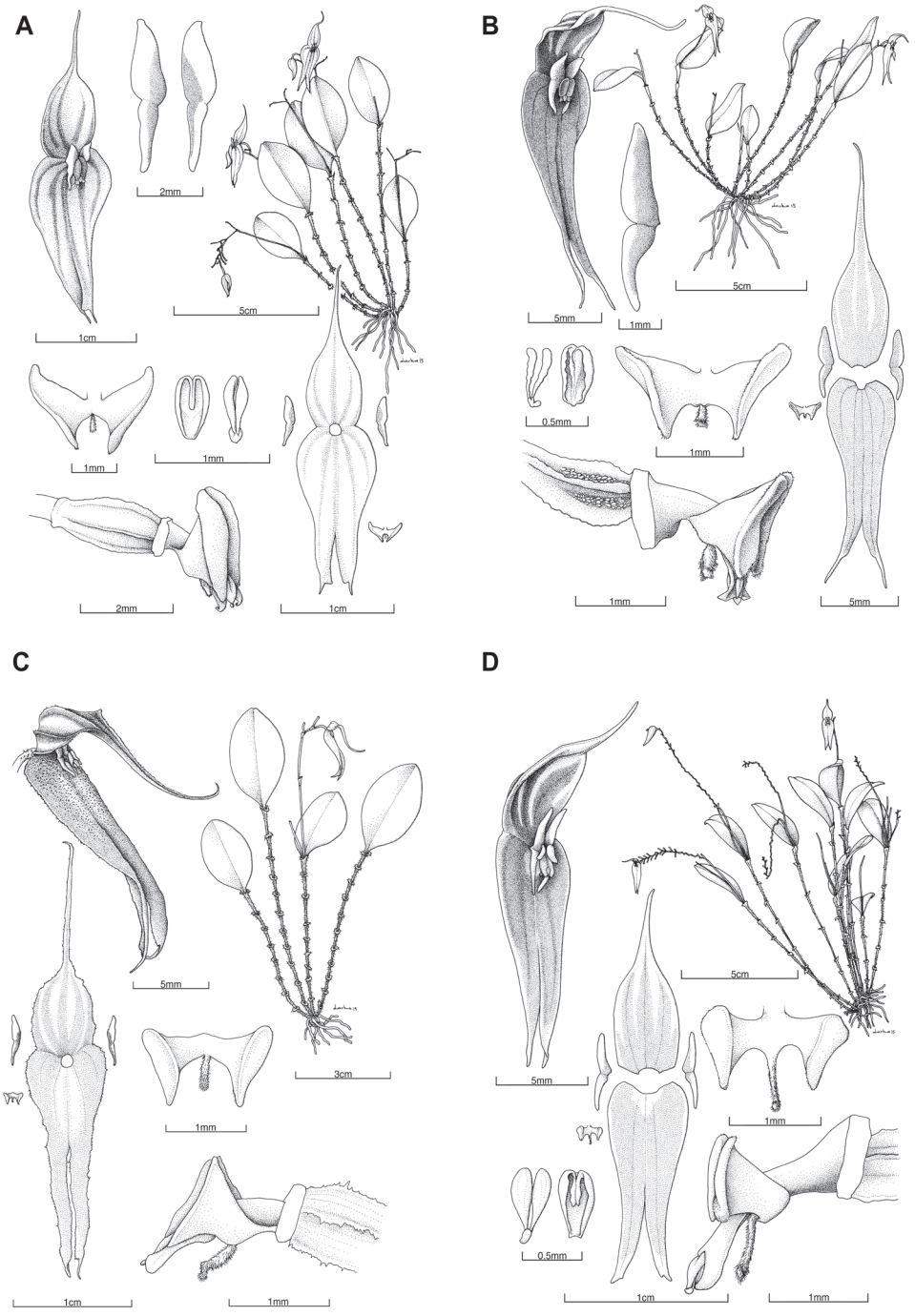

Figure 4.10. Composite-line drawings of **A.** *L. horrida* (Bogarín 272). **B.** *L. horrida* (Bogarín 11489). **C.** *L. chameleon* (Pupulin 4277). **D.** *L. genetoapophantica* (Bogarín 10644).

lower lobe subfalcate, acuminate, rounded at the apex. Lip 3-lobed, the lateral blades subtrapezoid, rounded, ca. 1 mm long, the apices diverging when erect, the connectives cuneate, connate to the column near the middle, the appendix pubescent, slender, ligulate, with an apical subquadrate gland. Column hemiterete, flat, dilated at apex into elliptic, acute wings, ca. 2 mm long; the anther dorsal, the stigma ventral Pollinia 2, narrowly obpyriform, sub attenuate at the base, on a elliptic, orange brown viscidium.

Distribution and ecology: Endemic to the Cordillera de Talamanca between south-eastern Costa Rica and western Panama in montane cloud forests at 2183-2624 m.

Etymology: Named after the Greek words γένεσις (genesis), origin, generation, and ἀποφαίνω (apophaino), to make visible, in reference to the genetic work, carried out with the aid of next generation sequencing techniques, that revealed the hidden identity of this species among its relatives.

Discussion: This species is closely related to *L. chameleon* and *L. horrida*, all with similar red flowers. It was confused by us with *L. maxonii* however, after studying the type material we realized that it corresponds to the yellow-flowered species most closely allied to *L. amicitiae*. Therefore, we proposed it here as a new species. From similar *L. horrida* it differs in the linear-subfalcate petals (vs. ovate, erect), the yellowish, diverging, sub-trapezoid lobes of the lip (vs. pink, parallel, ovate-elliptic), the lower apices of the lip not reaching the anther cap (vs. reaching the anther cap), the appendix extending beyond the lower apex of the lip (vs. shorter, not extending) and the truncate apex of the column (vs. cleft). From *L. chameleon* it differs in the glabrous, entire sepals (vs. hirsute, denticulate).

Additional specimens examined: Costa Rica. Limón: Talamanca: Bratsi, Parque Internacional La Amistad, Valle del Silencio, camio del refugio hacia el jardín (Turbera), orillas del Río Terbi, bosque pluvial montano, 2471 m, 9°07’05.12”N 82°57’40.95”W, 15.08.2012, *D. Bogarín et al. 9817* (JBL); same collection data, *D. Bogarín et al. 9842* (photo-JBL); Bratsi, Parque Internacional La Amistad, Valle del Silencio, camino del refugio hacia el jardín (Turbera) antes de cruzar el Río Terbi, bosque pluvial montano, 2411 m, 9°07’45.53”N 82°57’31.23”W, 18.9.2014, *A. Karremans et al. 6395* (JBL). Puntarenas: Buenos Aires, Buenos Aires, Olán, de la falda noreste del Cerro Tinuk hacia la falda sureste de Cerros Utyum, 9°17’33.9” N 83°09’47.5” W, 2587 m, bosque pluvial montano bajo, epífitas en bosque primario, 26 julio 2012, *D. Bogarín et al. 9745* (JBL); Buenos Aires, Buenos Aires, Olán, de la falda noreste del Cerro Tinuk hacia la falda sureste de Cerros Utyum, 9°17’37.1” N 83°09’40.1” W, 2624 m, bosque pluvial montano bajo, epífitas en bosque primario, 26 julio 2012, *D. Bogarín et al. 9752* (JBL); Puntarenas-Chiriquí: Coto Brus-Renacimiento, línea fronteriza hacia el Cerro Pando, después del mojón N.338, 8°55’11.22”N 82°43’18.18”W, 2446 m, bosque muy húmedo montano bajo, epífitas en bosque primario, “*in sylvis virginis versus montium Pando in itinere ad summum Costa Rica austro-orientalis in finibus utrimque Costa Rica et Panama*”, 19 abril 2011, *D. Bogarín et al. 8682* (photo). PANAMA. Chiriquí: Bugaba, Cerro Punta, Parque Internacional La Amistad, sendero Las Nubes, Mirador La Nevera, 8°54’00.3”N 82°37’12.8”W, 2436 m, bosque pluvial montano bajo, epífitas en *Podocarpus* sp., J & L. Harrison, Z. Samudio & Z. Serracín, 25 febrero 2014, florecieron en cultivo, 3 marzo 2014, *D. Bogarín 10974* (UNACHI). Renacimiento, Santa Clara, Cotito, camino a la divisoria de la sierra, 8°53’57.7”N 82°42’07.8”W, 2183 m, bosque pluvial montano bajo, epífitas en bosque secundario, 6 marzo 2014, *D. Bogarín et al. 10986* (UCH);

Guadalupe, camino de Finca Drácula al Parque Internacional La Amistad, 1200 m, epífitas en bosque secundario a orillas del camino, en cultivo Finca Drácula, 12 diciembre 2006, *D. Bogarín 2966 & R.L. Dressler* (JBL); Bugaba, Cerro Punta, Guadalupe, 2000 m, planta colectadas por Erick Olmos & A. Maduro, sin más datos de recolecta, en cultivo en Finca Drácula, 19 diciembre 2008, *D. Bogarín 5961* (JBL).

4.5.4 *Lepanthes horrida* Rchb.f., Beitr. Orchid.-K. C. Amer. 91. 1866. Figs. 4.8B and 4.10A-B.

Type: [Alajuela-Heredia]: Desengaño in Costa Rica, 9 May 1857, *H. Wendland s.n.* (holotype: W; illustration of type: AMES).
Distribution and ecology: Found in Costa Rica in secondary forests at elevations of 1500 to 2500 m. along the Cordillera Volcánica Central and Cordillera de Tilarán.
Etymology: Even though the Latin word *horridus* commonly refers to dreadful, horrible, or inspiring fear. However, the adjective bears the alternative meaning of bristly, referring to the stiff trichomes covering the ramicauls of this species.
Discussion: This species is closely related to *L. genetoapophantica* but it differs in the ovate, erect petals (vs. linear-subfalcate), pink, parallel, ovate-elliptic lobes of the lip (vs. diverging, sub-trapezoid) and the shorter appendix, not extending beyond the lower apex of the lip (vs. extending far beyond the lobes of the lip).

4.5.5 *Lepanthes maxonii* Schltr., Repert. Spec. Nov. Regni Veg. 12: 204. 1913. Figs. 4.8G and 9B.

Type: Panama. An Gaumstämmen in feuchten Wäldern zwischen Alto de Las Palmas und dem Gipfel des Cerro de Horqueta (Chiriqui), 2100-2268 m, blühend im Mar 1911, *W.R. Maxon 5494* (holotype: B, destroyed; isotypes: NY, US; illustrations of type: AMES-100633, AMES-100634).
Distribution and ecology: Endemic to Panama around Cerro Horqueta, Chiriquí in the Cordillera de Talamanca at 2100-2268 m in cloud forests.
Eponymy: Named after William Ralph Maxon (1877–1948), American botanist, who worked for the United States National Museum, a part of the Smithsonian Institution.
Discussion: This species is closely related and morphologically similar to *L. amicitiae*, both having similar yellow flowers. However, *L. maxonii* differs in the oblong-elliptic upper lobe of the petals (vs. rounded, suborbicular in *L. amicitiae*), the lower lobe of the petals being subequal to the upper lobe (vs. filiform) and the appendix shorter than the connectives of the lip (vs. longer or as long as the connectives).

4.5.5 *Lepanthes nymphalis* Luer, Phytologia 54: 357. 1983. Figs. 4.8E and 4.9C.

Type: Costa Rica. Heredia: epiphytic in cloud forest, Alto Gallito, alt. 2000 m, beyond the pass north of El Castillo, 21 June 1981, *C.A. Luer & J. Luer 6356* (holotype: SEL; isotype: CR).
Distribution and ecology: Endemic to the Cerro Delicias and Alto Gallito in the Cordillera Volcánica Central, Heredia, Costa Rica at around 2000 m. It grows epiphytically in cloud forests.

Etymology: From the Latin *nymphalis*, "of a nymph, a mythological woodland deity," referring to the dark, mossy, wooded habitat of the species.
Discussion: This is a very distinctive species related to *L. amicitiae* and *L. maxonii*, from which it differs in the red flowers and the long ciliate-hispid blades of the lip.

4.5.6 *Lepanthes wendlandii* Rchb.f., Beitr. Orchid.-K. C. Amer.: 91. 1866. Figs. 4.8A and 4.9D.

Type: Vulkan de Barba in Costa Rica, 11 Jul 1857, *H. Wendland s.n.* (holotype: W; illustration of type: AMES).
Distribution and ecology: This species is found in oak cloud forest along the Cordillera Volcánica Central and the Cordillera de Talamanca in Costa Rica and Panama at 2200-2800 m.
Eponymy: Named after the German botanist, collector and gardener Hermann Wendland (1825–1923), from the Herrenhauser Gardens in Hannover, Germany.
Discussion: This is the most divergent species of the group characterized by the glabrous sheaths of the ramicaul and the broadly ovate-orbicular synsepal with free obtuse apices (not elongated as in the other related species).

Appendix A. Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ympev.2018.07.014.