

Original Article

Using integrative taxonomy to clarify species boundaries in *Diervilla* (bush-honeysuckle, Caprifoliaceae)

Katherine G. Mathews^{1,*} , Brandon Wheeler^{1,2}, Luiz Silveira¹

¹Department of Biology, Western Carolina University, 460 Memorial Drive, Cullowhee, NC 28723, United States

²North Carolina Botanical Garden, The University of North Carolina at Chapel Hill, CB 3375, Chapel Hill, NC 27599, United States

*Corresponding author. Department of Biology, Western Carolina University, 460 Memorial Drive, Cullowhee, NC 28723, United States. E-mail: kmathews@wcu.edu

ABSTRACT

The taxonomic limits across *Diervilla*, a genus endemic to Eastern North America, are puzzling, compounded by broad morphological variation within species which may overlap across species. Here we used an integrative taxonomic approach to determine the degree of genetic, morphological, and geographical overlap across *Diervilla* species. We combined phylogenetic and quantitative analysis of ddRAD sequences, multivariate morphometrics, and qualitative morphological comparisons, and gathered occurrence data from herbarium specimens and samples collected anew. Species delimitation methods consistently identified *D. lonicera* as distinct, but disagreed on the limits between *D. rivularis* and *D. sessilifolia*. However, all three species can be consistently separated genetically, as revealed by principal coordinates analysis of the ddRAD sequences, as well as morphologically, as revealed by linear discriminant analysis of morphometric data. The distinctiveness across the three species was further confirmed by separate permutational analyses of variance. Moreover, the three species are largely separated geographically: *D. lonicera* is predominantly northern in distribution and restricted to higher elevations at lower latitudes, while *D. rivularis* and *D. sessilifolia* are parapatric in the Southern Appalachians. Informed by our integrative analyses, we provide information on geographical distributions and a dichotomous key for *Diervilla* species.

Keywords: Caprifoliaceae; *Diervilla*; integrative taxonomy; molecular phylogeny; multivariate morphometrics; RADSeq; species delimitation

INTRODUCTION

The Southern Appalachian Mountains of the Southeastern USA house high biodiversity and endemism for historical, geological, climatic, and geographical reasons (Estill and Cruzan 2001). For example, this region is at a crossroads of the extreme ends of many species' geographical ranges, leading to the accumulation of species from both northern and southern North America (Soltis *et al.* 2006). In addition, this area may have served as a refuge for many northern species during the most recent periods of glaciation. Several authors have hypothesized that during the last ice ages (between ~1.9 million and 10 000 years ago), northern plants were forced to move southward and found suitable habitat conditions in the mountains of the Southern Appalachians, which were not glaciated (Ramseur 1960, Wiser 1994). When the glaciers retreated and climates warmed, some plants migrated northward again, while others colonized the highest mountain peaks and are now considered relicts of an historical distribution. As a consequence of historical climate change and long-distance plant migrations, interesting distributional

patterns and phylogenetic relationships occur among Southern Appalachian plant species groups and their nonregional congeners (e.g. Quinlan *et al.* 2020, *Kalmia* L.; Stubbs *et al.* 2020, *Micranthes* Haw.).

In many cases, species complexes evolve in which there has been relatively little morphological differentiation in contrast to the amount of genetic differentiation within Southern Appalachian plant groups and their close relatives (e.g. Gould and Donoghue 2000, *Triosteum* L.; Spriggs *et al.* 2019, *Viburnum nudum* L. complex). This can also be accompanied by broad-ranging morphological variability within lineages. These processes often lead to taxonomic confusion and failure to recognize independently evolving lineages at the species level. Defining species as independently evolving metapopulation lineages (de Queiroz 2007) allows any of multiple lines of evidence to be used to identify and delimit such species.

The genus *Diervilla* Mill. (bush-honeysuckle, Caprifoliaceae) in eastern North America is an example of such a case, with three currently recognized species exhibiting broad morphological

variability within species and little morphological differentiation among species, leading to confusion in field identification and poor understanding of the geographical ranges of each species. Additionally, in past studies, using single chloroplast and nuclear genetic markers, scientists have been unable to resolve phylogenetic relationships among the species of *Diervilla*, or even to resolve populations into monophyletic species groups, leading to questioning of the taxonomic boundaries within this group (Kim and Kim 1999, Meeler 2018).

Divergence dating analyses have been done within Caprifoliaceae, but not within *Diervilla*. *Diervilla* and its sister genus, the eastern Asian *Weigela* Thunb., display a classic Arcto-Tertiary disjunction (e.g. Wen 1999). These two genera form subfamily Diervilloideae of Caprifoliaceae (Backlund and Pyck 1998, Kim and Kim 1999, Stevens 2001), which is differentiated from the rest of Caprifoliaceae by having dry, capsular (vs. fleshy) fruits. Fossil seeds of both taxa were found in the Mary Sachs Gravel deposits of Southern Banks Island, Northwest Territories, in Arctic North America, tentatively dated to the early Miocene (18 Mya) (Matthews and Ovenden 1990). The split between *Weigela* and *Diervilla* has been dated much earlier than the earliest known fossil, to the middle Eocene, at 56.59 ± 18.8 Mya (Bell and Donoghue 2005), although this disagrees with Donoghue et al. (2001), who placed the split in the Pliocene. Despite these important contributions to the natural history of the Caprifoliaceae in general, the age of divergence among the *Diervilla* species remains unknown.

Linnaeus (1753) described the northern bush honeysuckle as *Lonicera diervilla*; this was subsequently transferred to the new genus *Diervilla* as *D. lonicera* by Miller (1754, 1768). Between 1810 and 1940, four heterotypic names were described for entities in various parts of the broad range of *D. lonicera*, but all have subsequently been synonymized with *D. lonicera* [data retrieved 25 May 2024, from the Integrated Taxonomic Information System (ITIS) online database www.itis.gov, CC0 <https://doi.org/10.5066/F7KH0KBK>]. Two more species were later added to the genus: *D. sessilifolia* Buckley (1843), and *D. rivularis* Gatinger (1888). Three-quarters of a century later, Ahles (1964) reduced *D. rivularis* to a variety of *D. sessilifolia*, to which it most closely resembles. Hardin (1968) soon revived *D. rivularis*, thus returning the genus to three species, two endemic to the Southeastern USA (*D. sessilifolia* Buckl., *D. rivularis* Gatt.) and one mostly northeastern species, *D. lonicera* P.Miller, that also has a relictual distribution in high elevations of the Southern Appalachians.

Modern floras recognize all three species of *Diervilla* but have questioned whether *D. sessilifolia* and *D. rivularis* were conspecific (e.g. TN Flora Committee 2015). Because of high morphological variability within the southeastern endemics, as well as morphological similarity between the two southern endemics, available field identification keys are ineffective and sometimes misleading. They rely mainly on few, potentially overlapping, quantitative characters (petiole length and sepal length) and subjective (overall hairiness) and variable discrete characters (leaf margins ciliate or not, and stems terete or square in cross-section) (Radford et al. 1968, Chester et al. 2015, Weakley 2023). Furthermore, several populations intermediate in characters based on these keys among all three taxa have been found by field botanists, causing difficulty in identification, uncertainty

in the geographical distribution of each taxon, and calling into question their distinctness as species (D. Estes, pers. comm., A. Floden, pers. comm., O. Carson, pers. comm.). Therefore, revisionary studies are needed to facilitate identification of *Diervilla* species.

Integrative species delimitation studies have been successful at clarifying species boundaries and delimiting species in other taxa that display subtle or confusing morphological variation [e.g. *Heuchera parviflora* Bartl. group, Pantinople et al. 2024; *Symphytotrichum subulatum* (Michx.) G.L.Nesom group, Dahal et al. 2023; *Viburnum* L., Spriggs et al. 2019]. Spriggs et al. (2019) called for such studies, combining molecular, morphological, geographical, and ecological data, to potentially expose hidden diversity within known geographical hotspot regions, such as the Southern Appalachians. Pantinople et al. (2024) recently used such an approach to identify a cryptic, narrowly endemic novel species of *Heuchera* in Western North Carolina. Integrative taxonomic approaches, described as processes of using multiple lines of evidence to delimit species, retesting and refining species boundary hypotheses with new data sources, and searching for explanations for discordance among different datasets, were advocated independently by Dayrat (2005) and Will et al. (2005) and reviewed by Yeates et al. (2011).

To better understand the morphological variation in *Diervilla*, to clarify the taxonomy, species distributions, ages of divergence, and phylogenetic relationships in this genus, we integratively assessed the species limits in this genus. We obtained samples from multiple populations of each species throughout their ranges, examined and measured morphological traits on herbarium specimens, and sought new characters to aid in species identification and delimitation. We also generated a double-digest restriction-site associated DNA sequence (ddRADSeq) dataset from the samples to obtain a molecular dataset with potentially more informative variation and discriminatory power than single-gene sequences with which to study this group and used these data for clustering, admixture, species validation, phylogenetic and molecular dating analyses. With this study, we also hope to shed more light on the evolution of plant diversity in the Southern Appalachians.

MATERIALS AND METHODS

Sample collection

Collections of specimens in the field were made between 2016 and 2017, focusing on populations of all three species in the Southern Appalachians. During the summer of 2018, we made additional field collections, focusing on the purported range of *D. rivularis* in the Southeastern USA. Herbarium vouchers were made, and leaves were collected from living specimens and stored in silica gel desiccant for downstream analyses. Samples of northern populations of *D. lonicera* were obtained from herbarium specimens and living material. Vouchers of all samples are stored at WCUH and ACAD herbaria (Appendix). Taxon identifications were tentatively made on the basis of morphology. Because of the taxonomic confusion in the Southeastern USA, we sampled one to five individuals per locality, depending on the number of individuals seen, of tentatively identified *D. sessilifolia* ($N = 22$), *D. rivularis* ($N = 22$), and unknowns ($N = 9$). We sampled most of the known southeastern localities of *D. lonicera*

in the Southeast, but only one individual per locality when this taxon was easily distinguishable from the other two. We sampled from few localities throughout the northern portion of its range because of its consistent morphological distinctions from the other two taxa, and because, other than cultivated individuals, *D. sessilifolia* and *D. rivularis* have never been reported north of Western North Carolina. The total sample number for *D. lonicera* was $N = 10$. In addition, three samples of cultivated *Weigela florida* (Bunge) A. DC. were included as outgroups, giving a total of 57 samples.

Morphological observations for statistical analyses were made from all of the dried voucher specimens above. We measured continuous characters from leaves, flowers, and fruit (Table 1) using a dissecting microscope and ruler, and hairs were observed and imaged using a Leica M 205C dissecting microscope coupled with a DF5400 camera, and Leica Application Suite X. Characters to measure were chosen based on those used in type descriptions (Miller 1768, Buckley 1843, Gattinger 1888) and existing taxonomic keys (Radford *et al.* 1968, Chester *et al.* 2015, Weakley 2023) as well as additional characters we observed on herbarium specimens to potentially differentiate taxa. Morphological data were subject to multivariate analyses as described below. Because the morphometric analyses were sensitive to missing data, we opted to remove those samples that did not contain fruit character measurements, as fruit size and persistent sepal length were deemed highly important for species delimitation. The resulting dataset contained six *D. lonicera*, 16 *D. rivularis*, and 15 *D. sessilifolia* samples (all tentatively identified) giving a total of 37 samples. We also observed additional herbarium specimens loaned from CLEMS, JSU, NCU, NCSC, TENN, and UCHT, from throughout the ranges of each taxon to confirm the consistency of the morphological characters used in the subsequent synoptic key and to confirm the overall distribution of each taxon.

Molecular data collection

Total DNA was extracted from all 57 samples of silica-dried or herbarium specimens using DNEasy Plant mini kits (Qiagen, Valencia, CA, USA) following the manufacturer's protocol. Preserved leaf tissue was ground by hand in a mortar and pestle after freezing with liquid nitrogen at the beginning of the

process. Final extractions were stored in a -20°C freezer before ddRADSeq libraries were made.

We generated a ddRAD dataset to try to obtain a better resolved phylogeny and study the potential for incomplete lineage sorting or admixture, signalling lack of species differentiation or postspeciation introgression. Libraries were generated in the lab of Allan Strand at College of Charleston (Charleston, SC, USA). We prepared the genomic library by following a protocol in Parchman *et al.* (2012). Briefly, we digested genomic DNA with two restriction enzymes, EcoRI and MseI, and ligated adaptors with unique 8- to 10-bp barcodes to the digested DNA, assigning each sample a unique barcode. The restriction-ligation products were amplified in two independent PCRs using standard Illumina primers. The genomic library was size-selected using the Blue Pippin Prep (Sage Science, Beverly, MA, USA) to isolate the 300- to 500-bp fraction and sequenced as single-end reads on Illumina HiSeq 2000 and Illumina HiSeq 4000 platforms. All amplicons were pooled and sequenced at the University of Texas Genomic Sequencing and Analysis Facility (Austin, TX, USA).

Resulting sequences were demultiplexed and assembled *de novo* in ipyrad (Eaton and Overcast 2020). Raw fastq files were demultiplexed using a barcodes file. Most of the ipyrad default parameter settings were used for sequence assembly, including using a minimum Phred Qscore of 20 to retain bases for quality control. The `min_samples_locus` parameter was set to 25 (referred to hereafter as the 'min25' dataset) meaning for any locus to be retained, it must be present in at least 25 samples. This setting maximized the number of loci in the final dataset while minimizing the amount of missing data. All possible output formats were generated, including: (i) .phy file, containing 166 987 concatenated, sequenced nucleotides; (ii) .snp file, containing 11 998 SNPs (single nucleotide polymorphisms); and (iii) .usnp file, containing 2014 unlinked SNPs. The percentage missing data was 36.59% for the concatenated sequence matrix and 37.98% for the SNP matrix. Sequence coverage was visually inspected, and four samples were subsequently removed from the dataset because of poor sequencing results (accounting for much of the missing data), leaving 52 samples (Appendix), including 49 *Diervilla* samples (*D. sessilifolia*, $N = 21$; *D. rivularis*, $N = 19$; *D. lonicera*, $N = 9$) and three *Weigela* outgroup samples, in the final datasets.

Table 1. Morphological characters measured from dried herbarium specimens and used in the quantitative multivariate analyses in this study.

Category	Character name and description	Character abbreviation
Leaf	Petiole length (mm)	petiole_lg
	Blade width (mm)	blade_wd
	Blade length (mm)	blade_lg
Fruit	Pedicle length (mm)	fruit_pedicle_lg
	Fruit length (excluding beak) (mm)	fruit_lg
	Fruit width (mm)	fruit_wd
	Beak length (mm)	beak_lg
	Sepal length (mm)	fruit_sepal_lg
Inflorescence	Flower number per branch	total_flowers
	Average number of cymes per branch	avg_cymes

Species delimitation

Our integrative approach adopted multiple steps to identify operational taxonomic units. First, we ran a principal coordinates analysis [PCoA; function ‘cmdscale’; implemented in R::vegan (Oksanen *et al.* 2022)], which is an unsupervised statistical technique, on the genetic dataset, using concatenated ddRAD sequences generated from the ‘min25’ dataset converted to uncorrected ‘p’ distances in PAUP* (Swofford 2021). Using a newly discovered categorical character (hair type, see below) to give taxon assignments to our samples, we then ran a supervised technique, linear discriminant analysis [LDA; function ‘lda’, implemented in R::MASS (Venables and Ripley 2002)], on quantitative morphometric data, using a leave-one-out cross-validation approach to gauge the effectiveness of traditional measurements in sorting species out. This step is especially important given the historically challenging morphological discrimination of the three *Diervilla* species. Upon finding three clusters corresponding to named species groups in both analyses, we further tested the distinctiveness of these clusters with permutational analysis of variance (PERMANOVA; function ‘adonis2’, implemented in R::vegan), run separately on the genetic and morphometric datasets. Since PERMANOVA cannot distinguish between among-group multivariate mean differences and group dispersions (variances) (Warton *et al.* 2012), we complemented this step with separate permutation-based tests of multivariate homogeneity of group dispersions (variances) (function ‘permutest.betadisper’, implemented in R::vegan). All figures created in R were made with the packages ‘ggplot2’ (Wickham 2016) and ‘ggord’ (Beck 2024).

Additionally, we ran three complementary single-locus species delimitation methods on the genetic dataset: automatic barcode gap discovery [ABGD; Puillandre *et al.* 2012; implemented online (<https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html>)], multi-rate Poisson tree processes [mPTP; Kapli *et al.* 2017; implemented online (<https://mptp.h-its.org/#/tree>)], and the general mixed Yule-coalescent model [GMYC; Pons *et al.* 2006, implemented in R::splits (Ezard *et al.* 2021)]. ABGD is blind to phylogenetic relationships, whereas the latter two are phylogeny-aware. ABGD seeks for barcoding gaps in the dataset as surrogates for within- and across-species distances. mPTP accounts for within-species variation by allowing across-branch variation in divergence rates. Finally, GMYC utilizes a likelihood-based approach to estimate transition points between coalescent and speciation events.

To find ancestral groupings indicative of genetically independent groups that could represent species and to identify potential hybridization across species, we implemented the Bayesian program STRUCTURE (Pritchard *et al.* 2000) in the ipyrad pipeline on the unlinked SNP dataset. We removed the outgroups and used only one sample per population selected at random. STRUCTURE was run for 10 000 generations of burnin and 100 000 generations post burn-in. The *k* values (number of assumed a priori ancestral groups) ranged from 2 to 6 to test hypotheses of: two *Diervilla* species (*D. lonicera* + one southern species), three species as currently recognized, or four to six species (unrecognized species within any of the currently recognized species). Results from all runs were compiled and visualized using the R Shiny program ‘pophelper’ (Francis

2017). Optimal *k* values were chosen by selecting the number of clusters for which the slope of the log probability was highest (Δk), as described by Evanno *et al.* (2005).

Phylogenetic analysis

Maximum likelihood estimation was run on concatenated ddRAD sequences using RAxML GUI v.8.2.12 for Windows (Stamatakis 2018). The best evolutionary model (GTR+I+G4) was chosen by running ModelTest within RAxML. Branch support was estimated with 1000 rapid bootstrap replicates. Species tree estimation using the unlinked SNP dataset was performed using SVD Quartets in PAUP* to test for incomplete lineage sorting by comparing results from separate loci as opposed to the concatenated sequences. Clade support was estimated by bootstrap analysis.

Species distribution mapping

Species distributions were determined by using already georeferenced herbarium specimen labels or by georeferencing specimens in SERNEC (Data Portal. 2023. <http://www.sernecportal.org/index.php>. Accessed on multiple dates), including 185 *D. sessilifolia* specimens, 71 *D. rivularis* specimens, and 406 *D. lonicera* specimens. All species identifications were confirmed by K.G.M. based on results of the morphological and molecular analyses in this study. Maps were overlaid in R with elevational data [SRTM 90-m resolution data, using package ‘raster’ (Hijmans 2023)], Köppen–Geiger climate classification (1-km resolution; Beck *et al.* 2018), and bedrock type [Chorlton 2007 (compiled based on different resolutions)] to determine whether the taxa inhabit different elevational, soil, or climatic regions, which could provide further evidence for species delimitation.

Molecular divergence dating

BEAST2 v.2.7.3 (Bouckaert *et al.* 2019) was used to obtain divergence times for the clades of *Diervilla* using the unlinked SNP file. Two different MRCA (most recent common ancestor) priors were used: (i) using the preferred divergence date from Bell and Donoghue (2005), the age of the *Weigela*–*Diervilla* root was set to a lognormal distribution with offset = 37.79 (minimum constraint), *M* = 18.8, and *S* = 0.4, to achieve a mean of 56.6 Mya, and 2.5–97.5% quantiles (45.7–75.8 Mya); and (ii) using the Mary Sachs Gravel estimated fossil bed date of early Miocene (~18 Mya) as per Matthews and Ovenden (1990), the age of the *Diervilla* crown was set to a lognormal distribution with offset = 18 (minimum constraint), *M* = 1, and *S* = 1.25 (defaults), to give a mean of 23.9 Mya, and 2.5–97.5% quantiles (18.2–49.5 Mya). We used a GTR+gamma site model and an optimized relaxed clock model. The tree prior was a uniform rate Birth Death model with all other prior parameters left at default settings. The Markov chain Monte Carlo settings were chain length of 10⁸ generations and logging every 10 000 generations.

RESULTS

Morphological observations

Previous synoptic keys to the species of *Diervilla* relied mainly on overall hair density to distinguish the two Southeastern US



Figure 1. *Diervilla* hair types as viewed on twigs below peduncles and corresponding herbarium specimens of each taxon. A, *Diervilla sessilifolia* (soft, antrorsely curled hairs); B, *Diervilla rivularis* (stiff, straight hairs at right-angles to the stem); C, *Diervilla lonicera* (very short, soft, antrorsely curled hairs).

endemics, *D. rivularis* and *D. sessilifolia* (e.g. branchlets, leaves, pedicels, and calyx densely pubescent vs. branchlets, leaves, pedicels, and calyx glabrous, except for hairs on the twig angles; Weakley 2023). In our study of herbarium specimens, we found that overall hair density was inconsistent within and between purported *D. rivularis* and *D. sessilifolia* and could not reliably be used to separate them. Rather, ‘twig hair type’ observed on the internode below the inflorescence (Fig. 1) is the only consistent hair characteristic within and among all three taxa, as validated by visualizing this trait onto the phylogenetic tree (see below), and by multivariate statistical techniques run independently on genetic and morphometric data (see below). Typical *D. rivularis* has long, stiff hairs that project outward from the stems at a 90° angle [e.g. from the type locality of Lula Falls, Tennessee (TN)], whereas *D. sessilifolia* [e.g. from the type locality of North Carolina (NC) mountains] is glabrous or has short, soft hairs that are slightly curled toward the stem. *Diervilla lonicera* has the shortest hairs that also are slightly curled toward the stem. Figure 1 also shows the consistently different leaf shape (ovate) and flower number (few) of *D. lonicera* vs. the two strictly southeastern taxa (leaf shape lanceolate; flowers many) that can allow field taxonomists to distinguish these three in the

Southern Appalachians. While the presence of long-petiolate leaves is often used in keys to differentiate *D. lonicera* from the other two species, we have identified populations of *D. sessilifolia* (Polk Co., TN) and *D. rivularis* (Cheatham Co., TN) with long-petiolate leaves.

Species delimitation

The PCoA of genetic distance data showed three broadly separated clusters representing groups that are consistent with the morphological hair trait described above (Fig. 2A). LDA and cross-validation results based on morphometric data showed strong support and high model accuracy in sorting the three species [Fig. 2B; 0.8649, 95% CI: (0.7123, 0.9546); *P*-value of the model (Acc > NIR, where NIR is the proportion of the largest class in the dataset. For a model to be considered effective, its accuracy (Acc) should be significantly higher than the NIR. In this context, the *P*-value is the probability of randomly finding such an accurate model): 6.465e-08; Table 2].

In the LDA, *D. lonicera* is the most morphologically divergent of the three taxa on LD1, and samples of this taxon were correctly classified 100% of the time. *Diervilla sessilifolia* and *D. rivularis* were more similar to each other and occasionally were

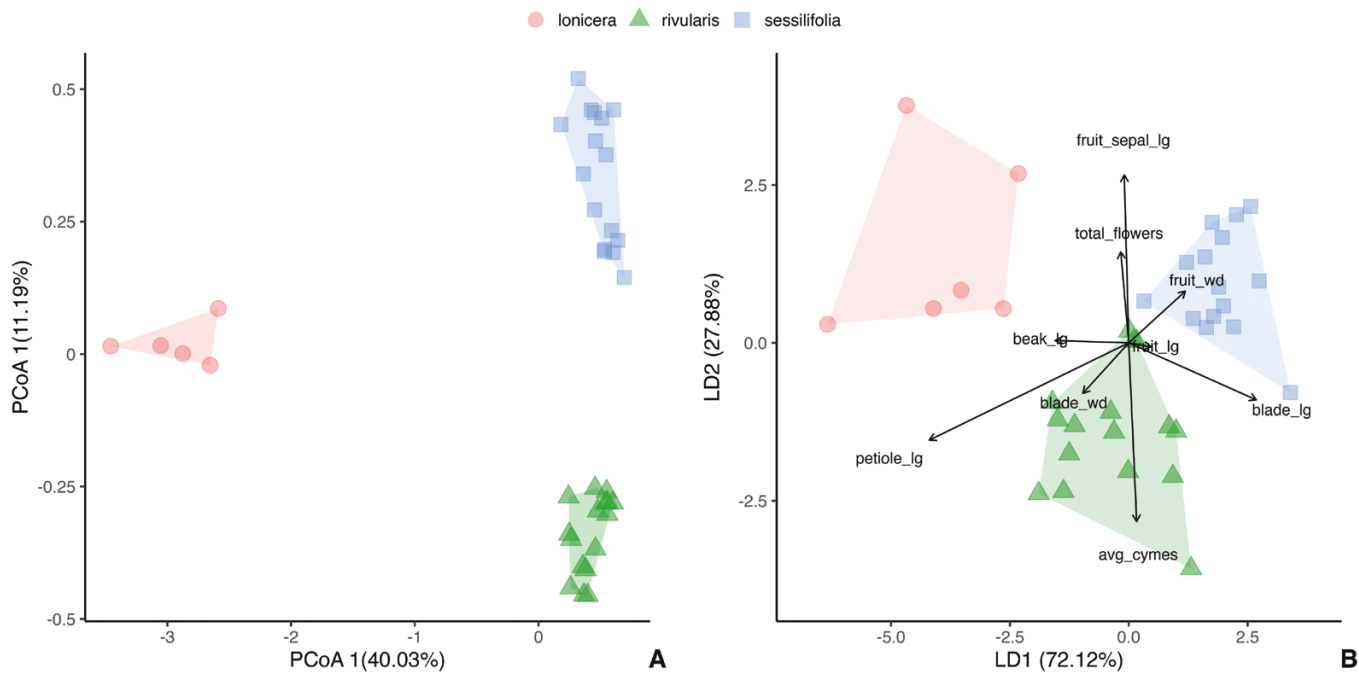


Figure 2. Principal coordinates analysis of genetic distances (A) and linear discriminant analysis from quantitative morphological data (B) reveal that the three *Diervilla* taxa are genetically and morphologically distinct. Coloured polygons represent the convex hulls containing all the sample points for each taxon. Taxon names were assigned to samples a priori based on twig hair type.

Table 2. Confusion matrix of the leave-one-out cross-validation technique of the linear discriminant analysis based on morphometric data, which calculates the final performance of the model by averaging the errors or accuracy across all rounds. The matrix shows a partial overlap between *D. rivularis* and *D. sessilifolia*, but a complete discrimination between these two and *D. lonicera*.

	<i>lonicera</i>	<i>rivularis</i>	<i>sessilifolia</i>
<i>lonicera</i>	6	0	0
<i>rivularis</i>	0	13	2
<i>sessilifolia</i>	0	3	13

misclassified in the cross-validation test; our model classified *D. rivularis* correctly 87% of the time and classified *D. sessilifolia* correctly 81% of the time (Table 2). Most of the variation in morphometric data was summarized in the first axis (72.12%). The characters that best separated *D. lonicera* from the other two taxa were petiole length (longest in *D. lonicera*) and leaf blade length (shortest in *D. lonicera*) (Fig.2B; Supporting Information Table S1). The characters that best separated *D. rivularis* and *D. sessilifolia* were fruit width (wider in *D. sessilifolia*) and leaf blade width (wider in *D. rivularis*) (Fig.2B; Table S1). Persistent sepal length was also shortest in *D. rivularis* and mean number of cymes per inflorescence branch was greatest in *D. rivularis*.

The three-species delimitation was further confirmed by PERMANOVAs run separately for sequence ($F = 40.55$, $P = .001$) and morphometric data ($F = 6.59$, $P = .001$) (Supporting Information Table S2). Dispersion effects were ruled out for either dataset (sequence: $F = 0.29$, $P = .75$; morphometrics: $F = 2.74$, $P = .08$; Table S3), despite a marginally greater morphometric dispersion found for *D. rivularis*. Pairwise

PERMANOVAs further confirmed the distinctiveness of the groups (Table S4).

The results of the species validation analyses were mixed, and in partial disagreement with the three-species arrangement found above (Supporting Information Fig. S1; Table S5). The small percentage divergences in the ddRADSeq dataset only yielded meaningful ABGD results (i.e. greater than one species) upon small α -values (relative barcode gap widths = 0.1; $P_{min} = 0.0001$; $P_{max} = 0.01$; Fig. S1), even using simple genetic distances. That is expected for widespread species, given that broad within-species variation will narrow barcode gaps between and among species. Both ABGD and mPTP delineated *D. lonicera* samples from the other two taxa but were unable to differentiate *D. sessilifolia* and *D. rivularis*, resulting in only two species being delineated. In strong contrast, GMYC delineated six species, two per named taxon, corresponding to the deepest divisions in each major clade of the phylogenetic trees, but not corresponding to any geographically defined groupings (Table S5).

In the STRUCTURE analysis, the mean log likelihood values of k [L(K)] plateaus at $k = 2-4$, but Δk , the second-order rate of change of the likelihood function with respect to k , indicates $k = 4$ as the most likely number of genetic groups (Evanno et al. 2005; Fig. 3). Figure 3 shows the k likelihood statistics for all k values tested and the genotype barplots by individual for $k = 2-4$. In the genotype bar plots, $k = 2$ divides the samples into a cluster containing all the *D. lonicera* samples and a cluster containing all the other samples, with small percentages of the ‘lonicera’ genotype found in two *D. sessilifolia* samples from Polk Co., TN. A value of $k = 3$ divides the samples into three clusters representing the three known taxa, with 1–25% admixture present between many samples of the two southeastern taxa, *D. sessilifolia* and *D. rivularis*. The $k = 4$ plot is similar to $k = 3$ but shows a small percentage of an

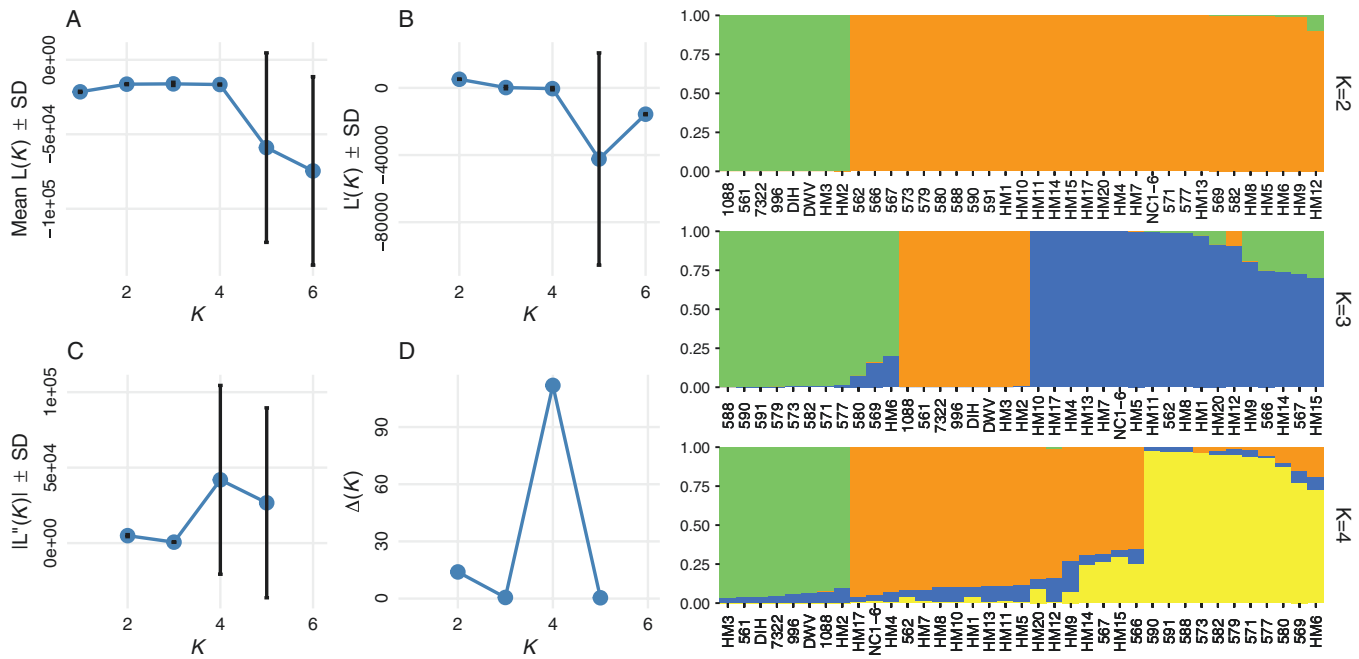


Figure 3. Graphs on the left show the STRUCTURE analysis posterior probability statistics output from the pophelterShiny application (Francis 2017). A, log likelihood (mean $\pm$ SD) plotted for each k value (2–6); B, rate of change of the likelihood distribution (mean $\pm$ SD); C, absolute values of the second-order rate of change of the likelihood distribution (mean $\pm$ SD); D, ΔK , the break in slope of the distribution of $L(K)$, the value likely to represent the true k according to Evanno *et al.* (2005). Bar plots on the right indicate by color the posterior probability of assignment of each individual sample to a particular cluster for $k = 2$ –4.

unknown cluster present in all the samples except one *D. rivularis*. Since no individuals show a majority of this fourth cluster and since it is found in almost all samples, it does not appear to reflect an unidentified species, but could represent a shared ancestral genotype. Therefore, we mapped the results based on $k = 3$, which had an equal $L(K)$ to $k = 4$ and allowed us to geographically explore more population structure than $k = 2$ (Fig. 4).

Phylogenetic results

A bootstrap support value of 94% was found for a clade containing all *D. lonicera* samples and 97% support for a clade containing both *D. sessilifolia* and *D. rivularis* samples (Fig. 5). The latter two taxa were also found to be reciprocally monophyletic, with 92% and 79% bootstrap support, respectively. Notably, each clade corresponded to the hair type character described above as assigned to each sample based on voucher specimens. Within each taxon clade, lineages received various levels of support, from 100% to less than 50%. Strongly supported clades often corresponded to samples taken near the same geographical locality, but not always.

Species tree estimation also showed the same species relationship results as the maximum likelihood analysis (tree not shown). Quartet assembly gave incompatible/compatible quartets results as 25%/75%, showing that incomplete lineage sorting did not affect the species tree results. Bootstrap support for the *D. lonicera* clade was 100%, while support for *D. sessilifolia* and *D. rivularis* clades were 90 and 76%, respectively.

Species distributions

Based on georeferenced herbarium specimen data and mapping overlaid with climatic and geological data layers (Fig. 6;

Supporting Information Fig. S2), *D. sessilifolia* is a Southern Appalachian endemic, found in the Blue Ridge province of the Appalachian Mountains, on ancient crystalline metamorphic, metasedimentary, or volcanic bedrock soils, at moderate to high elevations [1200–5400 ft (360–1650 m)] of west NC, east TN, northeast GA (Georgia), and northwest South Carolina (SC). *Diervilla rivularis* is found at moderate elevations (below 2000 ft), mainly on the southern Cumberland Plateau in TN, Alabama (AL), and GA, with one population on the Western Highland Rim of TN and several populations in western AL on the Coastal Plain. It grows on sedimentary bedrock (mainly sandstone) stream banks. These two species overlap in the Köppen–Geiger climate/vegetation zones they inhabit, but *D. rivularis* is restricted to the Cfa zone (temperate, no dry season, hot summer), whereas *D. sessilifolia* is found in Cfa and Cfb zones (temperate, no dry season, warm summer), but also in the Dfb zone (continental, no dry season, warm summer). All three species' distributions are allopatric, except for one occurrence, in Yancey Co., NC, on Potato Knob near Mt. Mitchell, at 5200 ft elevation (*T. Govus* 223, WCUH0031818 and *T. Govus* 222, WCUH31834), where both *D. sessilifolia* and *D. lonicera* occur.

Diervilla lonicera has the broadest and most northern geographical distribution. It is found in shady forests in the eastern US and Canada, from east TN, west NC and west VA (mostly 4000 ft and above in the Southern Appalachians) north to Minnesota (MN), north and west to Ontario, MN, and the Great Lakes region, and south to northern Illinois (IL). Based on our environmental mapping analysis, *D. lonicera* is found in the Dfa (continental, no dry season, hot summer) and Dfb Köppen–Geiger climate/vegetation zones, with the exception of a few specimens from the mid-Atlantic coastal plain of the USA

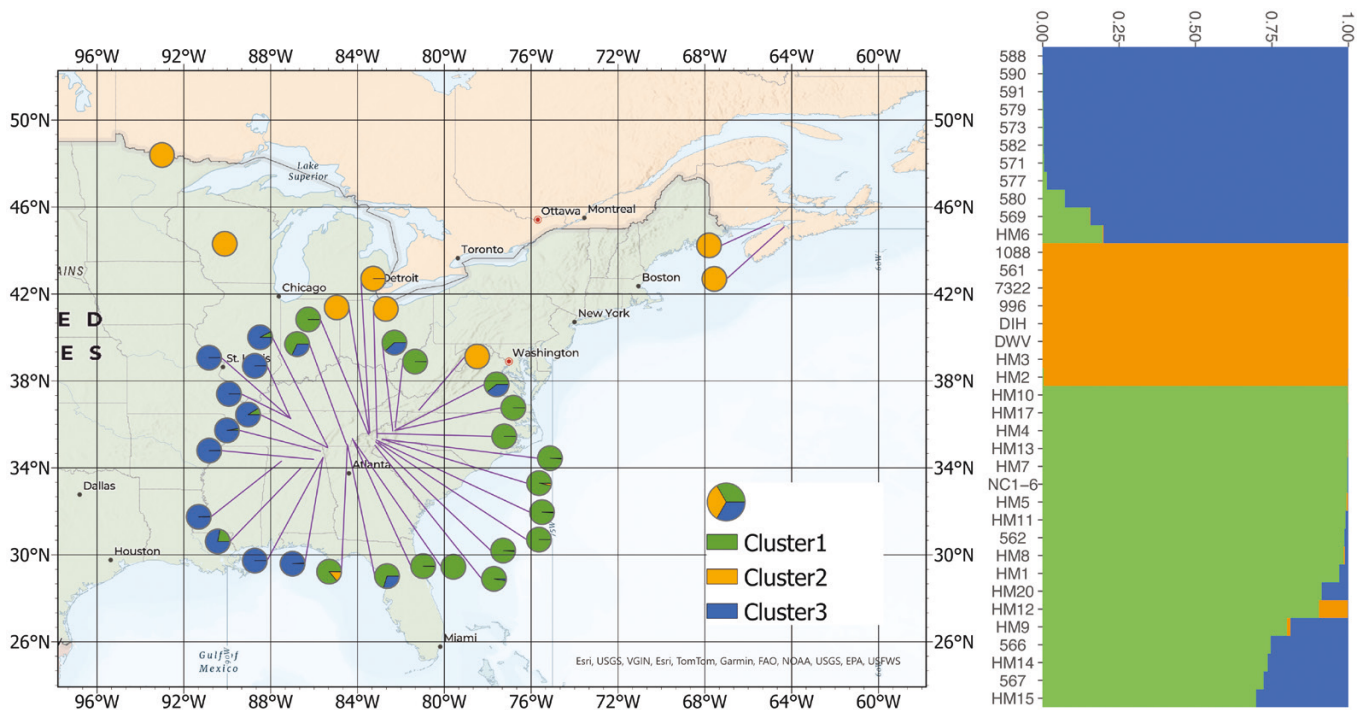


Figure 4. STRUCTURE analysis plot and mapped samples results for $k = 3$. In the inset, cluster 1 corresponds to the primary *D. sessilifolia* genotype, cluster 2 to the *D. lonicera* genotype, and cluster 3 to the primary *D. rivularis* genotype. Admixture (< 1 –25%) is evident in the samples with mixed genetic groups on the map and in the plot on the right.

in the Cfa zone. In the Southern Appalachians, it inhabits the same Appalachian Mountain bedrock types as *D. sessilifolia*.

Molecular divergence dating

The results from both calibration methods were consistent with age estimates for other high-elevation, Southern Appalachian species with northern relatives (e.g. Quinlan et al. 2020, Stubbs et al. 2020) and were consistent with each other within the 95% confidence intervals. The fossil calibration results (Fig. 7) had much narrower confidence intervals and estimated slightly older ages for the major divergences in *Diervilla*. The northern–southern *Diervilla* split was estimated at 21.03 (18.04–27.46) Mya, in the early Miocene. The age of extant *D. lonicera* was estimated at 7.76 (2.94–12.06) Mya, and the *D. sessilifolia*–*D. rivularis* split was estimated to be 8.61 (2.54–13.77) Mya. Calibrating the tree with the Bell and Donoghue estimate for the *Weigela*–*Diervilla* divergence (56.59 ± 18.8 Mya) (tree not shown) gave very broad confidence intervals. The split between the mainly northern *D. lonicera* and the two Southern Appalachian species was estimated at 17.15 (3.5–37.7) Mya, in the middle Miocene. The age of extant *D. lonicera* was estimated at 6.95 (1.24–18.39) Mya, at the end of the Miocene. The *D. sessilifolia*–*D. rivularis* split was estimated to be 8.43 (1.7–21.9) Mya, also in the late Miocene.

DISCUSSION

Our study has shown that multiple species delimitation criteria support the separation of three lineages of *Diervilla*, including properties considered by authors of various species concepts to be important secondary species properties, such as phenotypic

distinction (Freudenstein et al. 2017, de Queiroz 2007, and references therein). First, our morphological analyses show that the species of *Diervilla* can be phenetically distinguished from each other, in the field and from herbarium sheets, if the correct suite of morphological characters is used (see synoptic key below). *Diervilla lonicera* is most easily distinguished from the two Southeastern US endemics by its having lax, few-flowered inflorescences (typically three but up to 13 flowers, vs. congested, many-flowered inflorescences, often with hundreds of flowers), leaf blades that are ovate (vs. typically lanceolate), and petioles 5–10 mm (vs. petioles 0–8 mm). Morphological states are overlapping in *D. sessilifolia* and *D. rivularis*, but they have different medians: *D. sessilifolia* typically has sessile leaves (median petiole length 0 mm vs. 4 mm) and longer persistent sepals (median 3 mm vs. 2 mm) compared to *D. rivularis*. We have also identified one qualitative morphological characteristic, hair type, that corresponds to the three clusters found in our multivariate analyses and to the three major *Diervilla* clades in our phylogeny.

Second, our phylogenetic analysis shows multiple populations of each species to be reciprocally monophyletic. Third, our environmental analysis of distribution data shows that the three species of *Diervilla* are geographically differentiated, with *D. rivularis* and *D. sessilifolia* separated by the Valley and Ridge province of eastern TN (Fig. 6). The majority of all three taxon ranges are associated with different temperature regimes from each other, corresponding to elevational and latitudinal gradients. Underlying bedrock type is not diagnostic for any species, except for age of the geological formations (Supporting Information Fig. S2), and all appear to prefer acidic soils. Finally,

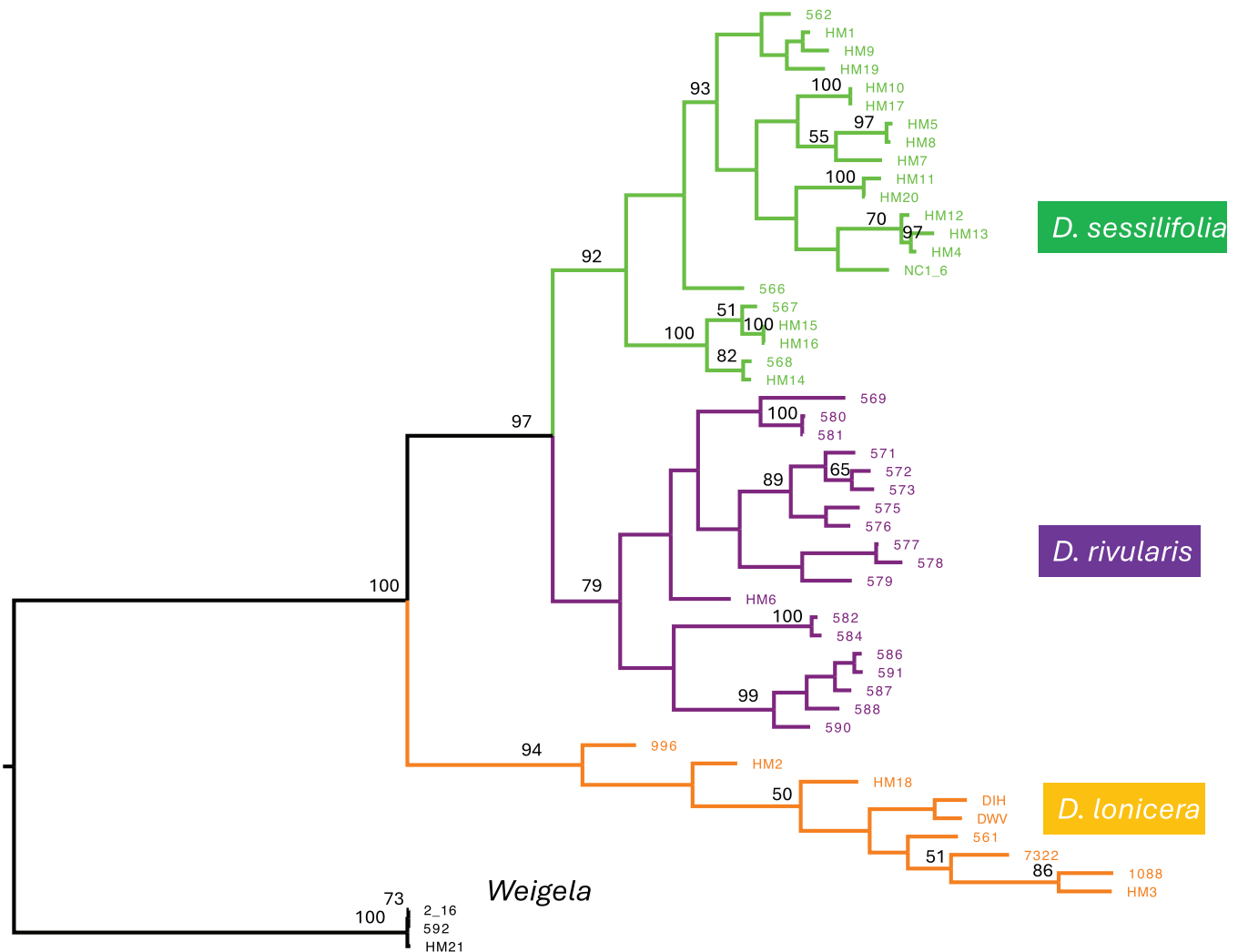


Figure 5. Maximum likelihood (ML) phylogenetic estimation based on concatenated ddRAD sequences showing that each major clade corresponds to the three *Diervilla* species. Numbers above branches are ML bootstrap support values $\geq 50\%$ from 1000 fast bootstrap replicates.

our multivariate and STRUCTURE analyses of molecular data strongly support three species of *Diervilla*.

We tested three molecular species delimitation methods with our ddRADSeq data to fully explore the molecular dataset. It has been shown that ABGD and GMYC tend to under- and oversplit species, respectively (Luo *et al.* 2018), and we see those differences when applied here to our *Diervilla* dataset, with ABGD finding only two species (as did mPTP) and GMYC finding six species. Our conclusions above could be seen as a consensus between these extremes and are based on the integration of morphological and geographical data with molecular data.

Even though *D. sessilifolia* and *D. rivularis* are currently allopatric, it is known that climatic oscillations of the Tertiary and Quaternary periods resulted in species migrations affecting the floristics and evolution of unglaciated eastern North America (Wen *et al.* 2010, Manos and Meireles 2015). Both *Diervilla* and its sister group, *Weigela*, were circumboreal according to fossils, and probably diverged during the early Miocene, when migratory routes between Asia and North America were cut off by the changing climate. Within *Diervilla*, our dating analysis is

consistent with the divergence of the sister species *D. sessilifolia* and *D. rivularis* from a common southern ancestor well before glaciation, with *D. sessilifolia* inhabiting the higher-elevation Southern Appalachian mountains and *D. rivularis* adapting to warmer, lower-elevation areas further south and west, while *D. lonicera* colonized northeastern North America. During the global cooling era, the two southern sister species could have been pushed into close contact with each other near the eastern and western edges of their ranges, possibly leading to hybridization and admixture as evidenced in our STRUCTURE analysis, while *D. lonicera* survived in the glacial refugia mountaintops of the southern Appalachians (and possible northern refugia). At the end of the ice ages, the two southern species would have reseeded into distinctive, allopatric ranges. *Diervilla lonicera* would have expanded its range and recolonized the north, while leaving behind high-elevation, southern populations as glacial relicts. An in-depth population genetics study utilizing more samples from the northern part of the range would be needed to explore relationships and migratory history among *D. lonicera* southern and northern populations.

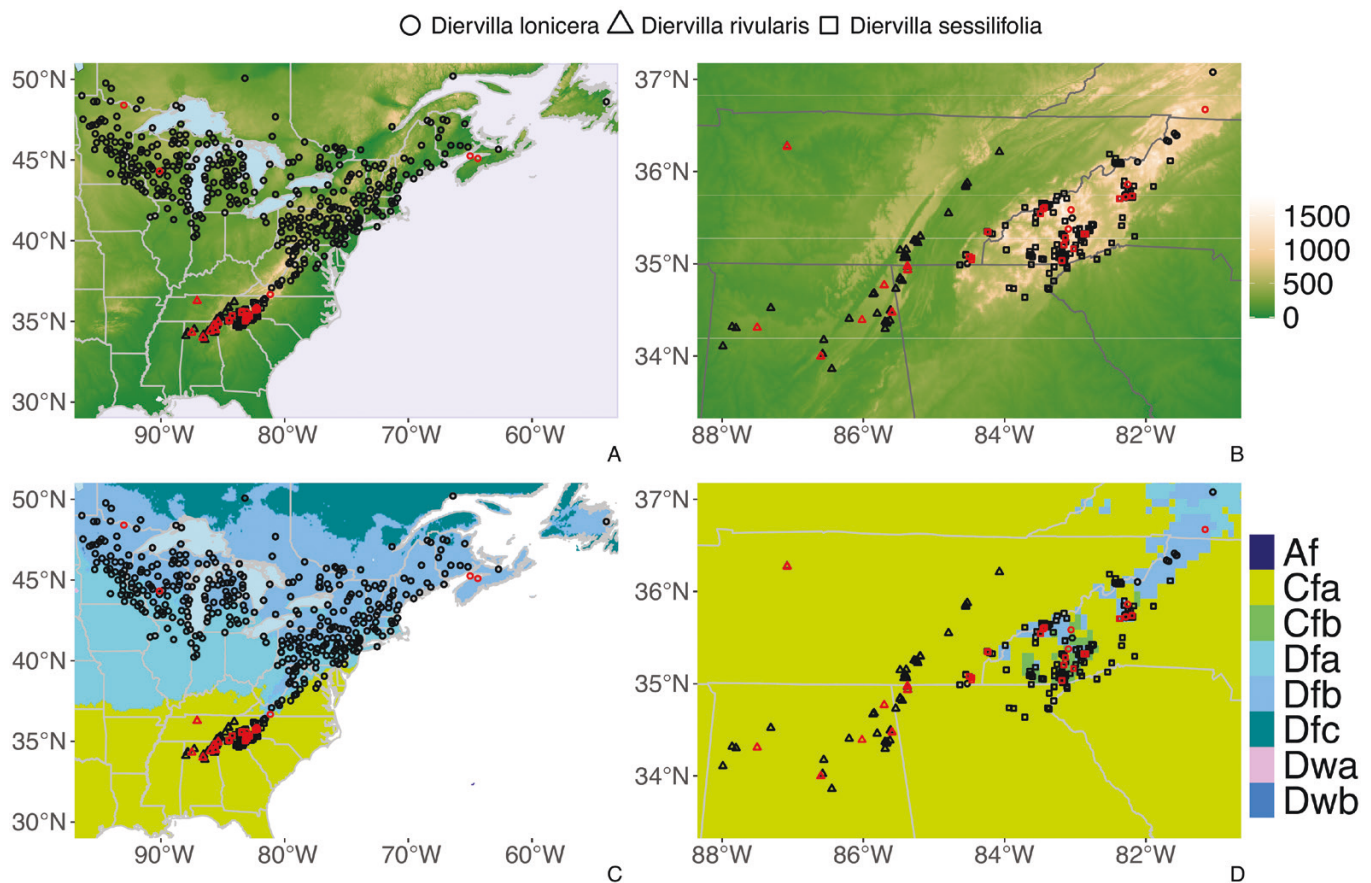


Figure 6. Partial segregation of *Diervilla* species by climatic niches as shown on occurrence maps [based on herbarium specimens in the SERNEC database (markers with black outlines) and vouchers for samples collected in the field and used in the analyses in this study (markers with red outlines)]. Species identifications were individually verified by K.G.M. Circles represent *D. lonicera*, triangles represent *D. rivularis*, and squares represent *D. sessilifolia*. Maps A and C on the left show the entire species distributions in eastern North America. Maps B and D on the right show details of all three species distributions in the southeastern portion of their ranges. A–B, map colours indicate elevation in metres. C–D, map colours indicate zones of the Köppen–Geiger world climate–vegetation classification system based on patterns of seasonal temperature and precipitation.

Finding a great age of their divergences is consistent with that of other Southern Appalachian species with close northern relatives. For example, [Quinlan et al. \(2020\)](#) dated the divergence between the northern circumboreal *Kalmia procumbens* and its sister species, the southeastern *K. buxifolia*, to ~9.4 Mya, also in the late Miocene. Climatic oscillations between warm/humid and cool/dry climates occurring throughout the Miocene are thought to be responsible for vicariant lineage splitting events, through range expansion and contraction, between North American and Eastern Asian disjunct clades such as *Weigela* and *Diervilla*, as well as divergences of clades of many Southern Appalachian woody species ([Wen et al. 2010](#), [Manos and Meireles 2015](#)).

The wide ecological amplitude encompassed in the *D. sessilifolia* distribution, where it ranges from 1200 to 5400 ft elevation and across three Köppen–Geiger climatic zones, could help explain the morphological overlap with *D. rivularis* despite the estimated old age of their divergence (~6 Mya). However, historical introgression could also have resulted in intermediate morphotypes. We discovered several, nonadjacent populations from herbarium specimens and plant collections made after

DNA data were collected, which are therefore not represented in this study, that show mixtures of ‘sessilifolia’ and ‘rivularis’ hair types. Notably, these are in between the ranges of the two taxa (Valley and Ridge province along the Clinch River and Walden Ridge escarpment in east TN) or at the northwestern edge of the *D. sessilifolia* range (Unaka Mountains along the Nolichucky River in TN and NC). These may represent populations of hybrid origin from glacial periods when populations of all three taxa are hypothesized to have shifted and may have come into contact.

In addition, historical hybridization may have occurred between *D. sessilifolia* and *D. lonicera*: plants from Polk Co., TN (Unaka Mountains), along the Ocoee River, have ‘sessilifolia’-type hairs and are firmly in the *D. sessilifolia* clade in our phylogenetic analysis, but they have atypically long petioles (3 mm) for this species and have been misidentified as *D. lonicera* based on this trait. The southernmost population of *D. lonicera* is nearby, on Big Frog Mountain, but at a much higher elevation (~3700 ft vs. ~1000 ft). Meanwhile, at the only known location where *D. sessilifolia* and *D. lonicera* are sympatric, in Yancey Co., NC, at 5200 ft elevation, these two species are clearly morphologically differentiated.

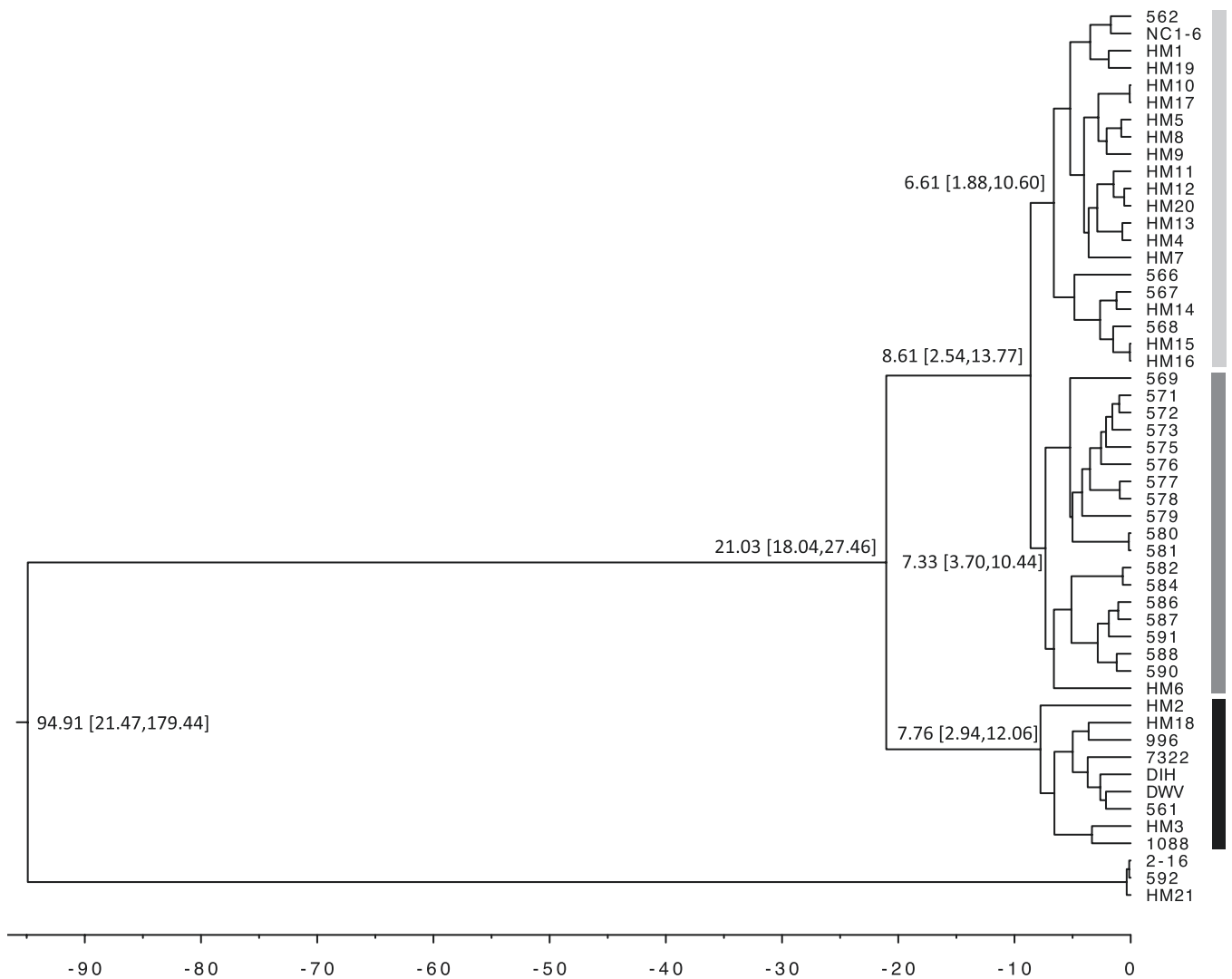


Figure 7. Fossil-calibrated chronogram for *Diervilla*. Estimated node ages (Mya) are shown for each major clade with 95% confidence intervals in square brackets. The light grey bar represents *D. sessilifolia* samples, the dark grey bar represents *D. rivularis* samples, and the black bar represents *D. lonicera* samples.

CONCLUSIONS

Multiple lines of evidence support the recognition of three species of *Diervilla* that diverged long enough ago to show distinctive genotypic ancestral groups, despite possible instances of introgression. Critical assessment of morphology in *Diervilla* revealed a consistent morphological character, twig hair type, as well as suites of quantitative vegetative characters that can reliably be used to distinguish

species, resulting in a more effective identification key given below, which includes accurate geographical ranges. Our phylogenetic analysis clearly shows for the first time that the two Southeastern USA taxa, *D. sessilifolia* and *D. rivularis*, represent sister species that are derived relative to the predominantly northeastern species, *D. lonicera*. This aligns with the fossil-evidenced arcto-Tertiary origins of *Diervilla* and its sister group, *Weigela*.

Synoptic key to the species of *Diervilla*

1. Inflorescences few-flowered (axillary: 1–3 flowers; terminal: 2–13 flowers); leaf blades broadly ovate with short-lanceolate apices, 5–9 cm long; petioles 5–10 mm long (median = 6 mm); twigs below peduncles puberulent with very short, soft hairs that are curled toward the stem; persistent sepal lobes 3–5 mm long (median = 3.5 mm) [high elevations in the Southern Appalachians, north to Great Lakes, northeastern USA and Canada].....*Diervilla lonicera*
1. Inflorescences many-flowered (axillary: 18+ flowers; terminal: 30–100 flowers); leaf blades narrowly ovate with long-lanceolate apices, 8–21 cm long; petioles 0–8 mm long (median = 2 mm); twigs below peduncles glabrous or pubescent with long hairs; persistent sepal lobes 1–3 mm long [Southeastern USA, generally at elevations below 4000 ft].....2

2. Twigs below peduncles moderately to densely hispid (rarely glabrous), with stiff, straight hairs at right-angles to the stem; leaf blades 8–21 cm long; petioles 1–7 mm (median = 4 mm); persistent sepal lobes 1–2 mm long (median = 2 mm) [of the Southern Cumberland Plateau of northern AL, northern GA, eastern TN, Western Highland Rim of TN, and northwest Coastal Plain of AL].....*Diervilla rivularis*
2. Twigs below peduncles glabrous or pubescent with soft hairs that are curled antrorsely toward the stem; leaf blades 8–14 cm long; petioles 0–3 mm (median = 0 mm); persistent sepal lobes 1.5–3 mm long (median = 3 mm) [of the Blue Ridge Mountains of western NC and eastern TN].....*Diervilla sessilifolia*

SUPPLEMENTARY DATA

Supplementary data are available at *Botanical Journal of the Linnean Society* online.

ACKNOWLEDGMENTS

The following helped K.G.M. with fieldwork and obtaining samples: Alain Belliveau, Cooper Breeden, Dwayne Estes, Sheryl Gould, Zack Irick, and Brian Keener. We thank the curators of the following herbaria for supplying specimen loans: CLEMS, JSU, NCU, NCSC, TENN, and UCHT. Initial *Diervilla* samples were collected by Hannah Meeler for use in her Master's thesis research at WCU. The authors would like to thank Dr Allan Strand for the use of his laboratory, sharing barcode adaptors, providing assistance with generating ddRAD libraries, and advice on sequencing. The WCU Scholarly Development Assignment Program allowed K.G.M. to collect morphological data and begin writing the manuscript for this project. Housing, research, and herb-arium space at Highlands Biological Station was kindly provided to K.G.M. during scholarly leave.

CREDIT STATEMENT

Katherine G. Mathews (conceptual development of this project, performed field work, morphological and molecular data collection, wrote the taxonomic key, wrote, edited, and revised the manuscript), Brandon Wheeler (conducted bioinformatics analyses, provided the mapped STRUCTURE figure, advised K.G.M. on analyses, software use, python and R coding, edited and revised the manuscript), and Luiz Silveira (conceptual development of this project, conducted multivariate and mapping analyses, wrote, edited, and revised the manuscript).

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

FUNDING

The Western Carolina University Faculty Research and Creative Activities Award to K.G.M. funded the ddRADSeq library development and sequencing.

DATA AVAILABILITY

The python code used in ipyrad and the data underlying this article are available in Zenodo (zenodo.org), and can be accessed with DOI 10.5281/zenodo.10685121 and DOI 10.5281/zenodo.11992334.

APPENDIX

Voucher specimens collected in the field and used in this study. Molecular and morphological data were taken from each

voucher. Entries are formatted as follows: Collector name and number, US state/Canadian province, county/city, latitude, longitude, elevation, herbarium digital accession number (em dash = missing data). All vouchers are deposited at WCUH except where indicated.

Diervilla lonicera (N = 9)

Mathews 561, NC, Jackson, 35.37219, -83.081932, 5284 ft, WCUH0033868; McNeilus 996, WI, Wood, 44.31004, -90.117728, 1017 ft; McNeilus 1088, MN, Koochiching, 48.39427, -93.008748, 1210 ft; Wieboldt 7322, VA, Grayson, 36.672, -81.1622, 4300 ft, NCU00145873; Belliveau (ACAD), NS, Isle Haute, 45.25324, -64.993683, 328 ft; Belliveau (ACAD), NS, Wolfville, 45.08722, -64.367783, 656 ft; Meeler 18, NC, Yancey/Buncombe, 35.7242, -82.285383, 4580 ft, —; Meeler 2, NC, Yancey, 35.86017, -2.2591, 4885 ft, —; Meeler 3, NC, Haywood, 35.58428, -83.060517, 5010 ft, —.

Diervilla rivularis (N = 19)

Mathews 569, GA, Dade, 34.97551, -85.376106, 1200 ft; Mathews 571, AL, DeKalb, 34.47366, -85.593789, 1240 ft; Mathews 572, AL, DeKalb, 34.47366, -85.593789, 1240 ft; Mathews 573, AL, DeKalb, 34.47357, -85.593739, 1241 ft; Mathews 575, AL, DeKalb, 34.47357, -85.593739, 1241 ft; Mathews 576, AL, DeKalb, 34.47357, -85.593739, 1240 ft; Mathews 577, AL, Jackson, 34.7693, -85.704978, 1316 ft, WCUH0031830; Mathews 578, AL, Jackson, 34.7693, -85.704978, 1317 ft, WCUH0031829; Mathews 579, AL, DeKalb, 34.39002, -86.019375, 987 ft, WCUH0031828; Mathews 580, AL, Blount, 33.99571, -86.603321, 440 ft; Mathews 581, AL, Blount, 33.99571, -86.603321, 440 ft; Mathews 582, AL, Lawrence, 34.30831, -87.503575, 855 ft, WCUH0031827; Mathews 584, AL, Lawrence, 34.30831, -87.503575, 855 ft; Mathews 586, TN, Cheatham, 36.27065, -87.077402, 407 ft; Mathews 587, TN, Cheatham, 36.27065, -87.077402, 407 ft; Mathews 588, TN, Cheatham, 36.27065, -87.077402, 407 ft; Mathews 590, TN, Cheatham, 36.26964, -87.080481, 520 ft; Mathews 591, TN, Cheatham, 36.26964, -87.080481, 520 ft; Meeler 6, GA, Walker, 34.93125, -85.373817, 1276 ft.

Diervilla sessilifolia (N = 21)

Mathews 562, NC, Jackson, 35.20371, -83.160145, 3040 ft, —; Mathews 566, TN, Polk, 35.06676, -84.46146, 1233 ft, WCUH0031822; Mathews 567, TN, Polk, 35.0779, -84.506367, 1040 ft, WCUH0031823; Mathews 568, TN, Polk, 35.0779, -84.506367, 1040 ft, WCUH0031824; Meeler 10, NC, Jackson, 35.163067, -83.027367, 1125 m, —; Meeler 11, TN, Monroe, 35.3461, -84.2426, 1175 ft, —; Meeler 12, TN, Polk, 35.040556, -84.466667, 959 m, —; Meeler 13, NC, Swain, 35.59583, -83.458611, 6644 ft, —; Meeler 14, TN, Polk,

35.04056, -84.46667, 1440 ft, —; Meeler 15, NC, Buncombe, 35.70427, -82.373833, 5650 ft, —; Meeler 16, NC, Buncombe, 35.70427, -82.373833, 5650 ft, —; Meeler 17, TN, Monroe, 35.3461, -84.2426, 358 m, —; Meeler 19, NC, Jackson, —, —, —; Meeler 20, TN, Monroe, —, —, —; Meeler 4, NC, Swain, 35.60917, -83.434167, 6644 ft, —; Meeler 5, TN, Monroe, 35.34492, -84.235517, 1040 ft, —; Meeler 7, NC, Haywood, 35.584283, -83.060517, 1527 m, —; Meeler 8, NC, Haywood, 35.323395, -82.877083, 1856 m, —; Meeler 9, NC, Haywood, 35.321967, -82.850467, 1535 m, —; Horton 2309 (NC1-6), NC, Macon, 35.03405, -83.191537, 4050 ft, WCUH0010946; Meeler 1, NC, Jackson, 35.2821, -83.1382, 2200 ft, —.

Weigela florida ($N = 3$)

Mathews 592, NC, Jackson, 35.37264, -83.22748, 2100 ft, —; Mathews 2-16, NC, Jackson, 35.37264, -83.22748, 2100 ft, —; Meeler 21, NC, Jackson, 35.37264, -83.22748, 2100 ft, —.

REFERENCES

- Ahles HE. *Diervilla sessilifolia* var. *rivularis*. *Journal of the Elisha Mitchell Scientific Society* 1964;**80**:173.
- Backlund AK, Pyck N. Diervillaceae and Linnaeaceae, two new families of Caprifolioids. *Taxon* 1998;**47**:657–61. <https://doi.org/10.2307/1223583>
- Beck M. *ggord: Ordination Plots with ggplot2*. R package version 1.1.8. 2024.
- Beck HE, Zimmermann NE, McVicar TR *et al.* Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data* 2018;**5**:1–12.
- Bell CD, Donoghue MJ. Dating the Dipsacales: comparing models, genes, and evolutionary implications. *American Journal of Botany* 2005;**92**:284–96. <https://doi.org/10.3732/ajb.92.2.284>
- Bouckaert R, Vaughan TG, Barido-Sottani J *et al.* BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* 2019;**15**:e1006650. <https://doi.org/10.1371/journal.pcbi.1006650>
- Buckley SB. *Diervilla sessilifolia*. *The American Journal of Science and Arts* 1843;**45**:174.
- Chester EW, Wofford EB, Shaw J, *et al.* *Guide to the Vascular Plants of Tennessee*. Knoxville: University of Tennessee Press, 2015.
- Chorlton LB. *Generalized geology of the world: bedrock domains and major faults in GIS format: a small-scale world geology map with an extended geological attribute database*. Geological Survey of Canada, 2007:48. <https://doi.org/10.4095/223767>
- Dahal S, Siniscalchi CM, Folk RA. Species delimitation in the *Symphyotrichum subulatum* group (Asteraceae) reveals a new species in Central Mexico. *Systematic Botany* 2023;**48**:595–607. <https://doi.org/10.1600/036364423x17000842213579>
- Dayrat B. Towards integrative taxonomy. *Biological Journal of the Linnean Society* 2005;**85**:407–15. <https://doi.org/10.1111/j.1095-8312.2005.00503.x>
- DeQueiroz K. Species concepts and species delimitation. *Systematic Biology* 2007;**56**:879–86. <https://doi.org/10.1080/10635150701701083>
- Donoghue MJ, Bell CD, Li J. Phylogenetic patterns in Northern Hemisphere plant geography. *International Journal of Plant Sciences* 2001;**162**:S41–52. <https://doi.org/10.1086/323278>
- Eaton DAR, Overcast I. Interactive assembly and analysis of RADseq datasets. *Bioinformatics* 2020;**36**:2592–4. <https://doi.org/10.1093/bioinformatics/btz966>
- Estill JC, Cruzan MB. Phytogeography of rare plant species endemic to the southeastern United States. *Castanea* 2001;**66**:3–23.
- Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 2005;**14**:2611–20. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Ezard T, Fujisawa T, Barraclough T. *splits: SPecies' LImits by Threshold Statistics*. R package version 1.0-20/r56, 2021. <https://R-Forge.R-project.org/projects/splits/> (4 December 2023, date last accessed).
- Francis RM. POPHELPER: an R package and web app to analyse and visualize population structure. *Molecular Ecology Resources* 2017;**17**:27–32. <https://doi.org/10.1111/1755-0998.12509>
- Freudenstein JV, Broe MB, Folk RA *et al.* Biodiversity and the species concept – Lineages are not enough. *Systematic Biology* 2017;**66**:644–56. <https://doi.org/10.1093/sysbio/syw098>
- Gattinger A. *Diervilla rivularis*. *Botanical Gazette* 1888;**13**:191–191. <https://doi.org/10.1086/326302>
- Gould KR, Donoghue MJ. Phylogeny and biogeography of *Triosteum*. *Harvard Papers in Botany* 2000;**5**:157–66.
- Hardin JW. *Diervilla* (Caprifoliaceae) of the southeastern U. S. *Castanea* 1968;**33**:31–6.
- Hijmans R. *raster: Geographic Data Analysis and Modeling*. R package version 3.6-20, 2023. <https://CRAN.R-project.org/package=raster> (15 November 2023, date last accessed).
- Kapli P, Lutteropp S, Zhang J *et al.* Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics* 2017;**33**:1630–8. <https://doi.org/10.1093/bioinformatics/btx025>
- Kim Y-D, Kim S-H. Phylogeny of *Weigela* and *Diervilla* (Caprifoliaceae) based on nuclear rDNA ITS sequences: Biogeographic and taxonomic implications. *Journal of Plant Research* 1999;**112**:331–41. <https://doi.org/10.1007/pl00013887>
- Linnaeus C. *Lonicera diervilla*. *Species Plantarum* 1753;**1**:175.
- Luo A, Ling C, Ho SYW *et al.* Comparison of methods for molecular species delimitation across a range of speciation scenarios. *Systematic Biology* 2018;**67**:830–46. <https://doi.org/10.1093/sysbio/syy011>
- Manos PS, Meireles JE. Biogeographic analysis of the woody plants of the southern Appalachians: Implications for the origins of a regional flora. *American Journal of Botany* 2015;**102**:780–804. <https://doi.org/10.3732/ajb.1400530>
- Matthews JV Jr, Ovenden LE. Late Tertiary plant macrofossils from localities in arctic/subarctic North America: a review of the data. *Arctic* 1990;**43**:364–92.
- Meeler HE. Taxonomy and phylogeny of the flowering plant genus, *Diervilla* (Diervillaceae). MS Thesis, Western Carolina University, 2018.
- Miller P. *Diervilla Tourn. ex Mill., Gardeners Dictionary*, 4th edn. London: Printed for the author and sold by John and James Rivington, 1754. [unpagged].
- Miller P. *Diervilla lonicera*. *Gardeners Dictionary*, 8th edn. London: Printed for the author and sold by John and Francis Rivington [and 23 others], 1768 [unpagged].
- Oksanen J, Simpson G, Blanchet FG, *et al.* *vegan Community Ecology Package Version 2.6-2*. The Comprehensive R Archive Network, 2022. <http://cran.r-project.org>
- Pantinople DJ, Engle-Wrye NJ, Folk RA. *Heuchera tuckasegeensis* sp. nov. (Saxifragaceae), a new species from Western North Carolina. *Systematic Botany* 2024;**49**:37–47. <https://doi.org/10.1600/036364424x17110456120677>
- Parchman TL, Gompert Z, Mudge J *et al.* Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology* 2012;**21**:2991–3005. <https://doi.org/10.1111/j.1365-294X.2012.05513.x>
- Pons J, Barraclough TG, Gomez-Zurita J *et al.* Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology* 2006;**55**:595–609. <https://doi.org/10.1080/10635150600852011>
- Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics* 2000;**155**:945–59. <https://doi.org/10.1093/genetics/155.2.945>
- Puillandre N, Lambert A, Brouillet S *et al.* ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology* 2012;**21**:1864–77. <https://doi.org/10.1111/j.1365-294X.2011.05239.x>
- Quinlan EJ, Mathews KG, Collins B *et al.* Phylogenetic divergence and ecophysiological variation in the disjunct *Kalmia buxifolia*

- (Sand-myrtle, Ericaceae). *Systematic Botany* 2020;**45**:900–12. <https://doi.org/10.1600/036364420x16033962925277>
- Radford AE, Ahles HE, Bell CR. *Manual of the Vascular Flora of the Carolinas*. Chapel Hill: The University of North Carolina Press, 1968.
- Ramseur GS. The vascular flora of the high mountain communities of the Southern Appalachians. *Journal of the Elisha Mitchell Scientific Society* 1960;**76**:82–112.
- Soltis DE, Morris AB, McLachlan JS *et al.* Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology* 2006;**15**:4261–93. <https://doi.org/10.1111/j.1365-294X.2006.03061.x>
- Spriggs EL, Eaton DAR, Sweeney PW *et al.* Restriction-site-associated DNA sequencing reveals a cryptic *Viburnum* species on the North American Coastal Plain. *Systematic Biology* 2019;**68**:187–203. <https://doi.org/10.1093/sysbio/syy084>
- Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2018;**30**:1312–3.
- Stevens PF. *Angiosperm Phylogeny Website*. Version 14, July 2017 [and more or less continuously updated since]. 2001 onwards. <http://www.mobot.org/MOBOT/research/APweb/> (5 June 2024, date last accessed).
- Stubbs RL, Folk RA, Soltis DE *et al.* Diversification in the arctic: biogeography and systematics of the North American *Micranthes* (Saxifragaceae). *Systematic Botany* 2020;**45**:802–11. <https://doi.org/10.1600/036364420x16033963649282>
- Swofford DL. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*, v. 4.0a, beta version, Windows GUI. Sunderland: Sinauer Associates, 2021.
- Tennessee Flora Committee. *Guide to the Vascular Plants of Tennessee*. Knoxville: The University of Tennessee Press, 2015.
- Venables WN, Ripley BD. *Modern Applied Statistics with S*, 4th edn. New York: Springer, 2002.
- Warton DI, Wright TW, Wang Y. Distance-based multivariate analyses confound location and dispersion effects. *Methods in Ecology and Evolution* 2012;**3**:89–101.
- Weakley AS, Southeastern Flora Team. Flora of the Southeastern United States Web App. Chapel Hill: University of North Carolina Herbarium, North Carolina Botanical Garden, 2023. <https://fsus.ncbg.unc.edu/main.php?pg=index.php> (11 December 2023, date last accessed).
- Wen J. Evolution of Eastern Asian and Eastern North American disjunct distributions in flowering plants. *Annual Review of Ecology and Systematics* 1999;**30**:421–55.
- Wen J, Ickert-Bond S, Nie Z, *et al.* Timing and modes of evolution of Eastern Asian – North American biogeographic disjunctions in seed plants. In: Long M, Gu H, Zhou Z (eds). *Darwin's Heritage Today: Proceedings of the Darwin 200 Beijing International Conference*. Beijing: Higher Education Press, 2010, 252–69.
- Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer, 2016.
- Will KW, Mishler BD, Wheeler QD. The perils of DNA barcoding and the need for integrative taxonomy. *Systematic Biology* 2005;**54**:844–51. <https://doi.org/10.1080/10635150500354878>
- Wiser SK. High-elevation cliffs and outcrops of the southern Appalachians: vascular plants and biogeography. *Castanea* 1994;**59**:85–116.
- Yeates DK, Seago A, Nelson L *et al.* Integrative taxonomy, or iterative taxonomy? *Systematic Entomology* 2011;**36**:209–17. <https://doi.org/10.1111/j.1365-3113.2010.00558.x>