Received: 14 July 2025 | Accepted: 27 March 2026

DOI: 10.1002/ajb2.70213

RESEARCH ARTICLE

Botany

Evidence of introgression amid phylogenetic conflict in *Brachyotum*, a plant radiation from the Tropical Andes

Diego Paredes-Burneo[1,2] | **Ana M. Bedoya**[1,3] | **Fabián A. Michelangeli**[3] | **Lucas C. Majure**[4] | **Carmen Ulloa Ulloa**[5] | **Laura P. Lagomarsino**[1]

[1]Shirley C. Tucker Herbarium, Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803, USA

[2]Departamento de Dicotiledóneas, Museo de Historia Natural UNMSM, Av, Arenales, Lima 1256, Peru

[3]Center for Biodiversity & Evolution, The New York Botanical Garden, Bronx, NY 10458-5126, USA

[4]University of Florida Herbarium, Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA

[5]Missouri Botanical Garden, 4344 Shaw Blvd., St. Louis, MO 63110, USA

Correspondence
Diego Paredes-Burneo, Shirley C. Tucker Herbarium, Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803 USA.
Email: diegop.francob@gmail.com

Funding information
Division of Environmental Biology, Grant/Award Numbers: DEB 2002270, DEB-2001357, DEB-0818399, DEB-2055525; National Geographic, Grant/Award Number: NGS Grant # 829107; International Association for Plant Taxonomy; American Society of Plant Taxonomists; Society of Systematic Biologists

Abstract

Premise: The species-rich flora of Tropical Andes underwent multiple rapid and recent diversifications, yet resolving their evolutionary histories remains challenging despite increasing phylogenomic data. Here, we examined phylogenomic conflict in *Brachyotum* (Melastomataceae) to identify sources preventing its resolution.

Methods: We sequenced multiple nuclear loci through target capture and plastid loci through genome skimming for approximately two thirds of *Brachyotum* species. We analyzed nuclear loci with concatenated maximum-likelihood, multispecies coalescent (MSC), and phylogenetic network approaches, whereas plastid loci were analyzed with a concatenated maximum-likelihood.

Results: Topological conflict was widespread among gene trees. *Brachyotum* was inferred as monophyletic by concatenated nuclear and ASTRAL approaches, but as paraphyletic by BEAST and phylogenetic network analyses, the latter involving four introgression events. Plastid analyses also recovered paraphyly. Across analyses, the backbone of *Brachyotum* showed low branch support, and individual nuclear loci yielded markedly divergent topologies across inference methods.

Conclusions: Despite pervasive gene tree discordance, the signal for multiple introgression events is robust. Cytonuclear discordance between plastid and nuclear data further supports introgression. Our results suggest that current phylogenomic approaches may not fully capture the evolutionary complexity of *Brachyotum*, a challenge that may extend to other rapid Andean plant radiations.

KEYWORDS

Andean diversification, gene tree discordance, genome skimming, interspecific hybridization, Melastomataceae, phylogenomics, target capture,

ABSTRACT IN SPANISH

Premisa: La flora de los Andes tropicales, rica en especies, está comprendida por múltiples linajes que se diversificaron rápida y recientemente; sin embargo, determinar la historia evolutiva de dichos linajes sigue siendo un reto a pesar del creciente volumen de datos filogenómicos. En este estudio, examinamos el conflicto filogenómico en *Brachyotum* (Melastomataceae) para identificar las causas que impiden su resolución.

Métodos: Se secuenciaron múltiples loci nucleares mediante captura por sondas y loci plastidiales mediante barrido genómico para aproximadamente dos tercios de las especies de *Brachyotum*. Se analizaron los loci nucleares con métodos de máxima verosimilitud bajo concatenación, coalescencia multiespecífica (MSC) y red filogenética, mientras que los loci plastidiales se analizaron con máxima verosimilitud bajo concatenación.

Resultados: Se encontró un alto nivel de conflicto entre topologías de árboles de genes. Se infirió con loci nucleares que *Brachyotum* era monofilético mediante métodos de concatenación y ASTRAL, pero parafilético según BEAST y la red filogenética, este último con cuatro eventos estimados. Los análisis de plástidos también recuperaron parafilia. En todos los análisis, la columna de *Brachyotum* tiene un bajo soporte de ramas y cada loci nuclear mostró topologías marcadamente distintas según el método filogenético usado.
Conclusiones: Pese al conflicto entre topologías de genes, la señal de introgresión es clara. La discordancia citonuclear presente entre árboles nucleares y plastidiales es también consistente con introgresión. Nuestros resultados sugieren que los enfoques filogenómicos usados actualmente no estarían captando la complejidad evolutiva de *Brachyotum*, un desafío que puede extenderse a otros casos de radiaciones andinas de plantas.

PALABRAS CLAVE

barrido genómico, captura por sondas, conflicto filogenético, discordancia entre árboles de genes, diversificación en los Andes, filogenómica, hibridación interespecífica, Melastomataceae

Understanding the origin of clades with high species richness across the tree of life is central to evolutionary biology. Whether these so-called evolutionary radiations (Simões et al., 2016) are the result of ancient (Whitfield and Lockhart, 2007) or recent events (Marques et al., 2019) or involve ecological divergence (e.g., adaptive radiations; Givnish, 2015), their study often entails a challenge to phylogenomics. Evolutionary radiations, especially when rapid (Givnish, 2015), are characterized by successive speciation events over short time spans, which can be strongly impacted by incomplete lineage sorting (ILS; sometimes referred to as deep coalescence; Maddison, 1997) and introgression (Whitfield and Lockhart, 2007), two phenomena that complicate phylogenomic inference (Maddison, 1997; Degnan and Rosenberg, 2006, 2009; Mendes and Hahn, 2018). Incomplete lineage sorting arises when a fraction of the ancestral allele polymorphism, with an evolutionary history incongruent with the underlying species tree, is inherited by one of the resulting lineages during speciation (Maddison, 1997), thus producing different evolutionary histories between genes and species (Degnan and Salter, 2005; Degnan and Rosenberg, 2006, 2009). Introgression, which results from interspecific hybridization with subsequent backcrossing, also yields gene tree discordance due to the transfer of already-diverged alleles from one species to another (Linder and Rieseberg, 2004; Mallet, 2005; Yu et al., 2012). It is generally challenging to tease apart the effects of ILS and introgression because their resulting patterns in trees may not be distinguishable (Linder and Rieseberg, 2004; Yu et al., 2012; Hibbins and Hahn, 2022).

These challenges to plant phylogenomics are observed across plant taxa in the Andes, an area of the world with a high proportion of evolutionary radiations (Guevara-Andino et al., 2025). The dynamic geological and climatic history of the Andean region has likely contributed to the complex phylogenomic patterns in its flora. Mountain formation promoted diversification in many plant taxa by isolating populations, thus increasing rates of allopatric speciation (Nevado et al., 2018; Flantua et al., 2019; Sosa-Pivatto et al., 2020; Vargas et al., 2023). Additionally, while mountains were gaining elevation, new ecosystems (e.g., high elevation grasslands) emerged, creating ecological opportunity for speciation for both the local flora and recent migrants in the youngest habitats (Pennington et al., 2010). Among global montane systems, Andean uplift is very recent, especially in the northern reaches of the mountain chain, which were still uplifting in the last couple million years and remain geologically active (Boschman, 2021; Pérez-Escobar et al., 2022). Compounding the impact of dramatic mountain uplift, the warm periods of Pleistocene glacial cycles promoted isolation of high montane taxa (Flantua et al., 2019). Together, the geological and climatic context created a perfect set of conditions for rapid plant radiations in the recent past (Dellinger et al., 2024). However, the cooler periods, associated with greater connectivity of high montane ecosystems, likely promoted secondary contact of the already isolated populations, leading to potential subsequent gene flow between incipient species (Nevado et al., 2018). These gene flow events may have thinned species boundaries, likely allowing allele interchange between different closely related species (Nevado et al., 2018; Sosa-Pivatto et al., 2020). The combination of orogeny and glacial oscillations might have generated constant changes in population size and gene flow, which may explain the complex evolutionary histories of high Andean taxa.

Given the complex geological and climatic history of the Andes, it is not surprising that difficulties in obtaining well-resolved and well-supported phylogenies exist in many Andean-centered plant evolutionary radiations. These difficulties persist even in the face of deep genomic sampling (e.g., Vargas et al., 2017; Meerow et al., 2020; Lagomarsino et al., 2022; Aguirre-Santoro et al., 2024; Frost et al., 2024; Segovia et al., 2025). Incomplete lineage sorting is widespread across Andean plant lineages (e.g., Morales-Briones et al., 2018; Pouchon et al., 2018; Murillo-A. et al., 2022), likely a result of short intervals between speciation events. Interspecific gene flow also contributes to the poorly resolved phylogenies, leaving traces such as cytonuclear discordance (e.g., Vargas et al., 2017; Pouchon et al., 2018; Lagomarsino et al., 2022) and paralogous sequences (e.g., Vargas et al., 2017; Nevado et al., 2018; Pouchon et al., 2018; Murillo-A. et al., 2022; Frost et al., 2024). Introgression is likely facilitated by many mechanisms, including

secondary contact after allopatric speciation (e.g., Nevado et al., 2018) and the occurrence of many closely related species diverging in ecological traits (e.g., Pouchon et al., 2021). Hybridization between species sometimes led to polyploidy, with allopolyploidy documented in some Andean plant clades (e.g., Morales-Briones et al., 2018, 2021).

Here we examined the evolutionary history of an Andean plant radiation, the genus *Brachyotum* (DC.) Triana (Melastomataceae Juss: Melastomateae Bartling), whose 55 species are restricted to the high Andes of Colombia, Ecuador, Peru, Bolivia, and northern Argentina (Wurdack, 1953; Veranso-Libalah et al., 2022). *Brachyotum* occurs on a high-elevational belt of the Andes, with most species occurring around 3000 m a.s.l. (Wurdack, 1953; Meyer et al., 2021), an elevation where ecosystems are younger than the surrounding lower areas (Pennington et al., 2010). *Brachyotum* has an easily distinguishable morphology compared to other melastomes. Its nodding flowers have a pseudotubular corolla and staminal nectaries coupled with generally smaller leaves than its relatives, providing a unique combination of traits among high Andean Melastomataceae (Wurdack, 1953; Dellinger et al., 2022). Its closest relatives possess common floral traits for melastomes: open or reflexed corollas, lack of nectaries, and exposed stamens that are usually dimorphic with prolonged connectives (Veranso-Libalah et al., 2022). Most members of the family are pollinated by bees and bumblebees, but *Brachyotum* is likely mainly pollinated by hummingbirds (Renner, 1989; Dellinger et al., 2022). Understanding the origin and evolution of these traits and their association with transitions to higher elevations and bird pollination motivated this study, which was conducted within a robust phylogenetic framework.

Brachyotum is hypothesized to be a "natural unit" based on its unique morphology (Wurdack, 1953), but molecular phylogenetics studies (Michelangeli et al., 2013; Guimarães et al., 2019; Meyer et al., 2021) suggest that it may not be monophyletic. Despite the lack of clarity around the monophyly of the genus, all sampled *Brachyotum* species form a well-supported "*Brachyotum* + allies" clade with a handful of non-*Brachyotum* Melastomateae species also found in the high Andes (Meyer et al., 2021). Some of these allies have been transferred to *Chaetogastra*, a genus that occurs across the Neotropics, based on phylogenetic evidence (Guimarães et al., 2019). The remaining allies have repeatedly shown unstable phylogenetic positions and are treated as incertae sedis under the genus *Tibouchina*, an unsatisfactory solution, because it is unclear what their relationship is to the current *Tibouchina* s.s. (Guimarães et al., 2019; Meyer et al., 2021). The impact of gene tree discordance on these relationships remains to be determined.

Phylogenetics using few loci have been insufficient to resolve the relationships within "*Brachyotum* + allies" (Michelangeli et al., 2013; Guimarães et al., 2019; Meyer et al., 2021). To address this knowledge gap, we used a target capture phylogenomic approach and appropriate evolutionary models to capture the expected complexity of this evolutionary radiation. Target capture is a method that isolates low- or single-copy genes sampled from a broad subset of the nuclear genome to improve overall phylogenetic signal and to facilitate the use of coalescent approaches to infer species trees (McKain et al., 2018; Johnson et al., 2019; Baker et al., 2021; Breinholt et al., 2021). The multispecies coalescent (MSC), which reconciles gene trees with species trees while accounting for ILS (Molloy and Warnow, 2023), is an appropriate framework for *Brachyotum* because we expect a high degree of gene tree discordance resulting from ILS during its rapid radiation. When ILS becomes substantial, trees inferred using concatenation of loci yield conflicting results when compared to a tree inferred under a coalescent framework (Kubatko and Degnan, 2007; Mendes and Hahn, 2018). Complementing the bifurcating species tree, phylogenetic networks infer introgression events in a phylogenomic framework by identifying branches that have been involved in historic gene flow (Hibbins and Hahn, 2022), while accounting for ILS (Solís-Lemus and Ané, 2016). Finally, the comparison of the species tree inferred with data from the diploid nuclear genome to the phylogeny of the uniparentally inherited plastome is widely used in phylogenomics to identify potential historic introgression that resulted in cytonuclear discordance (Baker et al., 2022). A plastome phylogeny can be obtained from off-target reads from target capture and from genome skimming data (Straub et al., 2012). Here we apply these complementary phylogenomic approaches with the aim of inferring a robust phylogeny for *Brachyotum* and its closest relatives while attempting to disentangle the complexity of the evolution of this charismatic member of the high Andean flora.

MATERIALS AND METHODS

Study system, sampling, and data generation

Our ingroup included species of a clade referred to as "*Brachyotum* + allies" (Meyer et al., 2021) and many representatives of the closely related *Chaetogastra* (Michelangeli et al., 2013; Guimarães et al., 2019; Meyer et al., 2021). The 34 sampled species of *Brachyotum* (~62% of total known species) reflect the diversity of floral and vegetative morphology (Figure 1) and the broad geographic range of the genus. The 14 sampled *Chaetogastra* species were included to test the monophyly of *Brachyotum* and mainly comprised Andean members whose phylogenetic placement is known to be closely to *Brachyotum* or within the "*Brachyotum* + allies" clade (Michelangeli et al., 2013; Guimarães et al., 2019; Meyer et al., 2021). Only one species across our sampling, *Chaetogastra scabriuscula* (Schltdl.) P.J.F. Guim. & Michelang., occurs outside the Andes, in the Trans-Mexican volcanic belt and the Sierra Madre del Sur. We also included three of the four members of the *Tibouchina* incertae sedis species as ingroup taxa (see Guimarães et al., 2019). Finally, we included the Andean species *Andesanthus lepidotus* (Bonpl.) P.J.F. Guim. & Michelang. and five Atlantic Forest

FIGURE 1 Members of the *Brachyotum* + allies. Color of frames around images correspond to the clades consistently inferred across different topologies with the nuclear data set (Figures 2–5): dark purple, *Brachyotum* II (A–H); fuchsia, *B. quinquenerve* (J); pale purple, *Brachyotum* I (K, L); red, *Chaetogastra* V (I); salmon, *Chaetogastra* III (M–O). (A) *B. gleasonii*, (B) *B. ledifolium*, (C) *B. harlingii*, (D) *B. confertum*, (E) *B. cutervoanum*, (F) *B. sanguinolentum*, (G) *B. rostratum*, (H) *B. coronatum*, (I) *Tibouchina calycina*, an *incertae sedis*, (J) *B. quinquenerve*, (K) *B. grisebachii*, (L) *B. nutans*, (M) and (N) *Chaetogastra* sp., (O) *C. pleromoides*. Photo credits: A–D, Carmen Ulloa Ulloa; E–H, J–K, Diego Paredes-Burneo; I, M–L, Fabián A. Michelangeli.

species of *Tibouchina* s.s. as outgroup taxa. Specimens and locality information are provided in Appendix S1.

We generated data for all ingroup taxa using tissue collected during recent fieldwork and from herbarium specimens from the Missouri Botanical Garden (MO) and New York Botanical Garden (NY). Fieldwork targeted localities in Ecuador and Peru, where richness of *Brachyotum* and *Chaetogastra* is highest, sometimes with co-occurring species from both genera. For the outgroup, the tissue of *A. lepidotus* was freshly collected in the field, and we retrieved phylogenomic data for species of *Tibouchina* s.s. from a published data set (Jantzen et al., 2022). Taxon sampling encompassed one sample per species, with 51 species in the ingroup and six in the outgroup.

We extracted DNA using a modified version of the CTAB protocol (Doyle and Doyle, 1987). Modifications included longer incubations for the cell-wall lysis (12 h in extraction buffer at 60°C) and the DNA precipitation steps (12 h in isopropanol at −18°C); we also did not use RNAase or proteinase K at any step.

Genome sampling was aimed at recovering both nuclear and plastid loci. Recovery of the nuclear loci used a target-capture sequencing method with a probe set tailored for Melastomataceae (Jantzen et al., 2020). This tailored probe set was designed to target 384 putatively single to low-copy loci. Plastome data was obtained through genome skimming data, where most loci had more than 2× coverage with a maximum of 10× (Appendix S2: Figure S1). We also added

a previously published genome-skimming accession of *B. microdon* to our data set (Appendix S1). We retrieved both nuclear and plastid data from the same samples through the inclusion of uncaptured library in the sequencing lane alongside samples that undergone target capture. Probe synthesis, library preparation, hybridization enrichment, and high-throughput sequencing were done by RapidGenomics (Gainesville, FL, USA).

Processing raw nuclear target-capture data

Cleaning raw sequences included removing adapters from paired-end reads and omitting low-quality reads with Trimmomatic v 0.39 (Bolger et al., 2014). Assembly, paralog identification, and recovery statistic determinations were done with HybPiper v 2.0 (Johnson et al., 2016) using default settings. For assembly, we used the melastome-specific target file (see previous section; Jantzen et al., 2020) with a filtering step: When a specific locus had multiple sources for target reference sequences, we kept only the sequence from a melastome (other sources were *Arabidopsis thaliana*, *Theobroma cacao* and/or the Angiosperms353 probe set). After assembly, we retrieved data for only coding regions rather than supercontigs (using the exonerate command in HybPiper) to reduce both the number of chimeric sequences (i.e., when sequences from different copies of a paralogous locus are misassembled across intronic regions) and the amount of missing data (i.e., there was a high proportion of missing data in intronic regions for multiple specimens).

Loci flagged by HybPiper as having multiple assembled copies underwent orthology assessment using the monophyletic outgroup (MO) approach (Yang and Smith, 2014; Morales-Briones et al., 2021). To this end, we used multicopy loci with nonchimeric paralogous sequences recovered by HybPiper. We generated alignments using MACSE v 1.02 (Ranwez et al., 2011) with default settings. We then performed a cleaning step to keep a minimum of 10% occupancy per site column using the pxclsq program from phyx (Brown et al., 2017). We inferred gene trees using IQ-TREE version 2.2.0.3 (Minh et al., 2020) under the GTR + G model, with 1000 ultrafast bootstrap replicates (UFBoot2; Hoang et al., 2017) coupled with a reduction of branch support overestimation during the pruning step (through the -bnni flag). Additionally, we manually removed loci with more than two copies in over 10% of the samples. The MO approach requires that the outgroup has only one copy per locus, a condition not met by the species we assigned as an outgroup, *A. lepidotus*. However, because all copies consistently formed a clade across all genes for this species, we selected only the copies listed as "main" by HybPiper (i.e., the longest copy). After keeping only one of the copies for *A. lepidotus*, we re-inferred gene trees with the same methodology described above. Finally, for the orthology inference step, we kept a minimum of five ingroup taxa (10% total).

We then identified loci with a single assembled copy to align them using MACSE version 1.02 with default settings (Ranwez et al., 2011). The single–copy data set, complemented with the inferred orthologous loci, constitutes the nuclear locus data set. We then removed samples with less than 30% of all loci recovered and loci that were recovered in fewer than 50% of samples. The scripts to identify single-copy loci and filter data based on recovery are available at https://github.com/ambed0ya/Palicourea. Finally, we trimmed the alignments to keep a maximum of 40% missing data per site column using the pxclsq program via phyx (Brown et al., 2017).

We further analyzed allele divergence and locus heterozygosity to assess the occurrence of potential hybrids in our sampling. We used HybPhaser 2.0 (Nauheimer et al., 2021) to process raw data, filter loci (that were recovered in less than 50% of samples), and analyze and visualize genetic variability in the data.

From the 57 samples included in this study, we removed four samples due to low recovery (i.e., with sequence data for <30% of the recovered loci; Appendix S2: Figures S2–S4). The remaining 53 samples had an average of 220 loci (median 224), representing 32.8–60.2% (126–231 loci) of all targeted loci. After inferring and visually assessing gene trees from these 54 samples, we also removed two samples due to possible contamination: They had significantly longer branches for most gene trees and sequences were highly dissimilar to those of any other samples in the alignments. Of the 384 targeted loci, we recovered sequence data for 270. We removed 45 loci due to low recovery across samples (<30%). Of the remaining 225 loci, 65 had single-copy sequences, and the other 160 loci were flagged as containing multiple assembled contigs (i.e., putative paralogs). The 160 loci with putative paralogs underwent subsequent orthology assessment with the MO pipeline (Appendix S2: Figure S5). After filtering for (1) < 2 contigs across 10% of samples, (2) not having identical sequences across samples in the same alignment, and (3) the outgroup being inferred to be monophyletic in each gene tree, we inferred 39 orthologous loci from loci flagged as putative paralogs, for a total of 104 nuclear loci in our data set. Finally, we removed the 54 shortest loci (<430 bp) to reduce error in gene tree estimation (Cai et al., 2021).

Phylogenetic inference and gene tree discordance of the nuclear data set

We inferred individual gene trees using IQ-TREE 2.2.0.3 with the GTR + G substitution model and 1000 UFBoot2 steps with a reduction in branch support overestimation. We also implemented the Shimodaira–Hasegawa-like approximate likelihood ratio test (Guindon et al., 2010), as another metric of branch support, using 1000 replicates, also with IQ-TREE 2.2.0.3. Branches with the lowest possible support (i.e. SH-aLRT = 0) were collapsed using Newick Utilities (Junier and Zdobnov, 2010) and the method of Simmons and Gatesy (2021).

We used both concatenation and multispecies coalescent (MSC) methods to infer the evolutionary history of the species. We concatenated sequences with Geneious Prime software v.2024.0 (Dotmatics; Boston, MA, USA). The concatenated tree was inferred using IQ-TREE 2.2.0.3 with GTR + G as the substitution model in a single partition, with 1000 steps for both the UFBoot2 and SH-aLRT and using a reduction in branch support overestimation (through the flag -bnni).

We also used quartet sampling (QS) scores (Pease et al., 2018) on the concatenated tree. The QS scores are a measure of the frequency of each of the three possible quartets per internal branch, i.e., a measure of branch support. The QS scores comprise four metrics: (1) quartet concordance, which measures how frequently the observed quartet was inferred; (2) quartet differential, which measures how equal the frequency of the alternative quartets are; (3) quartet informativeness, which measures the strength of the signal of a given quartet to distinguish between concordant versus discordant quartets; and (4) quartet fidelity, which measures the likelihood that a given taxon will produce a concordant quartet. The QS scores were calculated for each internal branch of the concatenated tree using the same alignment that generated that tree. Parameters were estimated over 10,000 replicates with a sampling threshold of 2 (ln likelihood threshold = 2; Pease et al., 2018) and a GTR + G model calling IQ-TREE for likelihood calculation.

For species tree inference using multispecies coalescent (MSC), we considered two different approaches: a co-estimation approach with StarBeast3 (Drummond et al., 2012) and a summary method approach with ASTRAL-IV (Zhang et al., 2018; Zhang and Mirarab, 2022). StarBeast3 simultaneously estimates both gene and species trees as a joint Bayesian inference using alignments, while ASTRAL-IV uses a set of gene trees as input (Molloy and Warnow, 2023).

For co-estimation of gene trees and species trees, we used a Bayesian analysis, where we specified the GTR + G substitution model, treated each gene matrix as a separate partition, specified a strict clock and a birth–death tree prior (net diversification prior and population mean prior following log-normal distributions), and an MCMC chain of 1 billion steps, with all other parameters kept as default. We ran two independent chains and assessed convergence of all parameters in Tracer 1.7.2 (Rambaut et al., 2018) (effective sample size [ESS] >200). We used the inferred results by combining two independent runs after a burn-in step of 10% for each. After combining the output, we estimated a maximum clade credibility tree using TreeAnnotator 2.7.6 (Drummond et al., 2012) with common ancestor heights for node heights.

We applied two summary method approaches with ASTRAL: the standard ASTRAL-IV (Zhang et al., 2018; Tabatabaee et al., 2023) and weighted-ASTRAL algorithms, the latter to incorporate branch support and branch length for species tree inference (Zhang and Mirarab, 2022). In both cases, we incorporated optimality for subsampling in the heuristic search (using the flag -R). We plotted trees using the R package GGTREE (Yu et al., 2017).

Plastome data assembly and phylogenetic analysis

Cleaning, de novo assembly, and alignment of raw data from genome skimming sequencing was performed with CAPTUS 1.0.1 (Ortiz et al., 2023) using default settings to output data per locus. We cleaned alignments to keep a maximum of 40% missing data per site column with the program pxclsq from phyx (Brown et al., 2017). We inferred a likelihood tree for the plastome data using IQ-TREE 2.2.0.3, with each individual locus considered as a different partition to perform substitution model selection (Chernomor et al., 2016). We ran 1000 steps for both the UFBoot and SH-aLRT and a reduction in branch support overestimation. The tree was plotted using the R package GGTREE (Yu et al., 2017).

Phylogenetic network inference using the nuclear data set

We used SNaQ in the Julia package PhyloNetworks to infer a phylogenetic network from quartets using a heuristic search approach based on pseudolikelihoods (Solís-Lemus and Ané, 2016; Solís-Lemus et al., 2017). The method infers networks where each branch may participate in a maximum of one introgression event (level-1 networks). Taxon sampling covered half of the species to keep the model computationally tractable, resulting in 24 ingroup taxa and one outgroup. This resampled taxon list included one sample per clade identified in Figures 2–5, except for *Brachyotum*, for which sampled more taxa to test whether cytonuclear discordance (Figure 2) is consistent with evidence of introgression; across all targeted groups in our taxon subsampling, we kept the samples with the most recovered loci. We used all the gene trees from nuclear loci previously inferred with StarBeast3 after combining the two independent runs (see the section Phylogenetic inference with the nuclear data set). To keep the selected species, we trimmed gene trees using the program pxtrt from phyx (Brown et al., 2017) to use them as input for SNaQ. We first calculated concordance quartet factors with PhyloNetworks. We then ran the algorithm specifying the StarBeast3 tree as the starting topology and allowing one additional hybridization event each time, starting from zero to a maximum of four. We plotted networks using PhyloNetworks, retrieving the major tree topology (i.e., the backbone of the network where introgression events occur), which includes the direction of the introgression events identified with their respective inheritance probability (γ).

Quantifying and visualizing topological conflict

We quantified differences in tree topology using the generalized Robinson–Foulds (RF) distance (Smith, 2020), which incorporates the traditional count of equal bipartition

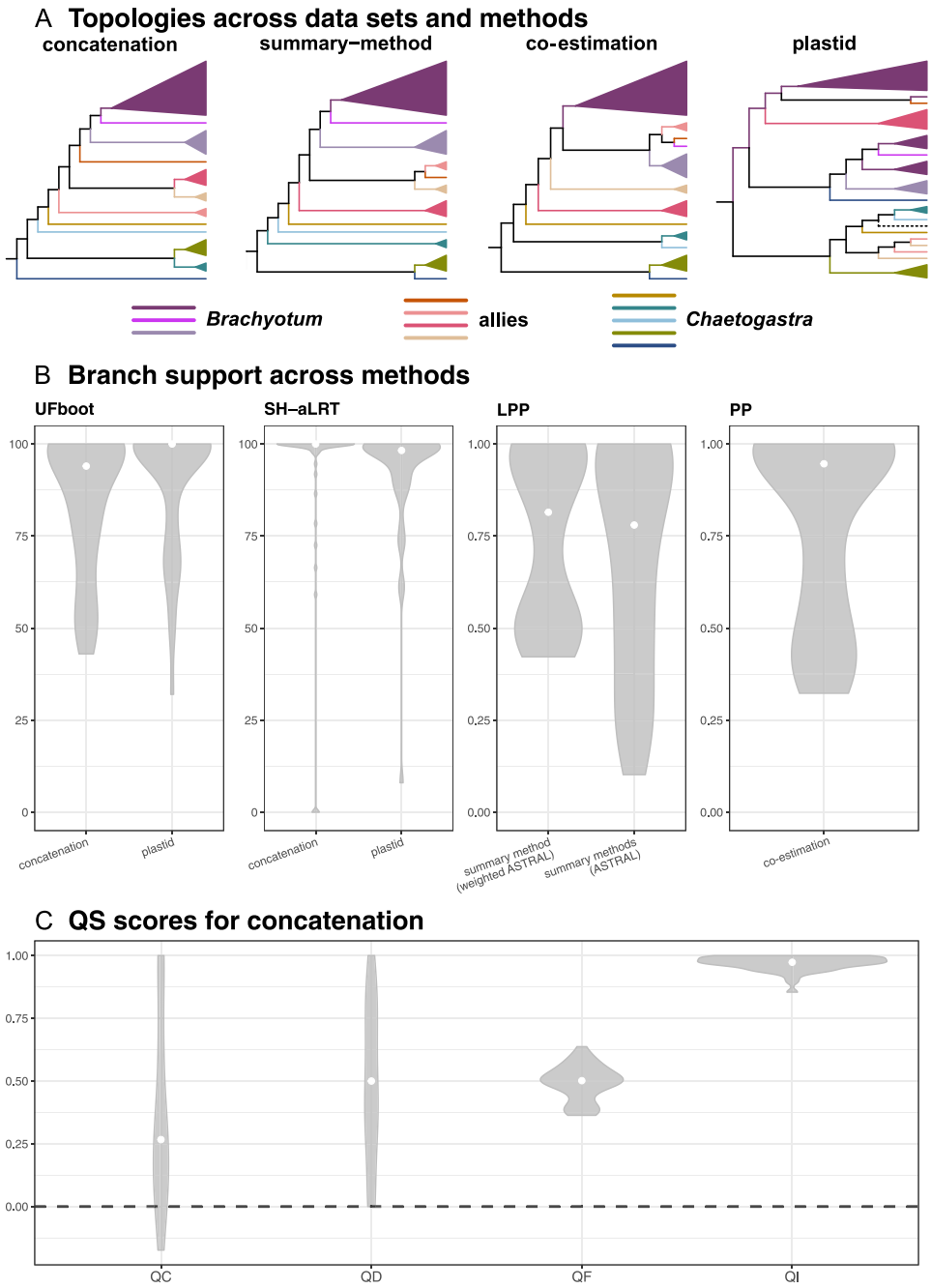

FIGURE 2 Phylogenetic conflict for evolutionary relationships of *Brachyotum* + allies and *Chaetogastra* across data sets and methods. (A) Different topologies resulting from different methods (concatenation, summary-methods, co-estimation) and datasets (nuclear and plastid loci). Colored branches indicate the different groups. (B) Branch support across methods in violin plots. UFboot (ultrafast bootstrap), and SH-aLRT (Shimodaira–Hasegawa approximate likelihood ratio test) are values for the maximum likelihood approach (IQ-TREE), local posterior probability (LPP) for the summary methods (ASTRAL-IV and weighted ASTRAL), and posterior probability (PP) for co-estimation (Beast, Bayesian inference). (C) Quartet sampling (QS) scores in violin plots for the topology inferred under concatenation (QC: quartet concordant, QD: quartet differential, QF: quartet fidelity, QI: quartet informativeness). The dashed line represents the threshold between support and counter-support for QC. White dots in panels B and C represent the median. Only the ingroup was included across these figures.

splits per pair of trees and scores almost-identical splits. Given the different topologies from the nuclear and plastid data sets across gene trees and between different species tree inference methods, we aimed to explore variation across phylogenetic methods and data type. Thus, we calculated pairwise generalized RF distances (1) among data sets, (2) across different phylogenetic inference methods, and (3) between gene trees and species trees. Pairwise distances were calculated with the R package TreeDist v 2.9.2 (Smith, 2020).

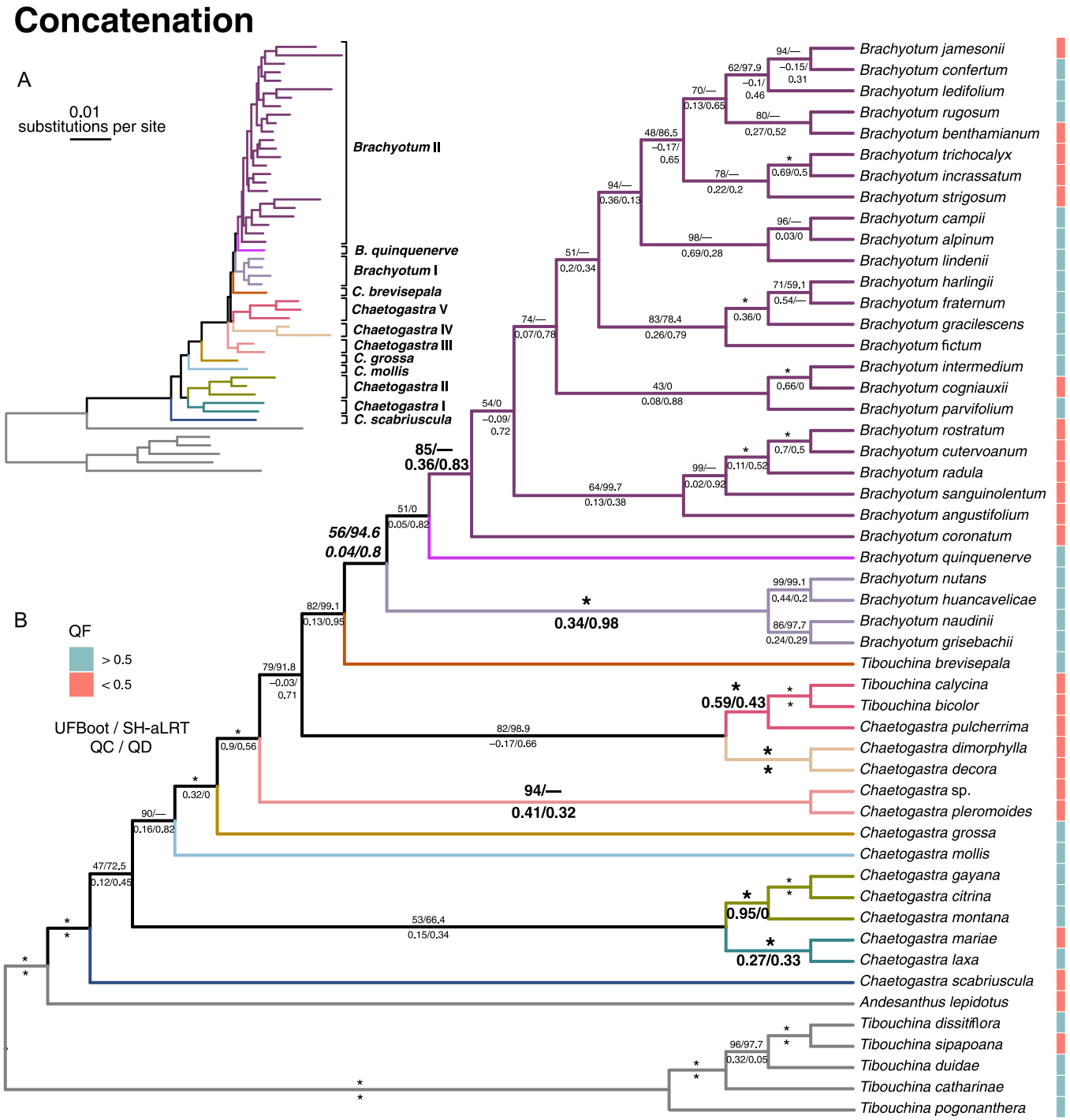

FIGURE 3 Evolutionary relationships of the *Brachyotum* + allies and *Chaetogastra* inferred under concatenation. (A) Consensus tree inferred under a concatenation approach (maximum likelihood with IQ-TREE) using 50 nuclear loci. Colored branches represent clades consistently inferred across analyses of the nuclear data set (as in Figure 2). The outgroup is shown in gray. (B) Concatenated tree from panel A as a cladogram with four different measures of branch support: above, UFBoot on the left, and SH-aLRT on the right; below, the quartet concordance score on the left (QC) and the quartet differential score on the right (QD). Values in bold correspond to consistent support across measures. The bold, italicized value is the branch support for the monophyly of *Brachyotum*. Full support for a given pair of measures (*) and full support for a single measure (—) are shown. The quartet fidelity (QF) is shown in solid squares next to the species names on the right.

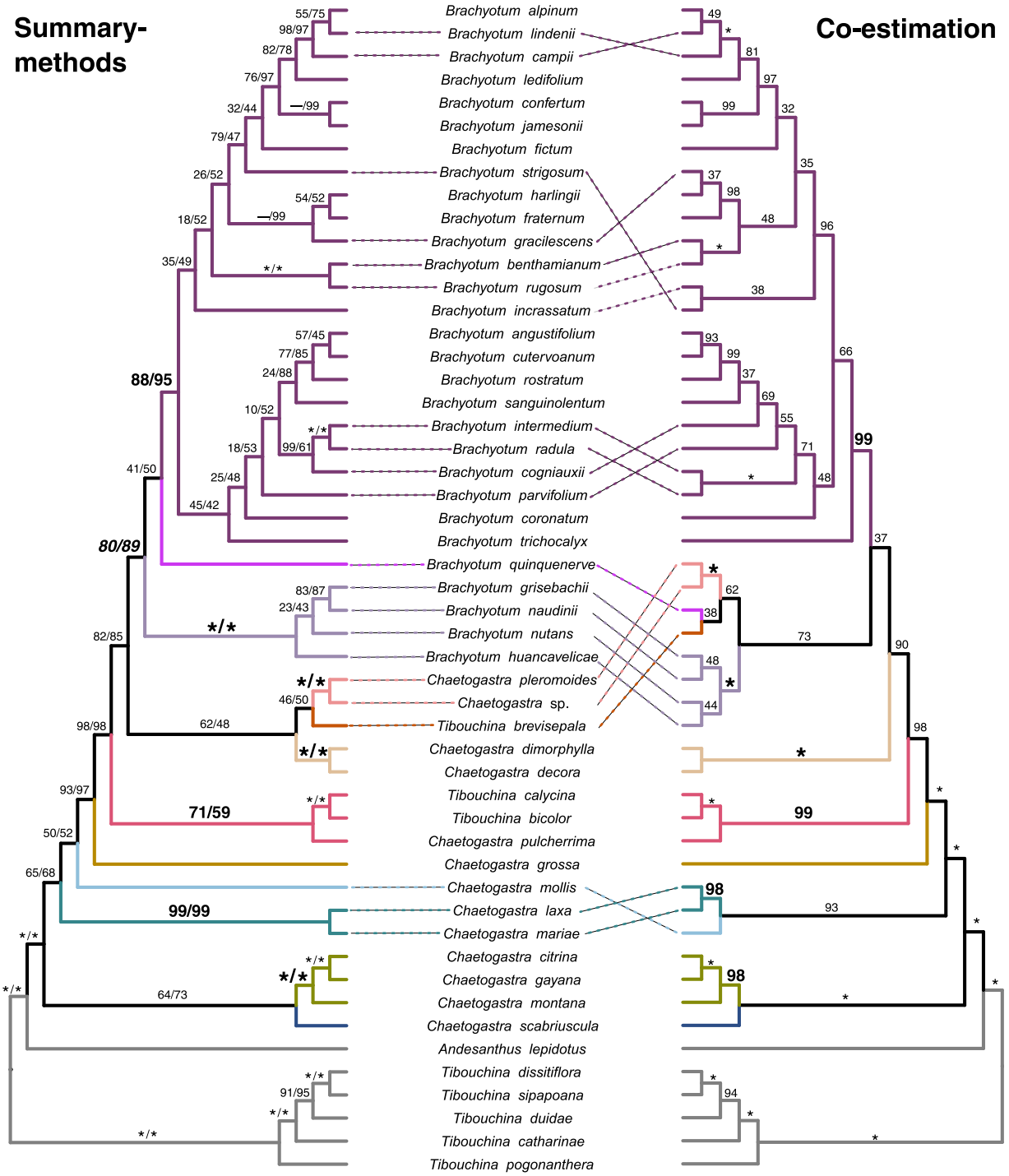

FIGURE 4 Topological conflicts between the two multispecies coalescent trees of *Brachyotum* and *Chaetogastra* using 50 nuclear loci. Left: consensus tree inferred under a summary-method approach (with ASTRAL-IV and wASTRAL inferring the same topology). Local posterior probability estimated with ASTRAL-IV and wASTRAL are shown above branches (right and left respectively). Right: Maximum clade credibility inferred under a co-estimation approach (with StarBeast3). Posterior probability shown above branches. Clades identified within each genus are shown with colors as in Figure 2. Black branches depict the internal branches between such clades and dashed lines connect the topological conflicts between both trees. Outgroup shown in gray.

RESULTS

Features and quality of the nuclear data set

We generated a data set of 51 samples for downstream analyses, which comprised 29 *Brachyotum*, 13 *Chaetogastra*, three incertae sedis species in *Tibouchina*, and six outgroup species (*A. lepidotus* and five species from *Tibouchina* s.s.). This final data set comprised 50 loci (hereafter, the nuclear data set) and was used for downstream analyses. Individual alignments ranged from 432 to 3312 bp with a median of 916.5 bp. The concatenated matrix of the 50 nuclear loci had

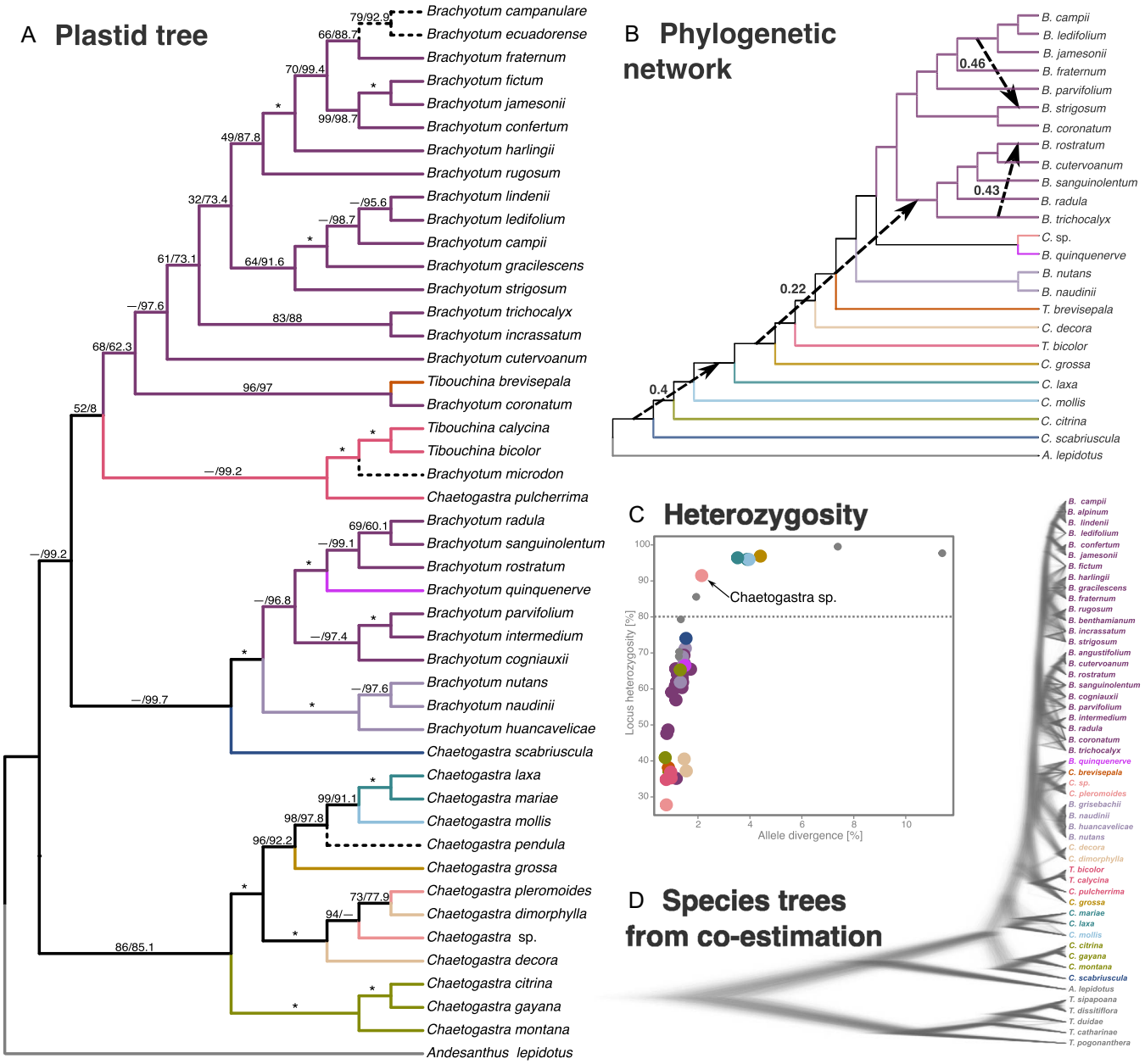

FIGURE 5 Evidence of introgression in *Brachyotum* and *Chaetogastra*. Colored branches highlight the clades consistently inferred across methods; the outgroup is depicted in gray. (A) Cladogram of the consensus tree of *Brachyotum* and *Chaetogastra* using 80 plastid loci. Branch support is shown (UFBoot on the left and SH-aLRT on the right). Full support for both metrics (*) and full support for one of the metrics (—) are shown. Dashed branches represent tips not included in the analyses with nuclear loci. (B) Phylogenetic network where the dashed branches represent the directionality of the introgression events and their associated inheritance probabilities (γ) for $H_{max} = 4$. (C) Allele divergence and locus heterozygosity plot of the nuclear data set. Each dot represents a sample; color indicates its clade. Dashed line indicates the threshold in locus heterozygosity for putative hybrids. Outgroup represented with a smaller dot size. (D) Densitree plot from the co-estimation inference using 50 nuclear loci. The individual trees are species trees (1000 total) inferred during the MCMC run after a burn-in of 10%.

a length of 54,024 bp with 75.54% (40,811 bp) invariable sites and 12.31% (6651 bp) parsimony informative sites.

Assessment of potential hybrids showed that five *Chaetogastra* and three members of the outgroup had more than 2% of allele divergence and more than 80% of locus heterozygosity, including *Chaetogastra* sp. (voucher ID Michelangeli 1865) as the only member of the "*Brachyotum* + allies" clade (Figure 5C).

Phylogenetic relationships and discordance in the nuclear data set

Brachyotum is monophyletic with moderate support in two of the three inference strategies, and paraphyletic only in the MSC co-estimation analysis (Figures 2–4) with low support. *Chaetogastra* was recovered as paraphyletic across all inferred trees (Figures 2–5). Twelve clades consistently

recovered across topologies, with high to full support, with *Brachyotum* having three clades (Figure 2; Appendix S3: Table S1).

Branch support in the ingroup ranged from low to maximum support (Figure 2B, C). Only 9.09% (four of 44) of internal branches in the concatenated tree had full support across all metrics, and 22.73% (10 of 44) had maximum values for two metrics (UFBoot2 and SH-aLRT). The QS scores varied greatly, except QI where the phylogenetic signal was high for all internal branches (quartet informativeness, QI > 0.95 in ~80% of internal branches, Figure 2C; Appendix S3: Table S2). More than 50% of branches showed weak support (quartet concordance, QC < 0.25), and 13.64% of branches (six of those 12) showed counter-support (QC < 0), in which discordance is more common than concordance. Most branches showed skewness toward one of the discordant quartets (quartet differential, QD < 0.5), where 11.36% of branches (5 of 44) had values of zero, a pattern consistent with introgression over ILS. Finally, many taxa in the ingroup mostly produced discordant quartets when sampled (quartet fidelity, QF < 0.5) (Figures 2, 3), pointing to potential "rogue" behavior of many included tips (Pease et al., 2018). More details of the QS scores are in Appendix S3 (Table S2).

The three clades of *Brachyotum* were consistently recovered even when the genus was not monophyletic (branch test, LPP or PP ≥ 95) (Figures 2–4). A four-species clade recognized as *Brachyotum* I, another clade represented by a single species, *B. quinquenerve*, and a clade with most species recognized as *Brachyotum* II (Figures 2, 3). The relationships between these clades were consistent under the concatenated ML and the MSC summary-method strategies (Appendix S2: Figures S6, S7), but differed for the topology given by MSC co-estimation, where *Brachyotum* was paraphyletic and included several *Chaetogastra* species, although with low support (Figures 2–4).

Relationships within each *Brachyotum* clade with multiple species differed among the inference methods, but the differing relationships usually had low support (Figures 3, 4). Relationships within *Brachyotum* I were consistent in the concatenation and the co-estimation analyses. Relationships within *Brachyotum* II were mostly consistent between the species trees inferred with both MSC approaches and greatly differed from the tree inferred in the concatenated ML analysis. Overall, branch support was low across internal branches within *Brachyotum*, sometimes with counter-support in the concatenated tree (i.e., the observed quartet was not inferred more frequently than the discordant quartets; quartet concordance < 0).

Relationships among the *Chaetogastra* and the incertae sedis species substantially varied across methods. They never formed a monophyletic group, but instead clustered in four separate clades. Similarly, the *Tibouchina* incertae sedis taxa never formed a single clade among them. Branch support for relationships including these "allies" tended to be low across methods (Figures 2–4, 5A).

Plastid phylogenetics

We recovered data for 46 samples (Appendix S2: Figure S4), and 80 loci were consistently recovered. The 80 loci represented 68,080 bp with 93.75% (63,828) invariable sites and 1.92% (1310) parsimony informative sites. We recovered plastid data for four samples not recovered for the nuclear data set due to low data: *B. campanulare*, *B. ecuadorense*, *B. microdon*, and *C. pendula*). Our final data set comprised 34 samples for *Brachyotum*, 14 for *Chaetogastra*, three for *Tibouchina* incertae sedis, and the assigned outgroup (*A. lepidotus*).

Phylogenetic relationships substantially differed between the plastid tree and the three nuclear topologies (Figure 5A). In the former, *Brachyotum* was inferred as non-monophyletic, including three *Chaetogastra* clades that contain all the *Tibouchina* incertae sedis species (Figure 5A). The largest *Brachyotum* clade in the nuclear trees (*Brachyotum* II; Figures 2, 5) was not monophyletic in the plastid tree. One relationship between genera is unique for the plastid, which placed a non-Andean species of *Chaetogastra* (*C. scabriuscula*) within *Brachyotum*. Also, the "allies" species were divided in two different clades: two (*Chaetogastra* V and *C. brevisepala*) grouped with *Brachyotum*, and two (*Chaetogastra* III and IV) grouped with the rest of *Chaetogastra* (Figures 2 and 5). Two of the three *Chaetogastra* clades nested in *Brachyotum* have high support for their placement within *Brachyotum*. Only one, *Chaetogastra* V, was poorly supported in its placement, with the lowest support values across the plastid tree.

Phylogenetic network inference from the nuclear data set

The inferred network with four maximum events of introgression ($H_{max} = 4$) was the one with the highest pseudo-likelihood score (Appendix S2: Figure S8). In the backbone tree (or major tree) where introgression events occur (Figure 5B), *Brachyotum* was resolved as paraphyletic due to the placement of a single *Chaetogastra* species (*Chaetogastra* sp.) as sister to *B. quinquenerve*. Two of our inferred introgression events occurred within *Brachyotum* II, on two different sides of the clade. These events had similar inheritance probabilities (γ): 0.43 between two terminal branches and 0.46 from an internal to a terminal branch. For the latter, *B. trichocalyx*, a species with unstable placement in our bifurcating phylogenies (Figures 3–5), was inferred to have experienced an introgression event with another species within *Brachyotum* II.

Two additional introgression events were identified between backbone branches. One of these involved gene flow from the branch bearing the *Brachyotum* + allies + *C. grossa* to the internal branch of one of the *Brachyotum* II subclades with $\gamma = 0.22$. Another event involved gene flow from the branch subtending the ingroup to another part of the backbone with $\gamma = 0.4$. Visualization of the posterior

distribution of all the MSC trees from our co-estimation analysis, from which gene trees for the network inference were taken, showed areas of uncertainty that may match with introgression events, for example between members of *Brachyotum* II (Figure 4C).

Quantifying topological conflict across methods and datasets

All our species-level phylogenies differed from each other in topology (Figures 3–5). The plastid tree was consistently the most different from the nuclear topologies (i.e., had lower generalized RF distances; Figure 6A) no matter which inference method was used for the nuclear data (concatenated ML, MSC summary-method, MSC co-estimation). In addition, the MSC trees were more similar to each other than to the concatenated tree.

For the nuclear gene trees, tree pairs inferred from the same gene but with two different methods (IQ-TREE and StarBeast3) tended to differ substantially (RF distance > 0.5; Figure 6B). Comparison of gene trees from each method against the species-level phylogenies had consistently lower generalized RF distances for StarBeast3 gene trees (Figure 6C). Again, the plastid tree differed most from the nuclear gene trees, independent of the method to generate them.

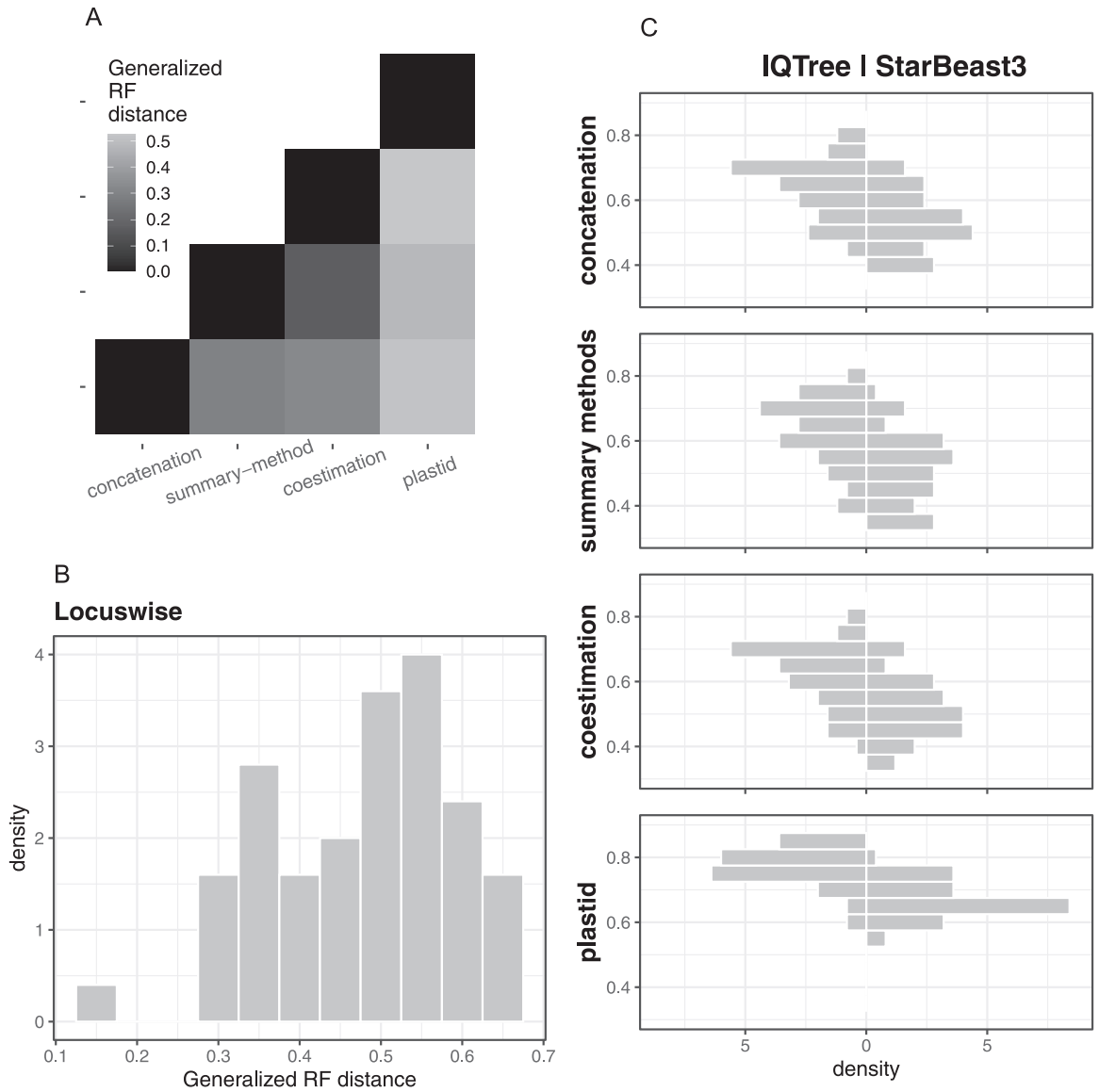

FIGURE 6 Topological conflict across inference methods and across nuclear and plastid datasets. (A) Pairwise generalized Robinson–Foulds distance between the species-level trees inferred, including the nuclear (concatenation, summary-method and co-estimation) and the plastid data sets. (B) Histogram of the pairwise generalized Robinson–Foulds distance between nuclear gene trees inferred from different methods (IQ-TREE vs. StarBeast3). (C) Histogram of the pairwise generalized Robinson–Foulds distance between each nuclear gene tree (inferred with IQ-TREE or StarBeast3) and the concatenated, ASTRAL, StarBeast3, and plastid tree.

DISCUSSION

The different evolutionary histories inferred for "*Brachyotum* + allies" variably supported and challenged the monophyly of *Brachyotum*. There was conflict across our different analyses: No two phylogenetic inference approaches produced the same topology, and there were significant differences across analyses and data sets. Differences were especially pronounced within *Brachyotum* and among the allied lineages that fell within *Brachyotum*. Despite morphological evidence suggesting that *Brachyotum* is a clade, its monophyly remains poorly supported. We found evidence consistent with ancestral introgression having occurred during the evolution of *Brachyotum*, which may have contributed to uncertainty about its monophyly. Evidence for extensive introgression is provided from (1) the QS scores on the concatenated tree (Figure 3), (2) the plastid topology (Figure 5A), which was substantially different from the nuclear topologies (Figures 2A and 6A), and (3) the phylogenetic network analysis, which favored the maximum number of introgression events modeled. Despite efforts to incorporate both incomplete lineage sorting and introgression in phylogenetic inference, we failed to find strong support for the monophyly of *Brachyotum*. This result may be due to the particularly complex evolutionary history of this high Andean taxon during its rapid diversification, which may have experienced a combination of historic introgression, ongoing gene flow (as evidenced by a putatively hybrid species, *Chaetogastra* sp.), and high levels of ILS. Despite topological uncertainty, *Brachyotum* is characterized by morphological traits that unite its species (notably including pseudotubular, nectar-producing flowers, which are unique within the family), providing concordant evidence in support of its monophyly as found in many of our analyses, and supporting its continued recognition as a taxon. Here, we discuss the scope and limitations of inferring an accurate phylogenetic history in such a complex group using current methods.

Topological conflict predominates in the phylogenomics of *Brachyotum*

Our analyses of *Brachyotum* yielded conflicting phylogenetic results. Each of the methods applied to the nuclear data set generated different evolutionary histories, and the nuclear and plastid phylogenies differed from each other (Figures 2–5). Phylogenetic conflict was clear despite our efforts to minimize systematic error and incorporate as much evolutionary complexity as possible in the approaches we used. Considering the conflict and low values of branch support across methods and data sets (Figure 2B, C), we suspect that the evolutionary history of "*Brachyotum* + allies" is more complex than the scope of currently available models for phylogenomic inference. As discussed below, our results for this group are consistent with a rapid radiation, ILS, and hybridization in *Brachyotum*'s evolutionary history.

Conflict between analyses mainly occurred in areas with the shortest branches. The relationships of lineages along the backbone of "*Brachyotum* + allies" were consistently different across analyses. This area of the phylogeny also had the shortest branch lengths (as measured by substitutions per site, Figure 3; Appendix S2: Figure S4). This pattern may result from ILS, which is elevated when the time between speciation events is short (Pamilo and Nei, 1988). However, conflict still remained across analyses that incorporate a coalescent framework (Figure 4), supporting the idea that potential additional factors have shaped the evolution of *Brachyotum*. These factors might include introgression and further changes in genomic regions caused by recombination and substitution rate heterogeneity (Steenwyk et al., 2023); the latter may be evidenced by unbalanced branch lengths among sister species (Figure 3A). Model assumptions, specifics of the search phylogenetic search algorithms employed, and data quality may also contribute to conflict among gene trees.

Concatenation-based and MSC analyses of nuclear data generated phylogenetic hypotheses that differed significantly. The two MSC methods produced a particularly notable difference when compared to each other: *Brachyotum* is monophyletic under the summary-methods approach (ASTRAL), but not under co-estimation (BEAST) (Figure 4). The tree inferred from the concatenated analysis, in which *Brachytoum* is monophyletic, markedly differed from each of the MSC trees (Figures 2A, 6A), which is expected given that the former does not incorporate relevant biological processes such as conflicting gene histories under ILS (Kubatko and Degnan, 2007). Especially when ILS levels are high, concatenation approaches under maximum likelihood fails (Mendes and Hahn, 2018). Given the differences in topology and the additional evidence of introgression, including intergeneric hybridization, we argue that the concatenation approach is not fully reliable for "*Brachyotum* + allies". Further, some empirical evidence also supports the application of MSC methods over phylogenetic inference using concatenated data sets due to the increase in bias (Jiang et al., 2020). Nonetheless, some phylogenetic relationships are consistent across the three methods, including the monophyly of *Brachyotum* I and the paraphyly of *Chaetogastra* (Figures 1–3).

The plastid topology had the greatest conflict with the other inferred trees (Figures 2, 5, 6). Most branches in the plastid tree had high to full support (Figures 1–5), although the placement of most *Tibouchina* incertae sedis species within *Brachyotum* II had weak support. These relationships among genera were not recovered in any nuclear topologies, where the *Tibouchina* incertae sedis species always fell outside of the clade that includes *Brachyotum*, in contrast to previous studies that concatenated both plastid and nuclear loci and where plastid information contributed more data (Michelangeli et al., 2013; Guimarães et al., 2019) or more variable sites (Meyer et al., 2021). Therefore, we suspect that concatenation of plastid and nuclear loci promotes a bias toward placing the *Tibouchina* incertae sedis within *Brachyotum*, due to cytonuclear discordance, i.e., the

combination of genes and organelles with conflicting evolutionary histories.

An additional area of phylogenetic complexity is a high degree of nuclear gene tree discordance (Figure 6B, C). The nuclear gene trees here differed greatly from inferred species trees, independently of the method used (Figure 6C). In addition, uncertainty (i.e., low bootstrap and branch test scores) was also substantial across gene trees (Figure 2B), despite evidence of high phylogenetic signal in the data set (high QI; Figure 2C). Gene tree discordance is affected by a combination of biological factors. For example, recombination after a hybridization event may lead to increased gene tree discordance (Posada and Crandall, 2002; Steenwyk et al., 2023), an outcome that is relevant considering our evidence supporting past introgression. Here we found signals of multiple introgression events (Figures 2, 5). Additionally, branch length variation in our concatenated tree (Figure 3) likely indicates substantial rate heterogeneity. This variation in evolutionary rates, either as lineage-, gene-, or site-specific or a combination, may bias introgression inference (Frankel and Ané, 2023), while also decreasing phylogenetic accuracy in general (Vankan et al., 2022). Heterotachy, or lineage-specific variation in rates through time (Lopez et al., 2002), for instance, has occurred during the evolution of the plastid in algae (Whelan et al., 2011), and its incorporation may improve model fit (Toups et al., 2025).

The combination of species-level phylogenetic conflict, gene tree discordance, and uncertainty helps explain why *Brachyotum* remains a recalcitrant group to resolve. As in *Brachyotum*, conflict occurs between nuclear and organellar genomes (plastid or mitochondria; Vargas et al., 2017; Pouchon et al., 2018) or between different methods of phylogenetic inference (Lagomarsino et al., 2022) in different Andean clades across angiosperms. Topological conflict across trees appears to be a general pattern for the phylogenomics of Andean plants.

Multiple lines of evidence support historical introgression in *Brachyotum*

Uncertainty about the monophyly of *Brachyotum* may be explained by introgression and is potentially supported by the evidence we presented here across topologies and data sets. The nuclear, network, and plastid analyses, either in isolation or comparison, inferred different and complementary events of introgression. These lines of evidence are consistent with gene flow having occurred across multiple areas of the ingroup, not only within the clade comprising *Brachyotum* + allies. Given the consistent inference of gene flow across data sets and methods, introgression likely at least partially explains the topological conflict observed here between the nuclear and plastid topologies.

Introgression within the clade comprising "*Brachyotum* + allies" likely involved multiple events. The phylogenetic network, which provides the most explicit piece of evidence for introgression in our study, inferred that gene flow occurred from the branch subtending the "*Brachyotum* + allies" clade to an internal branch within *Brachyotum* (Figure 5B). Approximately 22% of the genome sampled (given the estimated inheritance probability) was transferred from the broader "*Brachyotum* + allies" clade to *Brachyotum*. Three scenarios could explain this pattern. One is that of historic introgression between ancestors along each internal branch, affecting the ancestral populations on those internal branches or lineages. A second scenario is that of ghost introgression, due to unsampled or now-extinct species involved in the introgression event (Pang and Zhang, 2023). A third alternative involves a combination of the previous two scenarios. Even though we cannot clearly pinpoint which specific species were involved in this introgression event, the phylogenetic network provides strong evidence of introgression within the "*Brachyotum* + allies" clade.

The incertae sedis species of *Tibouchina*, whose phylogenetic positions were uncertain before the present study (and so retain the genus name *Tibouchina*), are all inferred to be involved in introgression events with *Brachyotum*, as evidenced by the conflict between the nuclear trees and the plastid topology (Figures 2A, 3–5A). In the plastid topology, *T. brevisepala* was inferred to be the sister group of *B. coronatum*, while *T calycina* and *T. bicolor* were sister to *B. microdon*; in contrast, in the nuclear phylogenies, these incertae sedis species mostly fell outside of the clade that includes *Brachyotum* species (the co-estimation MSC topology being the only exception). This cytonuclear discordance may be interpreted as direct evidence of introgression (Linder and Rieseberg, 2004; Lee-Yaw et al., 2019; Wang et al., 2022): After any given interspecific hybridization event with subsequent backcrosses, a mismatch may occur between the nuclear and plastid genomes due to the uniparental inheritance of the plastome (Rieseberg and Soltis, 1991; Linder and Rieseberg, 2004; Lee-Yaw et al., 2019; Wang et al., 2022). An additional indication of potential introgression involving the incertae sedis species comes from estimates of gene tree discordance from the concatenated nuclear tree. At least two internal branches subtending the incertae sedis species within the clade comprising "*Brachyotum* + allies" had a combination of negative values for the quartet concordance and low values for the quartet differential. This combination of QS scores suggests that an alternative quartet is favored (Figure 3) (Pease et al., 2018). In addition, *T. bicolor* had one of the lowest QF values and is thus highly likely to result in gene tree discordance when included in each analysis. Multiple factors can produce low QS scores, including ILS, introgression, and rate heterogeneity (Pease et al., 2018), although with the evidence of cytonuclear discordance, we argue that introgression is almost certainly a contributor. Thus, multiple historical introgression events involving incertae sedis species may explain their inferred phylogenetic position within *Brachyotum* in previous studies (Michelangeli et al., 2013; Guimarães et al., 2019; Meyer et al., 2021) and our inability to place these species into a genus given

phylogenomic evidence (Michelangeli et al., 2013; Guimarães et al., 2019; Meyer et al., 2021), including the analyses here. The phylogenetic uncertainty of the incertae sedis is consistent with previous taxonomic research that identifies their morphological resemblance to "*Brachyotum* + allies", including stamens without appendages and the presence of appressed hairs on vegetative structures, both found in *Brachyotum* and Andean *Chaetogastra* (Wurdack, 1965; Wurdack, 1988).

There are multiple additional cases of introgression supported by our phylogenetic network and cytonuclear discordance. A particularly unexpected case of introgression encompasses the only non-Andean species in our sampling, *C. scabriuscula*. This species is consistently placed—either alone or within a clade—as sister to the rest of the ingroup in the nuclear analyses (Figures 3–5). In contrast, this taxon is placed within *Brachyotum* in the plastid topology (Figure 5A). Further evidence of gene flow in the phylogenetic network involved the branch subtending *C. scabriuscula* and another internal branch close to *Brachyotum* + allies (Figure 5B). This report is the first of introgression between disjunct taxa involving *Brachyotum* and a Mexican *Chaetogastra* species, which was also placed in a clade sister to *Brachyotum* in previous analyses that concatenated nuclear and plastid loci (Guimarães et al., 2019; Meyer et al., 2021). This result may reflect historic introgression and possible long-distance dispersal between the Andes and Mesoamerican mountains.

Morphological evidence from *Chaetogastra* sp. (voucher ID Michelangeli 1865; Appendix S1) is consistent with ongoing hybridization. Indeed, this specimen has many traits consistent with its placement within *Chaetogastra*, although some characters seem to align better with *Brachyotum*. For example, this putative hybrid has well-developed connective appendages and a style that is bent to the side at anthesis, which both are traits that characterize *Chaetogastra* and that are very rare or absent in *Brachyotum* (Figure 1). On the other hand, its pendent flowers and isomorphic anthers with shortly elongated connectives are shared with all species of *Brachyotum* but are relatively rare to absent in *Chaetogastra*. Finally, the corolla morphology of this *Chaetogastra* sp. is intermediate between the pseudotubular corollas of *Brachyotum* and the open corollas of *Chaetogastra* (Figure 1M, N). There is also genetic evidence from allele divergence and locus heterozygosity that supports *Chaetogastra* sp. as a potential hybrid (Figure 5C) (Nauheimer et al., 2021), and this taxon exhibits cytonuclear discordance: It is grouped within *Brachyotum* with the nuclear data (in the phylogenetic network and the MSC co-estimation topology) and with *Chaetogastra* in the plastid topology (Figures 4, 5). Its co-occurrence with *C. pleromoides* and *B. quinquenerve* in southern Peru suggests these as possible parental species. High heterozygosity, morphological intermediacy, and cytonuclear discordance all point toward *Chaetogastra* sp. being a product of recent hybridization, and possibly even an F1 hybrid.

The history of introgression in "*Brachyotum* + allies" is likely even more complex than presented here. When inferring the phylogenetic network with the purportedly highly accurate SNaQ (i.e., high identifiability) (Solís-Lemus and Ané, 2016), the model fit, as assessed by the heuristic pseudolikelihood approach, increased as more introgression events were added. We stopped at four events due to excessive computational time and to avoid overfitting the model given our small taxon and locus set. However, we suspect that the occurrence of more events is realistic, due to the rampant presence of cytonuclear discordance in various areas of the ingroup, combined with weak support for multiple branches across nuclear-locus topologies. Given these multiple pieces of evidence, we suspect that the difficulty in determining whether *Brachyotum* is monophyletic results from the complexity introduced by introgression, which minimally involved the *Tibouchina incertae sedis* and *Chaetogastra* sp. but could be more widespread. This complexity is likely exacerbated by other sources of uncertainty, including ILS, rate heterogeneity, selection, and model misspecification.

Assumptions and inadequacy as sources of conflict

The conflict observed in our analyses may also be partially attributed to violation of assumptions by the models we used. Evidence for the importance of model choice on phylogenetic inference comes from the fact that each locus (i.e., individual alignments) in our data set produced gene trees with different topologies that were dependent on the method used to infer them (IQ-TREE vs. StarBeast3, Figure 6B–C). While ASTRAL uses already inferred gene trees from IQ-TREE here, StarBeast3 simultaneously infers the gene trees based on sequence alignments for each locus and the species tree using the multiple species coalescent. Neither method explicitly incorporates introgression.

Introgression can bias species-tree inference unless accounted for (Linder and Rieseberg, 2004; Yu et al., 2012). The topological conflict in the broader clade comprising "*Brachyotum* + allies" that we identified between the two multispecies coalescent approaches may reflect different model performance in the presence of introgression. Accuracy decreases for both methods with increasing levels of interspecific gene flow, although they differ in exactly how: in particular, ASTRAL performs better than StarBeast3 in the presence of introgression between nonsister taxa, while StarBeast3 performs better under scenarios of ghost introgression (Pang and Zhang, 2023). In general, unsampled diversity can impact the detection of introgression due to model violation (Tricou et al., 2022). Even though our sampling is the most comprehensive of "*Brachyotum* + allies" attempted to date, it covers only 60% of the extant species-level diversity of *Brachyotum*, and so ghost introgression is potentially relevant here.

Our MSC co-estimation approach (i.e., the fully Bayesian StarBeast3 analysis), is particularly parameter rich. Parameters inferred from our data included gene trees and the species trees, substitution rates, effective population sizes for each branch of the species tree, and species-tree priors including the clock model (Drummond et al., 2012; Bouckaert et al., 2014; Ogilvie et al., 2017). This large parameter space is biologically realistic but creates opportunities for model misspecification and likely overparameterization. In addition, we made pragmatic decisions to reduce parameterization of the StarBeast3 model to speed up computation time, in particular by assuming a strict clock. While each of our parameters performed well within the applied model (i.e., they reached convergence), this assumption may have impacted the accuracy of our phylogenetic analysis. It is notable that StarBeast3 is the only nuclear analysis in which *Brachyotum* is not inferred to as monophyletic, and its recovered topology implies a less parsimonious scenario of morphological trait evolution within the "*Brachyotum* + allies" clade (see last section in Discussion).

Model misfit or poor model performance in estimating the phylogeny of *Brachyotum* cannot be ruled out. In fact, there is likely at least some strong model misfit in our species tree analyses due to the strong evidence we found here in support of historic introgression. We took steps to reduce systematic error in our data sets by decreasing the amount of missing data (Xi et al., 2016; Nute et al., 2018), excluding short alignments (Cai et al., 2021), and collapsing poorly supported branches (Simmons and Gatesy, 2021). Regardless, uncertainty remains.

Phylogenetic complexity in high Andean radiations

Speciation events in "*Brachyotum* + allies" likely occurred in short evolutionary times given the short branch lengths (as substitution rates in the concatenated tree (Figure 3) and coalescent times inferred in the MSC tree; Appendix S2: Figure S7). Short evolutionary times are characteristic of a rapid radiation (Whitfield and Lockhart, 2007). Several radiations in the high elevation flora of the tropical Andes exhibit phylogenetic uncertainty and high rates of gene tree discordance. In some groups, such as *Espeletia* (Pouchon et al., 2018) and *Brunellia* (Murillo-A. et al., 2022), the incorporation of incomplete lineage sorting in the inference method significantly reduced the uncertainty in phylogenetic inference, while in others, such as *Alchemilla* (Morales-Briones et al., 2021), *Diplostephium* (Vargas et al., 2017) and *Lupinus* (Nevado et al., 2018), hybridization is predicted to have been a major factor underlying phylogenetic complexity. We add *Brachyotum* to this long and growing list of phylogenetically complex Andean plant radiations because it demonstrates a high level of gene tree discordance and uncertainty across our analyses.

Multiple unique features of the high tropical Andes where *Brachyotum* evolved may have contributed to its phylogenetic complexity. These include an elevational age gradient, with higher ecosystems that are home to younger taxa (Pennington et al., 2010), incredibly high topographic relief, and climatic heterogeneity (Mutke et al., 2014). The novel niches at higher elevations may have promoted the rapid radiation of *Brachyotum*, resulting in adaptations to lower temperatures, higher UV radiation, and hummingbird pollination, the latter of which is favored at high elevations in the tropics where insect pollinators are less effective (Renner, 1989; Stiles et al., 1992; Meyer et al., 2021; Dellinger et al., 2023). In addition, Pleistocene interglacial cycles may have repeatedly led to incipient species of *Brachyotum* being in secondary contact with congeners and close relatives in *Chaetogastra*, resulting in subsequent hybridization (Nevado et al., 2018; Flantua et al., 2019). Consistent with this scenario, we identify multiple instances of hybridization and introgression here, including at least one instance of a putative intergeneric hybrid sharing features of both parents (i.e., *Chaetogastra* sp., Figure 1M, N). Future work in the group may include testing the role of such glacial cycles on historical introgression using deeply sampled population genetic data.

Should *Brachyotum* still be recognized as a genus if neither monophyly nor paraphyly is definitively supported?

Different analyses, using models with different underlying assumptions, inferred *Brachyotum* variably as either monophyletic (in the MSC summary method and the concatenated analysis) or paraphyletic (in the MSC co-estimation, phylogenetic network, and plastid topology). Neither the inclusion of more data, nor the use of models that resemble different evolutionary scenarios relevant to *Brachyotum*, provided a definite answer here on its monophyly. Various morphological characters support the recognition of *Brachyotum* as a distinctive group. The unique arrangement of characters for the genus, including a hummingbird pollination syndrome (i.e., nectar production, pseudotubular flowers, and reduction of stamen appendages and pedoconnectives, Figure 1A–H, J–L) and its occurrence in high elevation grasslands, suggests a single origin for the genus. The alternative scenarios that are suggested by alternative topologies are less parsimonious, involving multiple losses or independent origins of these traits. In the case of pollination syndromes, our topologies in which *Brachyotum* is not monophyletic (Figure 4) imply either two parallel reversals to a bee pollination syndrome (e.g., pollen-rewarding, open flowers, Figure 1I, M–O) for the MSC co-estimation phylogeny (Figure 4), or two origins of hummingbird pollination with one reversal (both for the MSC co-estimation and phylogenetic network approaches). Additionally, chromosome counts lend support to a monophyletic *Brachyotum*: *Brachyotum* species consistently have 20 chromosomes ($2n = 20$ for a sampled 12.7% of species), while the *incertae sedis* species have 18 chromosomes ($2n = 18$ for *T. brevisepala*), which is also common in *Chaetogastra* (Meyer et al., 2021). Given the remaining

uncertainty on the monophyly of *Brachyotum*, but considering its distinct morphology and ecology, we do not advocate for any specific taxonomic change involving this genus.

Differing phylogenetic positions of incertae sedis species of *Tibouchina* between the nuclear and plastid topologies may suggest their involvement in historic introgression events with *Brachyotum* species. As we mentioned already, these taxa currently are recognized under their historical genus *Tibouchina*. These species have fallen within *Brachyotum* in previous studies that combined nuclear and plastid data in a concatenation framework (Guimarães et al., 2019; Meyer et al., 2021). However, we argue that these species—*T. bicolor*, *T. brevisepala*, and *T. calycina*—could be transferred to *Chaetogastra* because they are consistently placed as closely related to other *Chaetogastra* in the "*Brachyotum* + allies" clade in nuclear topologies. However, given that the phylogenetic status of *Brachyotum* remains unsolved here, we recommend assessing a more comprehensive sampling of the *Chaetogastra* + *Brachyotum* clade before taxonomic transfers are made (e.g., the incertae sedis species *T. octopetala* was not included in our sampling).

Despite our inability to definitively support *Brachyotum* as a monophyletic genus, our evidence indicates that its members are closely related. We hypothesize that selection may have acted on a small number of genomic regions in relatively few generations resulting in a novel set of traits that justified the recognition of *Brachyotum* as a distinct taxon. If so, either introgression or ILS involving *Chaetogastra* species may help to explain the nonmonophyly of the morphologically distinct *Brachyotum* species. The origin of this high Andean radiation remains unresolved and needs new approaches in phylogenomics.

CONCLUSIONS

Our study revealed that the evolutionary history of *Brachyotum* is marked by pervasive topological conflict, gene tree discordance, and robust evidence of introgression. Across data sets and inference methods, there is conflicting signal for both the monophyly and the paraphyly of *Brachyotum*, accompanied by low statistical support for either result. Phylogenetic network analysis, cytonuclear discordance, and quartet-based metrics suggest historical introgression, in agreement with previous hypotheses based on morphology. Signals for recent, ongoing hybridization are provided by assessment of locus heterozygosity and allele divergence, which again agree with interpretations of morphology.

Our findings emphasize the complexity of the phylogenomics of plant radiations in the Andes. We argue that the difficulty in resolving evolutionary relationships for *Brachyotum* stem not only from incomplete lineage sorting, but also from a combination of introgression, recent and rapid radiation, and model limitations—factors that are likely common in the study of Andean plant radiations. We add *Brachyotum* to the growing body of evidence that Andean plant radiations are challenging systems for phylogenetics and taxonomy. Future work should incorporate a broader sampling of the genome and assessing rate heterogeneity, while incorporating further model advancements to address the complexity in the evolutionary relationships of this emblematic high Andean clade.

AUTHOR CONTRIBUTIONS

D.P.B.: conceptualization and design, sampling, lab work, data analysis, funding, writing and draft; A.M.B.: conceptualization and design, data analysis; F.A.M.: sampling, sequencing, funding; L.M.: sequencing, funding; C.U.U.: sampling, funding; L.P.L.: conceptualization and design, funding, writing draft. All authors reviewed the manuscript.

ACKNOWLEDGMENTS

We thank Lucas Bacci for his support on logistics with DNA isolates and file organization, Johanna R. Jantzen and Léo-Paul M.J. Dagallier for guidance on the scope of the probe set, Elanor Fuller and Grace Brandhurst for support with bioinformatics, Jeremy M. Brown for guidance with phylogenomics and Bayesian inference, Edgardo Ortiz for guidance with CAPTUS, and the LSU High Performance Computing staff for assistance in maximizing the use of resources of the cluster. We thank the SERFOR from Peru for its permission to collect (0637-2011-AG-DGFFS/DGEFFS; 0637-2011-AG-DGFFS/DGEFFS; 007-2017-SERNANP-PNYCh; 031-2019-MINAGRI-SERFOR/DGGSPFFS-DGSPF; 091-2020-MIGA GRI-SERFOR/DGGSPFFS) and the Ministry of Environment of Ecuador for research permits (017-07 IC-FLO-DNBAPVS/MA, 007-09 IC-FLO-DNB/MA). Thanks to the herbaria staff of the Museo de Historia Natural UNMSM (USM), Missouri Botanical Garden (MO) and New York Botanical Garden (NY) for access to their collections. Portions of this research were done using high performance computing resources provided by Louisiana State University (http://www.hpc.lsu.edu). This work largely benefited with feedback from Laura A. Frost, Luis Y. Santiago-Rosario, Laymon Ball, Janet K. Mansaray, Katherin Arango-Gómez, Aislinn Mumford, Ya Yang and Rebekah A. Mohn. D.P.B. deeply appreciates the constant peer guidance from Sheila Rodríguez-Machado and Giovani Hernández-Canchola, who also helped us understand Mexican mountain systems. We also thank Janelle Burke, Karla Sosa, William Farfan, Sebastián Riva, and Asunción Cano for logistical help during fieldwork in Peru and Diana Fernández, Raffaella Ansaloni, Danilo Minga, Zhofre Aguirre, Jorge Caranqui, and David Neill for logistic support during fieldwork in Ecuador. We also thank Frank Almeda, Mac Alford, Michael Nee, and Jesus de Santiago for tissue of selected species. This work was partially funded by the Society of Systematic Biologists (SSB), the International Association of Plant Taxonomists (IAPT), and the American Society of Plant Taxonomists (ASPT) with grants to D.P.B., the National Science Foundation with grants to L.M. (DEB 2002270) and F.A.M. (DEB-2001357; DEB-0818399), and L.P.L. (DEB- 2055525), and the National Geographic Society to C.U.U. (NGS Grant # 829107). We thank our reviewers for their valuable feedback.

CONFLICT OF INTEREST STATEMENT

The authors declare they have no conflict of interest.

DATA AVAILABILITY STATEMENT

Final data sets and a Spanish version of this manuscript are available on Zenodo: https://doi.org/10.5281/zenodo.15678287. Code is on GitHub: https://github.com/dparedesburneo/Brachyotum_phylogenomics. Generated raw data are deposited in NCBI Bioproject PRJNA1442851.

ORCID

Diego Paredes-Burneo https://orcid.org/0000-0003-1246-928X
Ana M. Bedoya https://orcid.org/0000-0002-6245-3481
Fabián A. Michelangeli https://orcid.org/0000-0001-7348-143X
Lucas C. Majure https://orcid.org/0000-0002-0369-1906
Carmen Ulloa Ulloa https://orcid.org/0000-0003-2453-8131
Laura P. Lagomarsino https://orcid.org/0000-0003-4537-0761

REFERENCES

Aguirre-Santoro, J., A. Zuluaga, E. Stonesmyth, J. Betancur, and R. S. Jabaily. 2024. Phylogenomics of *Puya* (Bromeliaceae): evolution in the Andean slopes and sky island ecosystems. *Journal of Systematics and Evolution* 62: 257–274.

Baker, W. J., P. Bailey, V. Barber, A. Barker, S. Bellot, D. Bishop, L. R. Botigue, et al. 2022. A comprehensive phylogenomic platform for exploring the angiosperm tree of life. *Systematic Biology* 71: 301–319.

Baker, W. J., S. Dodsworth, F. Forest, S. W. Graham, M. G. Johnson, A. McDonnell, L. Pokorny, et al. 2021. Exploring Angiosperms353: an open, community toolkit for collaborative phylogenomic research on flowering plants. *American Journal of Botany* 108: 1059–1065.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.

Boschman, L. M. 2021. Andean mountain building since the Late Cretaceous: A paleoelevation reconstruction. *Earth-Science Reviews* 220: 103640.

Bouckaert, R., J. Heled, D. Kühnert, T. Vaughan, C. H. Wu, D. Xie, M. A. Suchard, et al. 2014. BEAST 2: A software platform for bayesian evolutionary analysis. *PLoS Computational Biology* 10: 1–7.

Breinholt, J. W., S. B. Carey, G. P. Tiley, E. C. Davis, L. Endara, S. F. Mcdaniel, L. G. Neves, et al. 2021. A target enrichment probe set for resolving the flagellate land plant tree of life. *Applications in Plant Sciences* 9: e11406.

Brown, J. W., J. F. Walker, and S. A. Smith. 2017. Phyx: Phylogenetic tools for unix. *Bioinformatics* 33: 1886–1888.

Cai, L., Z. Xi, E. M. Lemmon, A. R. Lemmon, A. Mast, C. E. Buddenhagen, L. Liu, and C. C. Davis. 2021. The perfect storm: gene tree estimation error, incomplete lineage sorting, and ancient gene flow explain the most recalcitrant ancient angiosperm clade, Malpighiales. *Systematic Biology* 70: 491–507.

Chernomor, O., A. Von Haeseler, and B. Q. Minh. 2016. Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology* 65: 997–1008.

Degnan, J. H., and N. A. Rosenberg. 2006. Discordance of species trees with their most likely gene trees. *PLoS Genetics* 2: 762–768.

Degnan, J. H., and N. A. Rosenberg. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology and Evolution* 24: 332–340.

Degnan, J. H., and L. A. Salter. 2005. Gene tree distributions under the coalescent process. *Evolution* 59: 24–37.

Dellinger, A. S., A. M. Hamilton, C. A. Wessinger, and S. D. Smith. 2023. Opposing patterns of altitude-driven pollinator turnover in the tropical and temperate Americas. *American Naturalist* 202: 152–165.

Dellinger, A. S., C. Kopper, K. Kagerl, and J. Schönenberger. 2022. Pollination in Melastomataceae: a family-wide update on the little we know and the much that remains to be discovered. *In* R. Goldenberg, F. A. Michelangeli, and F. Almeda [eds.], Systematics, evolution, and ecology of Melastomataceae, 585–607. Springer International, Cham, Switzerland.

Dellinger, A. S., L. Lagomarsino, F. Michelangeli, S. Dullinger, and S. D. Smith. 2024. The sequential direct and indirect effects of mountain uplift, climatic niche, and floral trait evolution on diversification dynamics in an Andean plant clade. *Systematic Biology* 73: 594–612.

Doyle, J. J., and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29: 1969–1973.

Flantua, S. G. A., A. O'Dea, R. E. Onstein, C. Giraldo, and H. Hooghiemstra. 2019. The flickering connectivity system of the north Andean páramos. *Journal of Biogeography* 46: 1808–1825.

Frankel, L. E., and C. Ané. 2023. Summary tests of introgression are highly sensitive to rate variation across lineages. *Systematic Biology* 72: 1357–1369.

Frost, L. A., A. M. Bedoya, and L. P. Lagomarsino. 2024. Artifactual orthologs and the need for diligent data exploration in complex phylogenomic datasets: a museomic case study from the Andean flora. *Systematic Biology* 73: 308–322.

Givnish, T. J. 2015. Adaptive radiation versus 'radiation' and 'explosive diversification': Why conceptual distinctions are fundamental to understanding evolution. *New Phytologist* 207: 297–303.

Guevara-Andino, J. E., L. M. Dávalos, F. Zapata, M. J. Endara, D. D. Cotoras, J. Chaves, S. Claramunt, et al. 2025. Neotropics as a cradle for adaptive radiations. *Cold Spring Harbor Perspectives in Biology* 17: a041452.

Guimarães, P. J. F., F. A. Michelangeli, K. Sosa, and J. R. de Santiago Gómez. 2019. Systematics of *Tibouchina* and allies (Melastomataceae: Melastomateae): A new taxonomic classification. *Taxon* 68: 937–1002.

Guindon, S., J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O. Gascuel. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* 59: 307–321.

Hibbins, M. S., and M. W. Hahn. 2022. Phylogenomic approaches to detecting and characterizing introgression. *Genetics* 220: iyab173.

Hoang, D. T., O. Chernomor, A. von Haeseler, B. Q. Minh, L. S. Vinh, and M. S. Rosenberg. 2017. UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution* 35: 518–522.

Jantzen, J. R., P. Amarasinghe, R. A. Folk, M. Reginato, F. A. Michelangeli, D. E. Soltis, N. Cellinese, and P. S. Soltis. 2020. A two-tier bioinformatic pipeline to develop probes for target capture of nuclear loci with applications in Melastomataceae. *Applications in Plant Sciences* 8: e11345.

Jantzen, J. R., P. J. F. Guimarães, L. C. Pederneiras, A. L. F. Oliveira, D. E. Soltis, and P. S. Soltis. 2022. Phylogenomic analysis of *Tibouchina* s.s.(Melastomataceae) highlights the evolutionary complexity of Neotropical savannas. *Botanical Journal of the Linnean Society* 199: 372–411.

Jiang, X., S. V. Edwards, and L. Liu. 2020. The multispecies coalescent model outperforms concatenation across diverse phylogenomic data sets. *Systematic Biology* 69: 795–812.

Johnson, M. G., E. M. Gardner, Y. Liu, R. Medina, B. Goffinet, A. J. Shaw, N. J. Zerega, and N. J. Wickett. 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in Plant Sciences* 4: 1600016.

Johnson, M. G., L. Pokorny, S. Dodsworth, L. R. Botigue, R. S. Cowan, A. Devault, W. L. Eiserhardt, et al. 2019. A universal probe set for

targeted sequencing of 353 nuclear genes from any flowering plant designed using *k*-medoids clustering. *Systematic Biology* 68: 594–606.

Junier, T., and E. M. Zdobnov. 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics* 26: 1669–1670.

Kubatko, L. S., and J. H. Degnan. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology* 56: 17–24.

Lagomarsino, L. P., L. Frankel, S. Uribe-Convers, A. Antonelli, and N. Muchhala. 2022. Increased resolution in the face of conflict: phylogenomics of the Neotropical bellflowers (Campanulaceae: Lobelioideae), a rapid plant radiation. *Annals of Botany* 129: 723–736.

Lee-Yaw, J. A., C. J. Grassa, S. Joly, R. L. Andrew, and L. H. Rieseberg. 2019. An evaluation of alternative explanations for widespread cytonuclear discordance in annual sunflowers (*Helianthus*). *New Phytologist* 221: 515–526.

Linder, C. R., and L. H. Rieseberg. 2004. Reconstructing patterns of reticulate evolution in plants. *American Journal of Botany* 91: 1700–1708.

Lopez, P., D. Casane, and H. Philippe. 2002. Heterotachy, an important process of protein evolution. *Molecular Biology and Evolution* 19: 1–7.

Maddison, W. P. 1997. Gene trees in species trees. *Systematic Biology* 46: 523–536.

Mallet, J. 2005. Hybridization as an invasion of the genome. *Trends in Ecology and Evolution* 20: 229–237.

Marques, D. A., J. I. Meier, and O. Seehausen. 2019. A combinatorial view on speciation and adaptive radiation. *Trends in Ecology and Evolution* 34: 531–544.

McKain, M. R., M. G. Johnson, S. Uribe-Convers, D. Eaton, and Y. Yang. 2018. Practical considerations for plant phylogenomics. *Applications in Plant Sciences* 6: e1038.

Meerow, A. W., E. M. Gardner, and K. Nakamura. 2020. Phylogenomics of the Andean tetraploid clade of the American Amaryllidaceae (subfamily Amaryllidoideae): unlocking a polyploid generic radiation abetted by continental geodynamics. *Frontiers in Plant Science* 11: 582422.

Mendes, F.K. and M.W. Hahn. 2018. Why concatenation fails near the anomaly zone. *Systematic Biology* 67: 158–169.

Meyer, F. S., M. Reginato, E. C. Smidt, J. R. De Santiago Gómez, F. A. Michelangeli, and R. Goldenberg. 2021. Phylogenetic relationships in *Brachyotum* and allies (Melastomataceae, Melastomateae): a reassessment of the limits of the genera. *Botanical Journal of the Linnean Society* 197: 170–189.

Michelangeli, F. A., P. J. F. Guimaraes, D. S. Penneys, F. Almeda, and R. Kriebel. 2013. Phylogenetic relationships and distribution of New World Melastomeae (Melastomataceae). *Botanical Journal of the Linnean Society* 171: 38–60.

Minh, B. Q., H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. Von Haeseler, and R. Lanfear. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* 37: 1530–1534.

Molloy, E., and T. Warnow. 2023. Large-scale species tree estimation. *In* L. S. Kubatko and L. L. Knowles [eds.], Species tree inference, 19–42. Princeton University Press, Princeton, NJ, USA.

Morales-Briones, D. F., B. Gehrke, C.-H. Huang, A. Liston, H. Ma, H. E. Marx, D. C. Tank, and Y. Yang. 2021. Analysis of paralogs in target enrichment data pinpoints multiple ancient polyploidy events in *Alchemilla* s.l. (Rosaceae). *Systematic Biology* 71: 190–207.

Morales-Briones, D. F., A. Liston, and D. C. Tank. 2018. Phylogenomic analyses reveal a deep history of hybridization and polyploidy in the Neotropical genus *Lachemilla* (Rosaceae). *New Phytologist* 218: 1668–1684.

Murillo-A., J., J. Valencia-D., C. I. Orozco, C. Parra-O., and K. M. Neubig. 2022. Incomplete lineage sorting and reticulate evolution mask species relationships in Brunelliaceae, an Andean family with rapid, recent diversification. *American Journal of Botany* 109: 1139–1156.

Mutke, J., R. Jacobs, K. Meyers, T. Henning, and M. Weigend. 2014. Diversity patterns of selected Andean plant groups correspond to topography and habitat dynamics, not orogeny. *Frontiers in Genetics* 5: 351.

Nauheimer, L., N. Weigner, E. Joyce, D. Crayn, C. Clarke, and K. Nargar. 2021. HybPhaser: a workflow for the detection and phasing of hybrids in target capture data sets. *Applications in Plant Sciences* 9: e11441.

Nevado, B., N. Contreras-Ortiz, C. Hughes, and D. A. Filatov. 2018. Pleistocene glacial cycles drive isolation, gene flow and speciation in the high-elevation Andes. *New Phytologist* 219: 779–793.

Nute, M., J. Chou, E. K. Molloy, and T. Warnow. 2018. The performance of coalescent-based species tree estimation methods under models of missing data. *BMC Genomics* 19: 286.

Ogilvie, H. A., R. R. Bouckaert, and A. J. Drummond. 2017. StarBEAST2 brings faster species tree inference and accurate estimates of substitution rates. *Molecular Biology and Evolution* 34: 2101–2114.

Ortiz, E. M., A. Höwener, G. Shigita, M. Raza, O. Maurin, A. Zuntini, F. Forest, et al. 2023. A novel phylogenomics pipeline reveals complex pattern of reticulate evolution in Cucurbitales. *bioRxiv* http://biorxiv.org/lookup/doi/10.1101/2023.10.27.564367 [preprint].

Pamilo, P., and M. Nei. 1988. Relationships between gene trees and species trees. *Molecular Biology and Evolution* 5: 568–583.

Pang, X. X., and D. Y. Zhang. 2023. Impact of ghost introgression on coalescent-based species tree inference and estimation of divergence time. *Systematic Biology* 72: 35–49.

Pease, J. B., J. W. Brown, J. F. Walker, C. E. Hinchliff, and S. A. Smith. 2018. Quartet sampling distinguishes lack of support from conflicting support in the green plant tree of life. *American Journal of Botany* 105: 385–403.

Pennington, R. T., M. Lavin, T. Särkinen, G. P. Lewis, B. B. Klitgaard, and C. E. Hughes. 2010. Contrasting plant diversification histories within the Andean biodiversity hotspot. *Proceedings of the National Academy of Sciences, USA* 107: 13783–13787.

Pérez-Escobar, O. A., A. Zizka, M. A. Bermudez, A. S. Meseguer, F. L. Condamine, C. Hoorn, H. Hooghiemstra, et al. 2022. The Andes through time: evolution and distribution of Andean floras. *Trends in Plant Science* 27: 364–378.

Posada, D., and K. A. Crandall. 2002. The effect of recombination on the accuracy of phylogeny estimation. *Journal of Molecular Evolution* 54: 396–402.

Pouchon, C., A. Fernández, J. M. Nassar, F. Boyer, S. Aubert, S. Lavergne, and J. Mavárez. 2018. Phylogenomic analysis of the explosive adaptive radiation of the *Espeletia* complex (Asteraceae) in the tropical Andes. *Systematic Biology* 67: 1041–1060.

Pouchon, C., S. Lavergne, A. Fernandez, A. Alberti, S. Aubert, and J. Mavarez. 2021. Phylogenetic signatures of ecological divergence and leapfrog adaptive radiation in *Espeletia*. *American Journal of Botany* 108: 113–128.

Rambaut, A., A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* 67: 901–904.

Ranwez, V., S. Harispe, F. Delsuc, and E. J. P. Douzery. 2011. MACSE: Multiple Alignment of Coding SEquences accounting for frameshifts and stop codons. *PLoS One* 6: e22594.

Renner, S. S. 1989. A survey of reproductive biology in Neotropical Melastomataceae and Memecylaceae. *Annals of the Missouri Botanical Garden* 76: 496.

Rieseberg, L. H., and D. E. Soltis. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants* 5: 65–84.

Segovia, R. A., E. Aguirre-Mazzi, C. E. Edwards, A. G. Linan, A. Fuentes, A. Chaspuengal, K. G. Dexter, et al. 2025. Phylogeny of *Weinmannia* (Cunoniaceae) reveals the contribution of the southern extratropics to tropical Andean biodiversity. *Molecular Phylogenetics and Evolution* 211: 108399.

Simmons, M. P., and J. Gatesy. 2021. Collapsing dubiously resolved gene-tree branches in phylogenomic coalescent analyses. *Molecular Phylogenetics and Evolution* 158: 107092.

Simões, M., L. Breitkreuz, M. Alvarado, S. Baca, J. C. Cooper, L. Heins, K. Herzog, and B. S. Lieberman. 2016. The evolving theory of evolutionary radiations. *Trends in Ecology and Evolution* 31: 27–34.

Smith, M. R. 2020. Information theoretic generalized Robinson–Foulds metrics for comparing phylogenetic trees. *Bioinformatics* 36: 5007–5013.

Solís-Lemus, C., and C. Ané. 2016. Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. *PLoS Genetics* 12: e1005896.

Solís-Lemus, C., P. Bastide, and C. Ané. 2017. PhyloNetworks: a package for phylogenetic networks. *Molecular Biology and Evolution* 34: 3292–3298.

Sosa-Pivatto, M., G. A. Camps, M. C. Baranzelli, A. Espíndola, A. N. Sérsic, and A. Cosacov. 2020. Connection, isolation and reconnection: Quaternary climatic oscillations and the Andes shaped the phylogeographical patterns of the Patagonian bee *Centris cineraria* (Apidae). *Biological Journal of the Linnean Society* 131: 396–416.

Steenwyk, J. L., Y. Li, X. Zhou, X.-X. Shen, and A. Rokas. 2023. Incongruence in the phylogenomics era. *Nature Reviews Genetics* 24: 834–850.

Stiles, F. G., A. V. Ayala, and M. Girón. 1992. Polinización de las flores de *Brachyotum* (Melastomataceae) por dos especies de *Diglossa* (Emberizidae). *Caldasia* 17: 47–54.

Straub, S. C. K., M. Parks, K. Weitemier, M. Fishbein, R. C. Cronn, and A. Liston. 2012. Navigating the tip of the genomic iceberg: next-generation sequencing for plant systematics. *American Journal of Botany* 99: 349–364.

Tabatabaee, Y., C. Zhang, T. Warnow, and S. Mirarab. 2023. Phylogenomic branch length estimation using quartets. *Bioinformatics* 39: I185–I193.

Toups, B. S., R. C. Thomson, and J. M. Brown. 2025. Complex models of sequence evolution improve fit, but not gene tree discordance, for tetrapod mitogenomes. *Systematic Biology* 74: 86–100.

Tricou, T., E. Tannier, and D. M. De Vienne. 2022. Ghost lineages highly influence the interpretation of introgression tests. *Systematic Biology* 71: 1147–1158.

Vankan, M., S. Y. W. Ho, and D. A. Duchêne. 2022. Evolutionary rate variation among lineages in gene trees has a negative impact on species-tree inference. *Systematic Biology* 71: 490–500.

Vargas, O. M., S. Madriñán, and B. Simpson. 2023. Allopatric speciation is more prevalent than parapatric ecological divergence in a recent high-Andean diversification (*Linochilus*: Asteraceae). *PeerJ* 11: e15479.

Vargas, O. M., E. M. Ortiz, and B. B. Simpson. 2017. Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-Andean diversification (Asteraceae: Astereae: *Diplostephium*). *New Phytologist* 214: 1736–1750.

Veranso-Libalah, M. C., R. D. Stone, G. Kadereit, and P. J. Fernandes Guimarães. 2022. Systematics and taxonomy of the tribe Melastomateae. *In* R. Goldenberg, F. A. Michalangeli, and F. Almeda [eds.], Systematics, evolution, and ecology of Melastomataceae, 429–463. Springer International, Cham, Switzerland.

Wang, Y., J. Huang, E. Li, S. Xu, Z. Zhan, X. Zhang, Z. Yang, et al. 2022. Phylogenomics and biogeography of *Populus* based on comprehensive sampling reveal deep-level relationships and multiple intercontinental dispersals. *Frontiers in Plant Science* 13: 1–17.

Whelan, S., B. P. Blackburne, and M. Spencer. 2011. Phylogenetic substitution models for detecting heterotachy during plastid evolution. *Molecular Biology and Evolution* 28: 449–458.

Whitfield, J. B., and P. J. Lockhart. 2007. Deciphering ancient rapid radiations. *Trends in Ecology and Evolution* 22: 258–265.

Wurdack, J. J. 1953. A revision of the genus *Brachyotum* (Tibouchineae: Melastomaceae). *Memoirs of the New York Botanical Garden* 8: 343–407.

Wurdack, J. J. 1965. Certamen Melastomataceis IX. *Phytologia* 11: 377–400.

Wurdack, J. J. 1988. Certamen Melastomataceis XXXVIII. *Phytologia* 64: 293–301.

Xi, Z., L. Liu, and C. C. Davis. 2016. The impact of missing data on species tree estimation. *Molecular Biology and Evolution* 33: 838–860.

Yang, Y., and S. A. Smith. 2014. Orthology inference in nonmodel organisms using transcriptomes and low-coverage genomes: improving accuracy and matrix occupancy for phylogenomics. *Molecular Biology and Evolution* 31: 3081–3092.

Yu, G., D. K. Smith, H. Zhu, Y. Guan, and T. T. Y. Lam. 2017. ggtree: An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* 8: 28–36.

Yu, Y., J. H. Degnan, and L. Nakhleh. 2012. The probability of a gene tree topology within a phylogenetic network with applications to hybridization detection. *PLoS Genetics* 8: e1002660.

Zhang, C., and S. Mirarab. 2022. Weighting by gene tree uncertainty improves accuracy of quartet-based species trees. *Molecular Biology and Evolution* 39: msac215.

Zhang, C., M. Rabiee, E. Sayyari, and S. Mirarab. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19: 153.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1. List of specimens and relevant information.

Appendix S2. Supplemental Figures S1–S8.

Figure S1. Plot of contig depth (coverage) per sample for plastid data set.

Figure S2. Features and quality of nuclear data set regarding recovery per sample and per locus.

Figure S3. Heatmap of recovery vs samples for the raw nuclear data set after assembly and before any filtering step applied.

Figure S4. Heatmap of the recovery vs samples for nuclear data set before doing orthology inference with MO.

Figure S5. Plot of number of taxa retrieved per ortholog inferred by the Monophyletic Outgroup (MO) pipeline.

Figure S6. MSC tree inferred by ASTRAL-IV.

Figure S7. MSC tree inferred by weighted ASTRAL.

Figure S8. Plot of the heuristic search from SNaQ.

Appendix S3. Supplemental tables S1 and S2.

Table S1. Branch support for the clades identified across all the different species trees from the nuclear data set.

Table S2. Quartet sampling (QS) scores for the concatenated tree in Figure 3.

How to cite this article: Paredes-Burneo, D., A. M. Bedoya, F. A. Michelangeli, L. C. Majure, C. Ulloa Ulloa, and L. P. Lagomarsino. 2026. Evidence of introgression amid phylogenetic conflict in *Brachyotum*, a plant radiation from the Tropical Andes. *American Journal of Botany* 113(6): e70213. https://doi.org/10.1002/ajb2.70213