

GEOGRAPHICAL HISTORY AND INFRASPECIFIC MORPHOLOGICAL
VARIATION OF AMERICAN COW-WHEAT (MELAMPYRUM LINEARE;
OROBANCHACEAE)

by

Karoline Oldham
A Thesis
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Master of Science
Biology

Committee:

_____	Dr. Andrea Weeks, Thesis Director
_____	Dr. James Lawrey, Committee Member
_____	Dr. Don Seto, Committee Member
_____	Dr. James D. Willett, Director, School of Systems Biology
_____	Dr. Donna M. Fox, Associate Dean, Office of Student Affairs & Special Programs, College of Science
_____	Dr. Peggy Agouris, Dean, College of Science
Date: _____	Spring Semester 2015 George Mason University Fairfax, VA

Geographical History and Intraspecific Morphological Variation of American Cow-
Wheat (*Melampyrum lineare*; Orobanchaceae)

A Thesis submitted in partial fulfillment of the requirements for the degree of Master of
Science at George Mason University

by

Karoline Oldham
Bachelor of Arts
Virginia Tech, 2011

Director: Andrea Weeks, Associate Professor
Department of Biology

Spring Semester 2015
George Mason University
Fairfax, VA



This work is licensed under a [creative commons attribution-noncommercial 3.0 unported license](https://creativecommons.org/licenses/by-nc/3.0/).

DEDICATION

This is dedicated to my ever-growing family, Sarah Cole Smith, and my cat, Loco.

ACKNOWLEDGEMENTS

I would like to thank the many friends and relatives who have helped me throughout this process. My adviser, Dr. Andrea Weeks, has been an amazing and reliable guide as I have navigated the twists and turns of my first big research project, and my lab mate, Morgan Gostel, has provided invaluable help and comedic relief along the way. Dr. Arndt Laemmerzahl provided crucial statistical advice, and Drs. Lawrey and Seto from my committee have given valuable feedback on my writing. Funding for field and lab work was provided by the New England Botanical Club, the Southern Appalachian Botanical Society, and the Cosmos Club Foundation.

TABLE OF CONTENTS

	Page
Abstract	vii
Introduction	1
The Varietal Delimitation of <i>Melampyrum lineare</i>	2
The Historical Biogeography of <i>Melampyrum lineare</i>	3
Objectives	6
Methods	7
The Varietal Delimitation of <i>Melampyrum lineare</i>	7
Loans and Collections	7
Determination of Specimens to Varietal Names	7
Measurements	9
Statistical Analyses	13
The Historical Biogeography of <i>Melampyrum lineare</i>	19
Taxon Sampling and DNA Extraction	19
Marker Selection and Molecular Protocol	19
Data Analysis	20
Phylogenetic Analysis	20
Results	22
The Varietal Delimitation of <i>Melampyrum lineare</i>	22
Character Selection	22
Multivariate Analysis of Variance and Discriminant Analysis	25
Chi-Squared Contingency Tests	27
The Historical Biogeography of <i>Melampyrum lineare</i>	28
Discussion	30
The Varietal Delimitation of <i>Melampyrum lineare</i>	30
Future Research	33
The Historical Biogeography of <i>Melampyrum lineare</i>	34
Future Research	34
Appendix A: Tables	36
Appendix B: Figures	50

Appendix C: Exsiccatae.....	82
References.....	90

ABSTRACT

GEOGRAPHICAL HISTORY AND INFRASPECIFIC MORPHOLOGICAL VARIATION OF AMERICAN COW-WHEAT (MELAMPYRUM LINEARE; OROBANCHACEAE)

Karoline Oldham, M.S.

George Mason University, 2015

Thesis Director: Dr. Andrea Weeks

The wildflower American cow-wheat (*Melampyrum lineare* Desr.) is the sole North American species of an otherwise Eurasian genus. There are four purported varieties of *M. lineare*, though it has not been thoroughly determined whether they merit designation as distinct taxonomic entities. Further, the evolutionary history of *M. lineare* has been poorly studied in light of this species' geographic isolation from the rest of its genus. The goal of this study was to reevaluate and revise the infraspecific taxonomy of *Melampyrum lineare* and to test hypotheses regarding the evolutionary history of this species using comparative genetic methods. The four purported varieties of *M. lineare* do not represent truly distinct taxonomic entities based on morphological data. Instead, morphological variation occurs along a continuum with no clear divisions between the putative varieties. Results from comparative genetic methods provide mixed support for

the competing hypotheses of European versus Asian ancestry for the sole North American species in *Melampyrum*.

INTRODUCTION

The wildflower American cow-wheat (*Melampyrum lineare* Desr.; Figure 1) is the sole North American species in the genus *Melampyrum*, which includes an additional 34 species, all of which are found in Eurasia (Pennell, 1935; Weakley, 2012).

Melampyrum and several other genera were recently moved from Scrophulariaceae to the broom-rape family, Orobanchaceae by Olmstead et al. (2001). Like most members of the broom-rape family, *M. lineare* is hemiparasitic - it is a facultative root parasite of various trees, such as *Acer saccharum*, *Pinus strobus*, and others (Cantlon et al., 1963).

Consequently, the distribution of *M. lineare* (Figure 1) is generally limited to temperate mid- to high-latitude or high-elevation forests (Pennell, 1935; Weakley, 2012) of northern and eastern North America that contain its preferred host species. The distribution of *Melampyrum lineare* may be further constrained by the behavior of granivorous ants that, as for many forest wildflowers, act as seed-dispersal vectors (Gibson, 1993a, 1993b). Specifically, *M. lineare* grows along the east coast of North America spanning Georgia to Newfoundland and Labrador and spreads west along the US – Canada border to Idaho and British Columbia. While there appear to be disjunct populations in eastern and western United States separated by unpopulated regions in North and South Dakota, *M. lineare* can be found from coast to coast in Canada, so eastern and western North American populations are not truly disjunct.

The Varietal Delimitation of *Melampyrum lineare*

There are four published varieties in the species. The precise morphological distinctions and geographical distributions of the varieties of *Melampyrum lineare* have not been revised since the work of Pennell (1935), nor is it clear whether all varieties warrant recognition as separate taxonomic entities (Gleason & Cronquist, 1991; Schori, 2001; Weakley, 2012). Pennell (1935) recognized three varieties based primarily on bracteal leaf width and tothing: *M. lineare* var. *latifolium* Barton, found between Massachusetts and New York, Georgia and Minnesota; *M. lineare* var. *pectinatum* (Pennell) Fernald, which mainly follows the Coastal Plain between Massachusetts and Virginia (Weakley, 2012); and *M. lineare* var. *typicum* Pennell, correctly referred to as *M. lineare* var. *lineare* and found between Newfoundland, New York and British Columbia. A fourth variety, *M. lineare* var. *americanum* (Michaux) Beauverd, has a similar distribution to *M. lineare* var. *latifolium* but reaches Quebec in the North and extends only as far south as Tennessee and North Carolina. However, Pennell (1935) expressed doubts about the distinctiveness of the varieties he described; “I wonder if they can ever be satisfactorily analysed [...] as the transitions between them are at times imperceptible and some specimens of each vary toward the condition of either of the other subspecies.”

More recently, Gleason & Cronquist (1991) and Weakley (2012) have cautioned that the distinction between *M. lineare* var. *lineare* and *M. lineare* var. *americanum* is in need of further investigation, as these varieties only differ slightly by the presence or absence of bracteal teeth. Schori (2001) concluded that within species variation occurs along a continuum, and the varietal designations should therefore be completely removed

from use. However, Schori (2001) examined 141 *M. lineare* specimens from only two New England herbaria. These specimens were primarily collected in northern New England and did not thoroughly represent populations and varieties from elsewhere in the distribution of the species. More extensive sampling is therefore necessary to thoroughly investigate the delimitation of infraspecific, or within species, taxonomic groups in this species.

The Historical Biogeography of *Melampyrum lineare*

Very little is known about how and when the current distribution of *Melampyrum lineare* became established in North America or how it is related to the Eurasian members of its genus. Based on the degree of capsule dehiscence and pollination mechanisms, Pennell (1935) proposed that *M. lineare* may be most closely related to a European species, *M. pratense* L. (Figure 2), which would suggest European ancestry. Pennell (1935) also mentions an alternative hypothesis: *M. lineare* may be most closely related to the Asian species *M. arcuatum*, now *M. laxum* var. *arcuatum* (Nakai) Soo (Figure 2). These two hypotheses imply drastically different ancestral histories. Pennell's hypotheses, however, were largely based on a subjective, pre-cladistic evaluation of the data available to him. Although he describes many qualitative differences between *M. lineare* and other members of the genus, he does not present actual specimen measurements, and he uses differences in pollination mechanisms as the primary evidence to support his evolutionary claims without supporting his designation of this character as more important than any other evolutionarily. Pennell notes that species in most of the genus rely on insect pollination, but *M. pratense* favors but does not

exclusively rely on self-pollination. In addition to this observation, Pennell speculates that *M. lineare* relies exclusively on self-pollination based on both personal observations and floral characters such as size, color, scent, presence or absence of nectar, and, primarily, the position of the stigma relative to the flower opening. Because Pennell failed to observe the stigma of *M. lineare* flowers ever protruding into the opening of the corolla, he reasoned that it relies on self-pollination and is most closely related to *M. pratense*, which frequently but not exclusively self-pollinates.

However, Pennell does not support his designation of pollination mechanism as the most informative character in determining evolutionary relationships. A more objective method for evaluating the relationship between *M. lineare* and the rest of the genus is therefore necessary. Fossil evidence for members of Orobanchaceae is lacking, and morphological comparison is made difficult by poor specimen preservation and representation in herbarium records (Wolfe et al., 2005). One study examined pollen morphology for select members of Orobanchaceae (Minkin & Eshbaugh, 1989), but only one species of *Melampyrum*, *M. pratense*, was included.

Molecular data has recently been used to develop family-level phylogenies with several representatives from *Melampyrum* (McNeal et al., 2013; Wolfe et al., 2005). DNA sequence data from several loci, including nuclear nrITS, PHYA, and PHYB, as well as plastid matK and rps2, have been collected and published for several European and Asian species in the genus *Melampyrum*, as well as representatives from several closely related genera (McNeal et al. 2013; Li et al., 2008), but *M. lineare* and several Asian members of the genus were not included among these data.

While the stem and crown diversification dates are not known for *Melampyrum*, a time-calibrated phylogeny constructed by Wolfe et al. (2005) estimates that the clade containing the genus *Melampyrum* diversified approximately 33.9 million years ago (Ma). However, Wolfe et al. (2005) calibrated their phylogeny using a divergence date for the order Lamiales, to which Orobanchaceae belongs, that was estimated by averaging Lamiales diversification dates from two separate publications (Magallon et al., 1999; Wikström et al., 2001). This method of calibration does not rely on fossil evidence for Orobanchaceae, so it only provides only a rough estimate of diversification dates within this group. These results suggest, albeit with only weak support, that members of this genus evolved too recently to have established a Northern Hemisphere distribution through continental vicariance in the Cretaceous Period (Wolfe et al. 2005). Rather, dispersal across Beringia or the North Atlantic land bridge during the Eocene is a more likely hypothesis (Tiffney & Manchester, 2001; Wolfe et al., 2005).

If *Melampyrum lineare* shares a most recent common ancestor with the European *M. pratense* or another European species, this would provide support for dispersal into eastern North America during the early Eocene, prior to the formation of the Atlantic Ocean barrier (Thorne, 1972; Tiffney & Manchester, 2001). However, if *M. lineare* shares a most recent common ancestor with an Asian species, this would suggest dispersal across Beringia into western North America (Thorne, 1972; Tiffney & Manchester, 2001). Long distance dispersal via animal or some other vector, though improbable, also cannot be discounted as a possibility.

Objectives

The goal of this study was to revise the infraspecific taxonomy of *Melampyrum lineare* and to test hypotheses regarding the evolutionary history and phylogeography of this species using comparative genetic methods. Specifically, morphological data collected from herbarium specimens was examined to assess the varietal delimitation of *M. lineare*. It was also the goal of this project to further resolve the phylogenetic relationships within the genus *Melampyrum* by incorporating DNA sequence data from *M. lineare* and other previously omitted species into already existing molecular datasets, and in doing so, infer the biogeographic history behind the sole North American member of this disjunct genus.

METHODS

The Varietal Delimitation of *Melampyrum lineare*

Loans and Collections

In addition to specimens included in George Mason University's herbarium, *Melampyrum lineare* specimens were requested from 26 herbaria across the range of the species (Table 1). Loans were received from 24 of these herbaria. In addition to herbarium specimens from GMUF and those loaned from other herbaria, over 440 field collections were made across the range of *M. lineare* in eastern United States during Summer 2013.

Only specimens with both fruit and flowers present on the same individual were considered in this study. The presence of flowers was especially important because the bracts, which subtend the flowers, are a crucial character in current botanical keys for the purported *M. lineare* varieties. Because the flowers of this species are small and fragile, only those specimens were assessed that had flowers that were not glued down and could therefore be dissected to measure floral characters. After selection for fruit and floral characteristics, specimens were sorted by variety and then by geographic region to ensure that individuals were selected across as much of each variety's range as possible from this pool of specimens. Based on these criteria, 248 specimens were selected for inclusion in this study.

Determination of Specimens to Varietal Names

After sorting loan specimens to find only those with both fruit and flowers present on single individuals, specimens were sorted by variety. Few herbarium specimens had

been annotated to include a varietal designation, and those that did were not always consistent with current botanical keys, so the variety of each specimen was determined without relying on label information.

Weakley (2012) provides the most recent description to include all four varieties, so it was used to determine specimen varieties in this study. It primarily distinguishes between the varieties based on bracteal leaf width and the length of bracteal teeth. Following Weakley's diagnosis, specimens with entire or nearly entire lower bracteal leaves, uppermost bracteal leaves with a few short basal teeth, and a maximum vegetative leaf width over 10 mm were considered to belong to var. *latifolium*. Specimens with narrower leaves than var. *latifolium* (ranging from 2 to 10 mm wide) and middle to upper bracteal teeth lengths that roughly matched the width of the undivided portion of the bracts were considered var. *pectinatum*. Specimens with leaf widths similar to var. *pectinatum* but with middle and upper bracteal teeth lengths that were shorter than the undivided portion of the bracts were considered var. *americanum*. Specimens with leaves narrower than 5 mm wide and entire bracts were considered var. *lineare*.

Specimens did not always fall neatly into these categories. For example, some specimens that had the broad leaves associated with var. *latifolium* also had uncommonly long bracteal teeth for this variety. In this case, the specimens were classified as var. *latifolium* based on the leaf width. Additionally, three specimens lacked teeth like var. *lineare* or var. *latifolium*, but had leaves wider than those of var. *lineare* (i.e., more than 5 mm) and narrower than those of var. *latifolium* (i.e., narrower than 10 mm). These specimens were classified as var. *americanum* based on the leaf width.

Another complicating factor was that the different varietal designations did not always appear to represent two distinct groups because of the amount of overlap in the morphology of purportedly different varieties. For example, leaf widths for var. *americanum* and var. *pectinatum* overlapped, so the primary difference between these varieties was the relative prominence of the bracteal teeth, represented as the ratio between bracteal teeth length and the width of the undivided portion of the bracts. However, this led to some specimens with short bracteal teeth and narrow bracts being classified as var. *pectinatum*, while specimens with bracteal teeth of the same length and comparatively wider bracts were classified as var. *americanum*. Conversely, specimens with long bracteal teeth and wide bracts were classified as var. *americanum*, while specimens with equally long bracteal teeth and comparatively narrower bracts were classified as var. *pectinatum*. Based on these criteria, 14 of the specimens included in this study were classified as var. *lineare*, 95 were classified as var. *americanum*, 64 were classified as var. *latifolium*, and 75 were classified as var. *pectinatum*.

Measurements

Specimens were measured for 44 different characters, 19 of which were vegetative, 17 of which were floral, and 8 of which were fruit or seed characters (Table 2). These 44 characters were pulled from a variety of sources, including diagnostic descriptions in floras (Gleason & Cronquist, 1991; Pennell, 1935; Weakley, 2012) and previous research on the treatment of the species varieties (Schori, 2001). Current botanical keys focus exclusively on vegetative characteristics to distinguish between the varieties, so the majority of characters measured in this study were vegetative. There was

little to no published information on whether floral, fruit or seed characters differed across the varieties, so measurements were collected for a list of these reproductive characters that was developed by examining descriptions of *Melampyrum lineare* in various floras.

Of the 19 vegetative characters, 5 were discrete: presence or absence of primary branching, presence or absence of secondary branching, presence or absence of tertiary branching, number of teeth per side of the bract used for the bract width measurement, and the maximum number of teeth per side of a bract. Four vegetative characters were ratios between two other vegetative characters: leaf width to leaf length, lowest bract width to lowest bract length, bract width to bract length, and longest bracteal tooth length to width of the undivided portion of the bract. The remaining 10 characters were continuous variables such as plant height or leaf length.

Of the 17 floral characters, 2 were discrete: presence or absence of pubescence on the calyx and the position of the stigma relative to the floral tube opening. Five floral characters were ratios between two other floral characters: length of the lower corolla lip to the total corolla length, length of the upper corolla lip to the total corolla length, length of the lower corolla lip to that of the upper corolla lip, length of the separated portion of the calyx to total calyx length, and total calyx length to total corolla length. The remaining 10 characters were continuous variables such as stamen length and pistil length.

Of the 8 fruit and seed characters, none were discrete. Three of the fruit and seed characters were ratios between other fruit and seed characters: capsule width to capsule

length, seed width to seed length, and elaiosome length to total seed length. The remaining 5 characters were continuous variables.

Discrete variables were measured according to their descriptions in Table 2. It was impossible to reliably tally the total number of primary, secondary and tertiary branches on each individual, so these characters were reduced to presence/absence of branching data and were recorded as a 0 if absent or a 1 if present. The number of bracteal teeth per side of a bract were counted on the same bract used for measuring the bract width and length, but the maximum number of bracteal teeth per side of a bract was also recorded to better capture the range in bracteal teeth number on a single specimen. The presence or absence of pubescence on the calyx was recorded as a 1 or 0, respectively. Pennell (1935) noted that except for small flowers from what he referred to as var. *typicum*, correctly referred to as var. *lineare*, the stigma always remained inside the corolla tube. Consequently, the position of the stigma was recorded as a binary character: 0 if the stigma was hidden inside the corolla tube and 1 if the stigma protruded out of the corolla tube and was visible before dissecting the flower.

When available, habitat moisture levels were recorded for each specimen following the description in Table 2. This was prompted by observing two specimens from the New York Botanical Garden (Figures 3 and 4), both collected from sites near the Biltmore in Asheville, North Carolina – one labeled “dry” habitat was easily identifiable as var. *americanum*, while the other, labeled “moist” habitat was easily identifiable as var. *latifolium* using current botanical keys.

All continuous variables were measured in millimeters using electronic calipers that were accurate to the nearest hundredth of a millimeter. Since the calipers could only measure a maximum of six inches, plant height often exceeded their limit and was instead measured with a ruler that was accurate to the nearest half-millimeter. All floral characters and other very small characters such as elaiosome length were measured with the calipers underneath a dissecting microscope to enhance visual accuracy during the measurement process.

The vegetative characters mentioned most often when distinguishing between varieties were bract width, the presence or absence of bracteal teeth, leaf width, the degree of branching, and the length of internodes on the midstem, so these characters were all included. Since bract morphology varies widely from lower to upper bracts, length and width measurements were collected for the lowest bract or, if the lowest bract was damaged or missing, from the lowest intact bract and were recorded as the lowest bract length and the lowest bract width. This was repeated for a bract halfway between the lowest bract measured and the uppermost bract and recorded as bract length and bract width. Bract and leaf widths were measured perpendicularly to the main vein at the widest portion of the leaf or bract. For toothed bracts, the bract width only included the undivided portion of the bract, not the length of bracteal teeth.

Bracteal tooth length varied greatly within individual specimens, so the maximum bracteal tooth length was recorded for each individual to better capture any variability between varieties. Bracteal tooth length was measured from the tip of the tooth to the base of the tooth where it joined the undivided portion of the bract. Because the ratio

between bracteal tooth length and the width of the undivided portion of the bract is important for distinguishing between varieties *americanum* and *pectinatum*, the width of the undivided portion of the bract used for the bracteal tooth length measurement was also measured.

Schori (2001) included corolla length as the only floral characteristic and capsule width and capsule length as the only fruit or seed characters in her analysis, but Pennell (1935) mentions a variety of other floral, fruit and seed characteristics when describing the species. While there is no implication from current literature that characters from these reproductive categories vary across varieties (Gleason & Cronquist, 1991; Weakley, 2012), several characters in these categories were measured to test whether this is indeed the case (Table 2).

Statistical Analyses

Characters were separated into three sets: vegetative, floral, and fruit and seed. Boxplots were made for each of the 37 continuous characters measured. Multivariate analysis of variance (MANOVA) was used to determine whether significant statistical differences existed between the purported varieties using each of these three character sets. Then, if significant differences were present, linear discriminant analysis (LDA) was used to determine whether morphological variation occurs along a continuum or if the varieties form distinct groups that can be distinguished from one another using specific characters.

Linear discriminant analysis is a common statistical method for determining which out of two or more independent variables are useful for discriminating between

natural groups. It is similar to analysis of variance (ANOVA) and MANOVA in its assumptions and theory, but because it also provides functions that can be used to predict to which group individuals belong (Hill & Lewicki, 2007), linear discriminant analysis serves as a common and useful tool for infraspecific taxonomic revisions based on morphological characters (Alexander et al., 2012; Chmielewski, 1997; Fritsch & Lucas, 2000; Kephart et al., 1999; Mwanyambo, 1996).

Linear discriminant analysis assumes multivariate normality and homogeneity of variances and covariances across groups. Measurements were separated into three categories: vegetative characters, floral characters, and fruit and seed characters. Multivariate normality requires that each variable be normally distributed individually, so each variable was tested for univariate normality using Q-Q plots. Those variables that were not normally distributed were log transformed to attain a normal distribution. However, individual normality for all variables does not guarantee multivariate normality; if variables are singly normally distributed but are not independent, they will not be jointly normal (Ash, 2013). Some dependence between characters is unavoidable since the features of an organism tend to scale proportionally to each other (Enquist, 2002). However, eliminating highly correlative relationships is an approximate means of minimizing the influence of dependence between characters. For this reason, pairwise correlations were checked using Pearson's correlation coefficient to determine if one or more variables were dependent. Variables were considered correlated if the correlation coefficient (r) was greater than 0.7, and the most correlated variables were removed from the dataset one by one until all correlative relationships with r greater than 0.7 had been

eliminated and the data had a multivariate normal distribution. Multivariate normality was assessed by creating a Q-Q plot that compared Mahalanobis distances calculated from the data to a chi-square distribution (Penny, 1996). Multivariate normality was checked for vegetative characters, floral characters, and fruit and seed characters separately.

Homogeneity of variances was tested using the Brown-Forsythe test. Both Levene's test and the Brown-Forsythe test are commonly used to test this assumption. However, authors (Brown & Forsythe, 1974; Glass & Hopkins, 1996; Hill & Lewicki, 2007) have questioned the power of Levene's test because it relies on the absolute deviations from means within groups, which can be highly skewed and thus violate the underlying assumption of normality. Brown and Forsythe (1974) proposed an alternative that uses absolute deviations from group medians, which proves to be a more robust test of equality of variances. However, Glass and Hopkins (1996) caution that both Levene's test and the Brown-Forsythe test suffer from the same "fatal flaw" of assuming homogeneity of variances, which, as Hill and Lewicki (2007) point out, gives reason to doubt the robustness of both tests when variances are heterogeneous, particularly when sample sizes are unequal.

The homogeneity of covariances assumption was tested using Box's M test. This test is very sensitive to deviations from a multivariate normal distribution, so Hill and Lewicki (2007) advise that its results "should not be taken too 'seriously.'" Relatively little is known about how violating the homogeneity of covariances assumption impacts the power of linear discriminant analysis. However, if the homogeneity of variances

assumption is met and sample sizes are large and roughly equal, the power of a linear discriminant analysis should not be highly impacted. To address this ambiguity, Hill and Lewicki (2007) propose running analyses both including and excluding one or two groups; if results vary notably based on the inclusion or exclusion of particular groups, this suggests that a violation of this assumption is impacting the power of the linear discriminant analysis to detect true differences between groups. Following this logic, linear discriminant analyses were run both including and excluding specimens belonging to the purported variety *lineare*. This group of specimens was chosen because its sample size is most different from those of the other varieties; only 14 specimens from var. *lineare* were sampled, versus 95 from var. *americanum*, 64 from var. *latifolium*, and 75 from var. *pectinatum*. This comparatively small sample sized posed the largest threat to the assumption of homogeneous covariances.

Authors (Hill & Lewicki, 2007; Lindman, 1974) note that MANOVA and linear discriminant analysis are fairly robust to violations of the assumptions of normality and homogeneity of the variances and covariances. For this reason, these statistical analyses were run both when these assumptions were met and when they were not.

Multivariate analysis of variance was used to determine whether statistically significant differences exist across the four purported *Melampyrum lineare* varieties before proceeding to use linear discriminant analysis, which seeks to evaluate how to differentiate between the varieties. To eliminate uninformative variables from consideration in the analyses, a forward stepwise analysis was run using Wilk's lambda criterion, which is the multivariate equivalent of the F statistic. This analysis builds a

model by starting with a single variable that best discriminates between groups and then adding other variables to the model one by one in a stepwise fashion. At each step, the variable which minimizes the Wilk's lambda of the model is incorporated as long as its p-value still shows statistically significant differences between groups (Hill & Lewicki, 2007; Weihs et al., 2005).

Linear discriminant analysis was then run using the subset of variables that stepwise analysis indicated as important for capturing the variation between groups. The ability of the LDA model to accurately predict the varietal classification of a specimen was then assessed. Because the LDA model is being used to predict the varietal classification of the same specimens used in its development, its ability to do this may be higher than if the LDA model were applied to an entirely new set of specimens. To address this issue, the analyses were run both with and without jackknifing. Jackknifing excludes measurements for a single specimen, the LDA model is built using fewer data, and then the model is adjusted as this process is repeated over many iterations. This causes a lower predictive ability of the LDA model than without jackknifing, but it provides a better estimate of how well the LDA model would be able to predict varietal classifications for specimens other than those used in the model's development (Miller, 1974). If predictive accuracy is similar for an LDA model developed with jackknifing and one developed without, this suggests robustness between the two models.

Six discrete characters did not have a normal distribution, so these variables could not be included in linear discriminant analyses. Instead, chi-square contingency tests were used to determine whether differences existed between the varieties for the

following five characters: degree of primary, secondary and tertiary branching, number of bracteal teeth per side of bract used in bract width and bract length measurements, and maximum number of bracteal teeth per side of bract.

When available, data on habitat moisture levels were recorded. Many labels contained ambiguous information about moisture levels, such as describing the collection site as shady or sunny, or describing the soil as rocky or sandy. Sandy soil is often well-drained, but this soil type is common along lakesides and other areas where water is readily available. Similarly, a shady collection site might have higher water availability than a sunny spot in the same vicinity, but the amount of sun or shade a spot receives can change over the course of a day as the sun's position changes, and the amount of moisture in a shady spot for one region does not necessarily have higher or lower water availability than a shady spot in a different region. For these reasons, habitat moisture levels were only recorded when labels clearly identified the collection site as “dry” or “moist,” and this information was recorded according to the description in Table 2. A chi-square contingency test was then used to determine whether habitat moisture levels and variety are dependent.

All statistical analyses were conducted with R statistical software version 3.1.2 using the basic package as well as the MASS, biotools, lawstat, klaR, and scatterplot3d packages (Gastwirth et al., 2013; Ligges & Machler, 2003; Rodrigo da Silva, 2014; R Core Team, 2014; Venables & Ripley, 2002; Weihs et al., 2005).

The Historical Biogeography of *Melampyrum lineare*

Taxon Sampling and DNA Extraction

Whole genomic DNA was extracted from nine herbarium specimens of *Melampyrum lineare* with at least two representatives from each of the four purported varieties, two herbarium specimens of Asian *M. laxum* Miq., one herbarium accession of European *M. nemorosum* L., and one herbarium specimen of European *M. cristatum* L. (see Table 3 for accession information) using the FastDNA® Kit (MP Biomedicals).

Marker Selection and Molecular Protocol

A PCR amplification survey of the nuclear ribosomal RNA internal transcribed spacer (nrITS) region (Baldwin, 1992; Schneeweiss et al., 2004; Wen & Zimmer, 1996; Wolfe & Randle, 2001) and nuclear gene phytochrome A (PHYA) (Bennett & Mathews, 2006) was conducted on two accessions of each of three varieties of *Melampyrum lineare* and three accessions of a fourth variety, as well as one accession of *M. laxum*, one accession of *M. laxum* var. *arcuatum*, one accession of *M. nemorosum*, and one accession of *M. cristatum*. The primers used and their sequences are provided in Table 4, and the primer combinations tried for PHYA amplification are provided in Table 5. For nrITS, temperature cycling parameters included 3 min (96° C) followed by 36 cycles of 1 min (94° C), 1 min (53° C), 45 s (72° C), ending with 7 min (72° C). PCR of PHYA used temperature cycling parameters of 4 min (94° C) followed by 35 cycles of 45 s (94° C), 45 s (70° C, decreasing 1°C per cycle), 1 min (72° C), ending with 4 min (72° C). PCR products were purified with the exo-sap method (Dugan et al., 2002), and samples were bidirectionally sequenced (1x coverage of the 5' to 3' strand, 1x coverage of the 3' to 5'

strand) by Macrogen USA (Bethesda, MD, USA) using ABI 3730x1 DNA Analyzers (Applied Biosystems, Foster City, CA, USA).

Data Analysis

Bidirectional sequence data were assembled in Sequencher® (Gene Codes Corporation) and viewed in Mesquite (Maddison & Maddison, 2011). Sequences of nrITS from the eight *Melampyrum lineare* accessions and one accession each of *M. laxum*, which did not have a varietal designation, *M. laxum* var. *arcuatum*, *M. cristatum*, and *M. nemorosum* were aligned using MUSCLE (Edgar, 2004) against the nrITS of 33 other *Melampyrum* accessions and 19 other species of Rhinanthae (McNeal et al., 2013), which were downloaded from GenBank (Table 6). Accessions obtained from GenBank were mostly derived from the published work of McNeal et al. (2013) and Li et al. (2008). Two accessions (McNeal et al., 2013), *Brandisia hancei* and *Pterygiella nigrascens*, were included as outgroups based on a recent family-level phylogeny (McNeal et al., 2013). Both species are members of Orobanchaceae. *Brandisia hancei* is sister to the clade containing all species in the Rhinanthae tribe, which includes *Melampyrum* species. *Pterygiella nigrascens* is included in Rhinanthae and is sister to the clade containing all other members of the tribe included in the analysis done by McNeal et al. (2013).

Phylogenetic Analysis

The aligned matrix was analyzed under parsimony in PAUP* v. 4 beta 10 (Swofford, 2002). Heuristic parsimony searches included 1000 random addition replicates, TBR with MulTrees, followed by swapping on all shortest trees. Bootstrap

support was assessed using 1000 pseudoreplicates and the same search criteria. To obtain Bayesian posterior probabilities, the data were run in MrBayes version 3.1.2 (parallel implementation) (Altekar et al., 2004) for 10 million generations, saving one tree per 1000 generations using the GTR+I+G model, as determined by jModeltest (Darriba et al., 2012; Guindon & Gascuel, 2003). This model allows for different mutation rates for every different type of base substitution, a portion of the sites to be considered invariable, and substitution rates to vary among sites. Because the starting point of the analysis creates a bias in initial estimates of the tree posterior probabilities, the early iterations of the algorithm should be discarded as burn-in. For this analysis, the burn-in was set at 5 million generations and trees were constructed from 10,000 trees pooled from two simultaneous runs. Based on the standard deviation of split frequencies at the end of this analysis (.003), it was determined that convergence between two independent runs had occurred. Stepping stone analysis was also conducted in MrBayes to test competing topological hypotheses using the GTR+I+G model, 10,000,000 generations, and a sampling frequency of every 1000 trees. Three topologies were compared: 1) a forced sister grouping between *Melampyrum lineare* and *M. pratense* to represent Pennell's hypothesized evolutionary relationships and denoted as the MLMP constraint, 2) forced monophyly of *M. lineare* and the European species of *Melampyrum* to represent the hypothesis of European descent and denoted as the EuroML constraint, and 3) forced monophyly of *M. lineare* and the Asian species of *Melampyrum* to represent the hypothesis of Asian descent, denoted as the AsiaML constraint.

RESULTS

The Varietal Delimitation of *Melampyrum lineare*

Character Selection

Of the 44 characters measured, 2 characters were constant or near constant for all specimens regardless of varietal designation; for all specimens, calices were pubescent, and all but a few specimens had a stigma that remained entirely inside the flower and did not protrude into the corolla tube opening. On occasions when the stigma was visible in the corolla tube opening, it was because a portion of the corolla had been folded back during pressing, revealing the stigma and making it difficult to reliably determine whether or not the stigma had been entirely inside the flower before the specimen was pressed. Both of these floral characters were omitted from analyses.

Boxplots for all 37 continuous characters are shown in Figures 5 – 11. Of the 14 continuous vegetative characters measured, 6 did not have a univariate normal distribution: leaf width, lowest bract width, bract length, bract width, the width of the undivided portion of the bract used for the bracteal tooth length measurement, and the ratio between bracteal tooth length and the width of the undivided portion of the bract. A log transformation of these variables greatly improved univariate normality for all 6 characters (Figures 12 – 19). After confirming univariate normality for each vegetative character, the set of all vegetative characters still did not have a multivariate normal distribution. A list of pairwise correlations between vegetative characters showed ten relatively high correlative relationships ($r > 0.7$) (Table 7). The 4 most correlated characters – lowest bract width, lowest bract length, bract width, and leaf width – were

eliminated from the analyses one by one until a multivariate normal distribution was achieved for the vegetative character set (Figure 20).

All floral characters had univariate normal distributions without transformations (Figures 12 – 19). The floral character set had a nearly multivariate normal distribution without univariate transformations or further character selection (Figure 21). However, multivariate normality greatly improved when 5 of the 15 floral characters were omitted from analyses due to relatively high correlations with other floral characters: lower corolla lip length, upper corolla lip length, total calyx length, the length of the unfused portion of the calyx, stamen length, and pistil length (Table 8).

Of the 8 characters in the fruit and seed character set, all had univariate normal distributions (Figures 12 – 19). However, the fruit and seed character set did not have a multivariate normal distribution. Of the 14 specimens from var. *lineare*, 10 had missing seed data, so the sample size for this variety in the fruit and seed character set was reduced to 4. To improve multivariate normality, the ratio between elaiosome length and seed length was omitted because it correlated highly with elaiosome length ($r = 0.7898$, $p < 0.001$), and seed length was removed because it correlated somewhat highly with seed width ($r = 0.6943$, $p < 0.001$) (Figure 22).

Based on the results of the Brown-Forsythe test for homogeneity of variances (Table 9), 4 vegetative characters failed to meet this assumption: lowest fruiting internode length, maximum bracteal tooth length, the ratio between lowest bract width and lowest bract length, and the ratio between leaf width and leaf length. These 4 characters were omitted from further analyses. All floral characters met the assumption

for homogeneity of variances. Of the 8 fruit and seed characters, only 1 – capsule width – failed to meet the assumption of equal variances and was consequently excluded from analyses.

In total, 8 of the 17 original continuous vegetative characters were excluded to improve multivariate normality and meet the assumption of equal variances. Six of the 15 original continuous floral characters and 3 of the 8 original continuous fruit and seed characters had to be excluded to meet these assumptions. Table 10 contains lists of variables included in each analysis and whether multivariate normality and homogeneity of variances assumptions were met. Because various authors suggest that these two assumptions can be violated with minimal impact on the power of linear discriminant analysis, two sets of analyses were run for each set of characters: one where the assumptions of a multivariate normal distribution and equal variances across groups were met, and one where all characters were included, resulting in a failure to meet these assumptions.

The vegetative character set, floral character set, and fruit and seed character set all failed Box's M test for homogeneity of covariances ($p < 0.001$ in all cases). For Box's M test, the null hypothesis is equal covariances, so a small p-value supports a conclusion of unequal covariances. However, because of issues with this test's reliability (Hill & Lewicki, 2007), these results do not necessarily translate into a violation of this assumption or notably decreased power of the linear discriminant analysis. Following Hill and Lewicki's suggestion (2007), an additional linear discriminant analysis was run for each character set excluding the var. *lineare* group because of its comparably small

sample size. If the LDA results were relatively consistent regardless of one group's inclusion or exclusion, the assumption of homogeneous covariances was most likely not violated.

Multivariate Analysis of Variance and Discriminant Analysis

In total, eight sets of MANOVA and LDA analyses were run (Table 10). Analyses 1 through 3 in used only vegetative characters, analyses 4 and 5 used only floral characters, and analyses 6 through 8 used only fruit and seed characters. The characters and varieties included in each are summarized in Table 10, along with which assumptions each analysis met or failed to meet. All eight MANOVA analyses confirmed that there were statistically significant differences across the four varieties ($p < 0.001$ in all cases).

The discriminant analysis that was able to most accurately predict varietal designation was analysis 1, which included all continuous vegetative characteristics with no transformations, despite the fact that this analysis failed to meet the assumptions of multivariate normality and homogeneous variances and covariances. The model developed in this analysis was able to accurately classify specimens 75.00% of the time. However, when jackknifing was incorporated to estimate how well the model would perform with new data, the accuracy of the model's specimen classification predictions dropped to 69.35%.

Both 2-dimensional and 3-dimensional plots for analysis 1 showed the clearest pattern of clustering within varietal designations of all analyses, but there was no obvious separation between the different varieties (Figures 23 – 30). Rather, the plots showed a continuum of morphological variation that primarily occurred along the axis for the first

discriminant function. The two characters that contributed most to the first discriminant function for this analysis were the ratio between bract width and bract length and the ratio between lowest bract width and lowest bract length (Table 11), which is consistent with the prominent role bract width plays in current varietal descriptions.

The highest predictive capability with jackknifing was accomplished by the model from analysis 3 in Table 10, at 73.50% accuracy. The same analysis had almost identical predictive accuracy (73.93%) without jackknifing. This vegetative dataset contained only those characters that met the assumptions of multivariate normality and equal variances. Although this analysis failed Box's M test for equal covariances, the results were almost identical when var. *lineare* was included (analysis 2 in Table 10) versus when it was excluded (analysis 3). This suggests that either the conclusions from Box's M test were misleading and the assumption of homogeneous covariances was met, or the violation of this assumption was not strongly impacting the model's power to detect differences between the purported varieties. The 2-dimensional and 3-dimensional plots for analysis 2 showed no obvious separation between the different varieties (Figure 24). Analysis 3 only yielded two discriminant functions, so only a 2-dimensional plot could be generated. This plot also showed no obvious separation between the varieties (Figure 25).

Both the floral character set and fruit and seed character set were unable to accurately predict specimen classifications in analyses 6 through 8 (Table 10). Discriminant function plots generated with these character sets did not show any separation of the varieties (Figures 26 – 30). This is not surprising given the complete

lack of floral, fruit or seed characters in current keys used to distinguish among the varieties of *Melampyrum lineare*.

Chi-Squared Contingency Tests

Discrete characters were not normally distributed and could therefore not be included in MANOVA nor LDA analyses. Instead, they were organized into tables (Tables 12 – 17) and subjected to chi-squared contingency tests. Five characters in total, all vegetative, were evaluated this way: the presence or absence of primary branching, secondary branching, and tertiary branching, the number of bracteal teeth per side of the bract used for the bract width and bract length measurements, and the maximum number of bracteal teeth per side of a bract on each specimen.

All of these characters were found to be dependent on varietal designation (Tables 12 – 17). However, these tests are not sufficient for establishing rules that can be used to predict varietal designations. For example, tertiary branching was totally absent from var. *lineare*, but otherwise, at least some specimens from every variety had primary, secondary and tertiary branching. However, several specimens from all other varieties also lacked tertiary branching, so the absence of tertiary branching does not guarantee that a specimen belongs to var. *lineare*. These tests do suggest that var. *pectinatum* tends to have a higher degree of branching than other varieties, while var. *lineare* tends to have the lowest degree of branching.

These tests also suggest that var. *pectinatum* and var. *americanum* have roughly the same number of bracteal teeth per side of a bract and both of these varieties tend to have more bracteal teeth than var. *latifolium* or var. *lineare*, but as was the case with

branching patterns, these trends are not enough to predict variety based on these characters. For example, only specimens from var. *americanum* and var. *pectinatum* had bracts with 5 bracteal teeth per side. However, the majority of specimens from these two varieties did not have bracts with that many bracteal teeth per side, so the presence of bracts with 5 bracteal teeth per side is not sufficient to determine that a specimen belongs to var. *americanum* or var. *pectinatum*.

Habitat moisture levels were only reliably recorded for 62 of 248 specimens. Of these 62 specimens, 14 were determined to belong to var. *americanum*, 20 were determined to be var. *latifolium*, 9 were determined to be var. *lineare*, and 19 were determined to be var. *pectinatum* (Table 17). A chi-squared contingency test to determine if varietal designations depended on habitat moisture levels failed to find evidence supporting such a conclusion ($p = 0.121$).

The Historical Biogeography of *Melampyrum lineare*

Only nrITS amplified for all accessions of *Melampyrum* species and was sequenced successfully. PHYA did not yield PCR products despite trying different primer pairs and cycling parameters.

The aligned 704 bp dataset of 65 individuals from 34 species yielded 265 (38%) phylogenetically informative characters and 45,409 most-parsimonious trees (CI = 0.545, RI = 0.830). *Melampyrum lineare* is well nested within the *Melampyrum* clade with high statistical support. There is no statistical support for the phylogenetic relationships of accessions within the *M. lineare* clade, nor do the varieties of *M. lineare* form monophyletic groups. The 50% majority-rule consensus tree (Figure 31) places *M.*

lineare as sister to a clade containing three East Asian species, *M. laxum*, *M. roseum*, and *M. klebelsbergianum*. The strict consensus tree (Figure 32) is less resolved than the 50% majority-rule consensus tree; it places *M. lineare*, the East Asian clade, European *M. pratense*, and the remaining Eurasian species as separate clades in a polytomy. The bootstrap 50% majority rule consensus tree did not provide statistical support (60% BS) for the clade containing only *M. lineare* and the East Asian species. The Bayesian 50% majority rule tree (Figure 33) did not resolve *M. lineare* as sister to East Asian species; rather, this tree supported the European *M. pratense* as sister to the rest of the European species, *M. lineare* sister to this group, and the Asian clade sister to the *M. lineare* and European clade. There is strong support (0.91) for the node joining *M. pratense* to the rest of the European members of the genus. However, the posterior probability for the node joining *M. lineare* to the European clade was not statistically significant (0.60).

Stepping stone analysis calculated mean marginal likelihoods (natural log units) of -6632.37, -6625.12, and -6613.46 for the MLMP, AsiaML, and EuroML topological constraints. Since a difference of 5 ln units is considered strong support for the less negative of the two values being compared (Kass & Raftery, 1995), there is strong evidence suggesting that both the AsiaML and EuroML constraints are better than the MLMP constraint, and the EuroML constraint is better than the AsiaML constraint. These findings conflict with the parsimony 50% majority-rule consensus tree but do not conflict with the relatively unresolved parsimony strict consensus tree.

DISCUSSION

The Varietal Delimitation of *Melampyrum lineare*

It is not uncommon for infraspecific taxonomic groups to be identified based on a limited number of specimens, only to see these distinctions disappear when a larger scale revision of the species is done (Alexander et al., 2012; Brunell & Whitkus, 1998; Fritsch & Lucas, 2000). While MANOVA consistently indicated statistically significant differences between purported varieties of *Melampyrum lineare*, this alone is not enough to justify the recognition of these varieties because it is still possible for morphological variation to occur along a continuous spectrum. Ideally, infraspecific groups should be reliably distinguishable by morphological characters (Fritsch & Lucas, 2000). There is a precedent for combining infraspecific groups when those groups do not always separate out into distinct morphological types but instead contain intermediate morphologies as well (Alexander et al., 2012; Fritsch & Lucas, 2000). Linear discriminant analysis is therefore a useful tool because it not only allows one to visualize plots that capture very well the variation among groups, but it also allows one to use the model to predict assignments. This provides a less objective way to evaluate the model's utility in classifying individuals to variety rather than subjectively deciding whether a plot shows distinct groups or not.

However, none of the linear discriminant analyses run in this study showed distinct groupings of specimens based on varietal designation, and all models failed to reliably predict a specimen's varietal classification. Rather, all analyses showed a

continuum of variation, with specimens of each variety overlapping in morphology with specimens of at least one other variety.

Hill and Lewicki (2007) caution that correlations between means and variances across groups can be a potentially serious violation of the homogeneous variances and covariances assumption. As Hill and Lewicki (2007) explain: “Intuitively, if there is large variability in a group with particularly high means on some variables, then those high means are not reliable. However, the overall significance tests are based on pooled variances, that is, the average variance across all groups. Thus, the significance tests of the relatively larger means (with the large variances) would be based on the relatively smaller pooled variances, resulting erroneously in statistical significance.”

However, the type of statistical error Hill and Lewicki (2007) described would result in recognizing differences between groups when in truth there are none. Since the analyses included in this project failed to identify distinct groups, this particular source of error does not seriously challenge the validity of the conclusion that the varieties of *Melampyrum lineare* do not form truly distinct taxonomic groups.

Analysis of discrete variables that quantified degree of branching and bracteal toothiness suggested that trends do exist among these characters and varietal classifications. For example, var. *pectinatum* often has a higher degree of branching and more bracteal teeth than other varieties, while var. *lineare* often branches minimally and has fewer bracteal teeth than other varieties. However, none of the characters analyzed in this fashion identified trends pronounced enough to be useful for classifying individual specimens as belonging to one of the purported varieties versus another.

Varietal classifications appeared not to be dependent on habitat moisture levels. This was an unexpected result based on subjective observations in the field and information from herbarium specimens, but it does not rule out the possibility that moisture levels influence the morphological variation in this species, in particular leaf width. Not all of the specimens analyzed in this project had labels that clearly identified whether a habitat was dry or moist. Moreover, relying on herbarium label information to investigate such an ecological hypothesis is inexact at best. Using a binary system of “dry” and “moist” gives only a rough estimate of how moist a particular habitat truly is, and herbarium label information only provides the collector’s subjective account of habitat moisture conditions. In order to establish the presence or absence of a causal relationship between moisture levels and morphology, several other variables would need to be considered and controlled for, such as host species, how well connected an individual is to its host the plant, amount of sunlight, nutrient concentrations, soil type, and ambient temperatures; yet herbarium label information rarely provides this type of information. It is also likely that the phenotypic response of *M. lineare* to multiple environmental variables is affected by the population genetic diversity of this species, which was not characterized in this study.

Although these results lead to the conclusion that there is no strong evidence for a relationship between habitat moisture levels and varietal classifications, more research is needed to determine exactly what role, if any, habitat moisture levels, other environmental variables, or genetic factors have on the morphology of *Melampyrum lineare*. Interestingly, highly variable phenotypes are common in the genus *Melampyrum*.

Tutin et al. (1972) identified 4 common “ecotypes” for European *Melampyrum* species. They noted that these ecotypes are not intended to represent formal taxonomic groups but are instead an attempt to recognize infraspecific variation that “cut[s] across the normal pattern of geographical speciation and subspeciation” (Tutin et al., 1972). Similarly, there is insufficient evidence to support formal varieties based on the morphological variation observed in *M. lineare*.

Future Research

It is likely that the morphological variation observed in *Melampyrum lineare* is the result of the interaction between environmental factors and population genetics. Controlled laboratory or common garden experiments would provide the best means of conclusively establishing a causal relationship between environmental factors and morphological variation. Lacking such controlled experiments, including more detailed habitat information for herbarium collections could also shed light on the relationship between environmental factors and morphological variation in this species. If a clear relationship between habitat and morphological variation were established, it could be useful to identify common “ecotypes” as done for European *Melampyrum* species (Tutin et al., 1972), but these groups would likely not represent formal taxonomic classifications. Investigating the population genetics of the species could also potentially shed new light on the question of infraspecific taxonomic groups or ecotypes. However, using these data to create botanical keys for any such genetic groups would prove challenging for *Melampyrum lineare* because the morphological variation in the species occurs along a continuum.

The Historical Biogeography of *Melampyrum lineare*

Contrary to the previous hypothesis (Pennell, 1935), *Melampyrum lineare* is not sister to European *M. pratense*. Rather, the results from this study provide mixed support for the competing hypotheses of European versus Asian ancestry. Analysis of the nrITS region does not provide support for the hypothesis that *M. lineare* is of East Asian ancestry, nor does it provide strong support for the hypothesis that *M. lineare* is of European ancestry. Parsimony analysis yielded no significant results but weakly favored a topology in which *M. lineare* shared its most recent common ancestor with the three East Asian species included in this analysis, while Bayesian inference weakly favored a topology that suggest European ancestry, again with no statistical support.

One alternative hypothesis, with which these results do not conflict, but for which there is also no statistical support, is that a widespread ancestor may have become distributed across the Northern Hemisphere and subsequently evolved into separate continental lineages. This alternative scenario does not specify whether the original migration was west-bound (i.e., North America to East Asia) or east-bound (East Asia to North America). However, based on the work of Wolfe et al. (2005), it is likely that the genus *Melampyrum* diversified too recently for this scenario to be possible.

Future Research

Further tests and refinements of these biogeographical hypotheses will require more extensive taxon sampling than the current phylogeny contains, as *Melampyrum* contains an additional ca. 22 species not included in this analysis. Including insertion/deletion characters of the nrITS data set and adding DNA sequence data from additional genetic markers would potentially improve phylogenetic resolution and

therefore increase support for one biogeographical hypothesis over another. Maximum likelihood phylogenetic analysis methods could also be used to test conclusions drawn from parsimony and Bayesian analysis.

APPENDIX A: TABLES

Table 1. Herbaria from which specimens were contributed to this study.

Herbarium	Institution Code	No. Specimens Contributed	No. Specimens Measured
Academy of Natural Sciences, Philadelphia	PH	705	49
College of William and Mary	WILLI	20	3
Cornell	BH	165	16
George Mason University	GMUF	27	5
Harvard	GH	59	13
Longwood University	FARM	2	2
Lynchburg College	LYN	29	4
Montana State University	MONT	26	3
New York Botanical Garden	NY	390	26
Ohio State University	OS	68	6
Old Dominion University	ODU	13	0
Royal British Columbia Museum	V	0	0
Université de Montréal	MT	340	19
University of British Columbia	UBC	64	7
University of Connecticut	CONN	180	20
University of Idaho	ID	1	1
University of Michigan	MICH	219	25
University of Minnesota	MIN	241	20
University of Montana	MONTU	20	1
University of North Carolina	NCU	102	8
University of Richmond	URV	3	2
University of Tennessee, Knoxville	TENN	107	6
University of Vermont	VT	70	0
University of Washington	WTU	26	1
University of Wisconsin	WIS	0	0
Virginia Tech	VPI	35	5
West Virginia University, Morgantown	WVA	68	6
Total		2,980	248

Table 2. Characters recorded for each specimen measured.

No.	Character Type	Character	Abbreviation	Description
	Label Information	Herbarium	herb	
	Label Information	Accession Number	accno	
	Label Information	Location	loc	BC = British Columbia QU = Quebec ON = Ontario PE = Prince Edward Island MT = Manitoba NB = New Brunswick NF = Newfoundland and Labrador NS = Nova Scotia SK = Saskatchewan AB = Alberta Others = US State Abbreviations
	Label Information	Habitat Moisture Level	hab	Discrete 0 = Dry 1 = Moist
	Label Information	Variety	var	A = var. <i>americanum</i> LA = var. <i>latifolium</i> LI = var. <i>lineare</i> P = var. <i>pectinatum</i>
1	Vegetative	Primary branching	br1	Discrete 0 = Absent 1 = Present
2	Vegetative	Secondary branching	br2	Discrete 0 = Absent 1 = Present
3	Vegetative	Tertiary branching	br3	Discrete 0 = Absent 1 = Present
4	Vegetative	Plant height	height	Distance (mm) from uppermost point on main axis to just above beginning of roots and underground stem
5	Vegetative	Lowest fruiting internode length	lfi	Distance (mm) between the lowest capsule pair on the main axis and the node above it.
6	Vegetative	Leaf length	lfl	Length (mm) from the base of the petiole to leaf tip of the lowest leaf or lowest in tact leaf on the main axis.
7	Vegetative	Leaf width	lfw	Width (mm) of the widest part of the leaf used for the leaf length measurement, measured perpendicular to main leaf vein.
8	Vegetative	Leaf width/leaf length	lfwtol	The ratio between leaf width and leaf length.
9	Vegetative	Lowest bract length	lbl	Length (mm) from the base of the petiole to bract tip of the lowest bract or lowest in tact bract on the main axis.
10	Vegetative	Lowest bract width	lbw	Width (mm) of the widest part of the bract used for the lowest bract

				length measurement, measured perpendicular to main leaf vein.
11	Vegetative	Lowest bract width/lowest bract length	lbwtol	The ratio between lowest bract width and lowest bract length.
12	Vegetative	Bract length	brl	Length (mm) from the base of the petiole to bract tip of the bract approximately halfway between the lowest bract and the apex of the plant.
13	Vegetative	Bract width	brw	Width (mm) of the widest part of the bract used for the bract length measurement, measured perpendicular to main leaf vein.
14	Vegetative	Bract width/bract length	brwtol	The ratio between bract width and bract length.
15	Vegetative	Teeth per side of bract	tpers	Discrete Number of bracteal teeth per side of bract used in bract width and bract length measurements. Bracteal teeth grow symmetrically, so tallying teeth on one side sufficiently quantifies number of teeth per bract.
16	Vegetative	Maximum bracteal teeth/plant	maxt	Discrete Maximum number of bracteal teeth per side of bract on entire specimen
17	Vegetative	Length of longest tooth	tlen	Length (mm) from tip of longest bracteal tooth to point where the tooth base joined with undivided portion of the bract.
18	Vegetative	Width of undivided portion of bract used for longest tooth measurement	tlwid	Width (mm) of the undivided portion of the bract used for longest tooth measurement, measured at widest part of bract excluding teeth.
19	Vegetative	Maximum tooth length/width of undivided portion of bract	tlntotlw	Ratio between length of longest tooth and width of undivided portion of bract.
20	Floral	Corolla length	corlen	Length (mm) of corolla, measured from the base of the calyx to the tip of the lower lip of the corolla. Measurement did not include the pedicel.
21	Floral	Corolla width	corwid	Width (mm) of the widest part of the corolla, usually from the top of the upper lip to the bottom of the lower lip.
22	Floral	Lower corolla tube lip length	lliplen	Length (mm) from the tip of the lower lip of the corolla to the point where the lower and upper lips fused together to form the corolla tube.
23	Floral	Upper corolla tube lip length	uliplen	Length (mm) from the tip of the upper lip of the corolla to the point where the lower and upper lips

				fused together to form the corolla tube.
24	Floral	Lower corolla lip length/corolla length	llltocol	Ratio between the lower lip length and total corolla length.
25	Floral	Upper corolla lip length/corolla length	ulltocol	Ratio between the upper lip length and total corolla length.
26	Floral	Upper corolla lip length/lower corolla lip length	ulltolll	Ratio between upper lip length and lower lip length.
27	Floral	Total calyx length	callen	Length (mm) from the tip of the longest unfused portion of the calyx to base of the calyx.
28	Floral	Length of separated portion of calyx	seplen	Length (mm) from the tip of the longest unfused portion of the calyx to the point where the calyx fuses.
29	Floral	Length of separated portion of calyx/total calyx length	sltocal	Ratio between the length of the separated portion of calyx to total calyx length.
30	Floral	Total calyx length/corolla length	caltocol	Ratio between total calyx length and total corolla length.
31	Floral	Calyx pubescence	calpub	Discrete 0 = Absent 1 = Present
32	Floral	Stamen length	staml	Length (mm) of longest stamen, measured from the base of the calyx to the tip of the anther.
33	Floral	Anther length	anthl	Length (mm) of the anther from stamen used in longest stamen measurement.
34	Floral	Anther width	anthw	Width (mm) of the widest part of the anther used for the anther length measurement.
35	Floral	Pistil length	pistl	Length (mm) of pistil, measured from the base of the calyx to the tip of the pistil.
36	Floral	Stigma position	stigp	Discrete 0 = Positioned inside corolla; not visible without dissecting flower. 1 = Positioned outside corolla; visible without dissecting flower
37	Fruit/Seed	Capsule length	capl	Length (mm) of lowest undehiscent capsule, measured in a straight line from calyx base to capsule tip.
38	Fruit/Seed	Capsule width	capw	Width (mm) of the widest part of the capsule used for capsule length measurement.
39	Fruit/Seed	Capsule width/capsule length	cpwtocpl	Ratio between capsule width and capsule length.
40	Fruit/Seed	Seed length	seedl	Length (mm) of seed including elaiosome, measured in a straight line from one end of the seed to the other.
41	Fruit/Seed	Seed width	seedw	Width (mm) of the widest part of the seed used for seed length measurement.

42	Fruit/Seed	Elaiosome length	elsl	Length (mm) of elaiosome on seed used for seed length measurement. Measured along the same axis as seed length, from tip of elaiosome to point where it meets with the rest of the seed.
43	Fruit/Seed	Seed width/seed length	sdwtosdl	Ratio between seed width and seed length.
44	Fruit/Seed	Elaiosome length/seed length	elltosdl	Ratio between elaiosome length and seed length.

Table 3. Accession information for specimens used in DNA extractions for phylogenetic analyses. Blank cells indicate information that was absent from the specimen label.

Institution Code	Accession No.	Date Collected	State/Province	Collector	Collector No.	Varietal Designation
GMUF	46295	7/29/1997	VA	Ted Bradley		<i>latifolium</i>
GMUF	46316	6/29/1980	PA	Fred H.		<i>americanum</i>
GMUF	46320	8/2/1973	MA	H.E. Ahles		<i>lineare</i>
GMUF	46321	8/24/1971	MA	T. and S. Grimshaw		<i>pectinatum</i>
GMUF	46323	7/2/1974	QU	Dave Williams		<i>lineare</i>
ODU		7/11/1983	ID	Douglass Henderson	6703	<i>americanum</i>
VPI	83120	7/1972	QU	D.E. Boufford	6577	<i>lineare</i>
VPI	89206	5/15/1991	GA	T. F. Weiboldt	7600	<i>latifolium</i>

Table 4. Primer sequences used in PCR amplification survey for nrITS and PHYA.

Primer Name	Region	Forward/Reverse	Sequence	Citation
N-nc18S10	nrITS1	Forward	AGGAGAAGTCGTAACAAG	Wen & Zimmer, 1996
nrITS2m.Scroph	nrITS1	Reverse	CTTGCGTTCAAAGACTCG	Wolfe & Randle, 2001
nrITS3	nrITS2	Forward	GCATCGATGAAGAACGCAGC	(Baldwin, 1992)
AB102	nrITS2	Reverse	TAGAATTCCCCGGTTCGCTCGCCGTTA	(Schneeweiss et al., 2004)
212f	PHYA	Forward	TCWGGNAARCCNTTYTAYGC	Bennett & Mathews, 2006
a230f.ORO	PHYA	Forward	GAYTTYGAGCCYGYNAADCCYYAYG	Bennett & Mathews, 2006
a236f.ORO	PHYA	Forward	CCYYAYGAKGTBCCCHATGASYGC	Bennett & Mathews, 2006
377f	PHYA	Forward	CARTAYATGGCNAAYATGG	Bennett & Mathews, 2006
444f	PHYA	Forward	CARGTNTTYGCHATHCAYG	Bennett & Mathews, 2006
a624f.ORO	PHYA	Forward	GAYTWYGARATGGAYGCRAT	Bennett & Mathews, 2006
432r	PHYA	Reverse	CRCANGCRTANCKNARNGGRWANGG	Bennett & Mathews, 2006
444r	PHYA	Reverse	CRTGGATGGCRAANACYTG	Bennett & Mathews, 2006
a575r.ORO	PHYA	Reverse	KCHGTGTKNGACCKRAACCA	Bennett & Mathews, 2006
a624.ORO	PHYA	Reverse	ATYGCRTCCATYTCRSARTC	Bennett & Mathews, 2006
a678r	PHYA	Reverse	GTYTCMATBARDCKRACCATYTC	Bennett & Mathews, 2006
a788r.ORO	PHYA	Reverse	GHGCDATGAARCAYRCKCC	Bennett & Mathews, 2006
a832r.2	PHYA	Reverse	RTTCCAYTCNGARCACCANCC	Bennett & Mathews, 2006

Table 5. Primer combinations tried in PHYA PCR. Dark grey boxes indicate primer combinations that were tried because they captured >1000 bp sequences from three *Melampyrum* species in GenBank: *M. pratense* (GenBank accession number AM233981), *M. carstiense* (GenBank accession number AM233980), and *M. arvense* (GenBank accession number AM233919). This indicates a higher probability for successful amplification in *M. lineare*. Light grey boxes indicate primer combinations that were tried but did not meet these criteria; white boxes indicate primer combinations that were not tried. Boxes with X's in them indicate primer combinations that overlap with each other and therefore only capture the primer sequence, so they are unlikely ever to amplify.

Primer	432r	444r	a575r.ORO	a624r.ORO	a678r	a788r.ORO	a832r.2
212f							
a230f.ORO							
a236f.ORO							
377f							
444f		X					
a624f.ORO				X			

Table 6. References and accession numbers for sequences obtained from GenBank.

Species Name	Author	Citation in GenBank	GenBank No.
<i>Melampyrum arvense</i>	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503876
<i>Melampyrum arvense</i>	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503874
<i>Melampyrum arvense</i>	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503875
<i>Melampyrum carstiense</i>	Dong, L.N et al.	J Syst Evol 49 (3), 189-202 (2011)	GU445314
<i>Melampyrum carstiense</i>	Morawetz, J.J. et al.	Unpublished	EU259252
<i>Melampyrum cristatum</i>	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503872
<i>Melampyrum cristatum</i>	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503873
<i>Melampyrum cristatum</i>	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503871
<i>Melampyrum italicum</i>	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503858
<i>Melampyrum italicum</i>	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503856
<i>Melampyrum italicum</i>	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503857
<i>Melampyrum klebelsbergianum</i>	Dong, L.N et al.	J Syst Evol 49 (3), 189-202 (2011)	GU445315
<i>Melampyrum nemorosum</i>	Dong, L.N et al.	Unpublished	GU445316
<i>Melampyrum nemorosum</i>	Tesitel, J. et al.	Unpublished	FJ797592
<i>Melampyrum pratense</i>	Tesitel, J. et al.	Unpublished	FJ790039
<i>Melampyrum pratense</i>	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503870
<i>Melampyrum pratense</i>	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503868
<i>Melampyrum pratense</i>	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503869
<i>Melampyrum roseum</i>	Yang, F.S. & Wang, X.Q.	Plant Syst. Evol. 264 (3-4), 251-264 (2007)	AY881140
<i>Melampyrum roseum</i>	Bae, Y.-M.	Unpublished	GU359046
<i>Melampyrum saxosum</i> Haplotype A	Tesitel, J.	Unpublished	EU624125
<i>Melampyrum</i> sp. JW_13_02_07	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503860
<i>Melampyrum sylvaticum</i>	Wolfe, A.D. et al.	Folia Geobot. 40, 115-134 (2005)	AY911232
<i>Melampyrum sylvaticum</i>	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503864
<i>Melampyrum velebiticum</i>	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503866
<i>Melampyrum velebiticum</i>	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503867
<i>Melampyrum velebiticum</i>	Li, M. et al.	Cladistics 24 (5), 727-745 (2008)	AM503865

Table 7. List of ten strongest correlations for vegetative characters.

Characters	Correlation Coefficient (r)	p-value
lfw; lfwtol	0.879	< 0.001
lbw; brw	0.860	< 0.001
lbl; brl	0.814	< 0.001
lfw; lbw	0.799	< 0.001
lbl; lbw	0.791	< 0.001
lbl; brw	0.748	< 0.001
brl; brw	0.755	< 0.001
lbw; lfwtol	0.745	< 0.001
lfl; lbl	0.721	< 0.001
lfl; lfw	0.716	< 0.001

Table 8. List of seven strongest correlations for floral characters.

Characters	Correlation Coefficient (r)	p-value
corlen; pistl	0.941	< 0.001
staml; pistl	0.923	< 0.001
corlen; staml	0.922	< 0.001
callen; seplen	0.821	< 0.001
llltocol; lliplen	0.804	< 0.001
corwid; staml	0.698	< 0.001
uliplen; ulltocol	0.695	< 0.001

Table 9. Results from Brown-Forsythe test for equal variances for all continuous characters. A p-value ≤ 0.01 was considered significant, indicating unequal variances among varieties for that character, and was marked with an asterisk.

Character	p-value
height	0.042
lfi	<0.001*
lfl	0.012
lfw	0.710
lfwtol	0.004*
lbl	0.112
lbw	0.171
lbwtol	0.002*
brl	0.150
brw	0.644
brwtol	0.147
tlen	< 0.001*
tlntotlw	0.103
corlen	0.525
corwid	0.974
lliplen	0.073
uliplen	0.456
llltocol	0.116
ulltocol	0.028
ulltolll	0.231
callen	0.509
seplen	0.946
sltocal	0.825
caltocor	0.794
staml	0.608
anthl	0.767
anthw	0.034
pistl	0.236
capl	0.063
capw	0.0104
cpwtocpl	0.958
seedl	0.796
seedw	0.045
elsl	0.429
sdwtosdl	0.134
elltosdl	0.248

Table 10. Summary of linear discriminant analyses run and results.

No.	Character Set	Varieties Excluded	Characters Included	Characters Excluded	Multivariate Normality	Equal Variances	Equal Covariances	Percent Specimens Accurately Classified (No Jackknifing)	Percent Specimens Accurately Classified (Jackknifing)
1	Vegetative	none	height, lfi, lfl, lfw, lbl, lbw, brl, brw, tlen, tlwid, lfwtol, lbwtol, brwtol, tlntotlw	none	no	no	no	75.00	69.35
2	Vegetative	none	height, lfi, log(brl), tlwid, brwtol, log(tlntotlw + 1)	lfi, lfw, lbl, lbw, brw, tlen, lfwtol, lbwtol	yes	yes	no	73.39	71.37
3	Vegetative	<i>lineare</i>	height, lfi, log(brl), tlwid, brwtol, log(tlntotlw + 1)	lfi, lfw, lbl, lbw, brw, tlen, lfwtol, lbwtol	yes	yes	no	73.39	73.50
4	Floral	none	corlen, corwid, lliplen, uliplen, callen, seplen, staml, anthl, anthw, pistl, llitocol, ulltocol, ulltrolll, sltocal, caltocor	none	no	yes	no	51.21	44.35
5	Floral	none	corlen, corwid, anthl, anthw, llitocol, ulltocol, sltocal, caltocor	staml, pistl, callen, lliplen, uliplen, seplen	yes	yes	yes (p = 0.020)	45.97	43.15
6	Fruit/Seed	none	capl, capw, seedl, seedw, elsl, cpwtocpl, sdltosdw, elltosdl	none	no	no	no	51.48	49.37
7	Fruit/Seed	none	capl, seedw, elsl, cpwtocpl, sdltosdw	capw, seedl, elltosdl	yes	yes	no	51.90	49.37
8	Fruit/Seed	<i>lineare</i>	capl, seedw, elsl, cpwtocpl, sdltosdw	capw, seedl, elltosdl	yes	yes	yes (p = 0.039)	52.36	48.93

Table 11. Discriminant function coefficients for analysis 1.

Character	Function 1	Function 2	Function 3
lbw	-0.0754	0.3029	-0.3079
tlntotlw	-0.5548	3.3830	-0.5060
brwtol	7.1758	7.3968	17.0360
lbl	0.0793	-0.0434	0.0605
lbrwtol	7.7290	-10.6632	6.8826
lfl	-0.0288	-0.0446	-0.0277
lfwtol	4.7198	6.6449	-1.1346
tlen	-0.0189	-0.1929	0.3765
height	-0.0001	0.0073	0.0014
brl	0.0366	0.1020	0.1545
brw	-0.0258	-0.2163	-0.4171
Percentage of variance explained	86.6	7.7	5.7

Table 12. Data on the presence or absence of primary branching (br1). Chi-Squared contingency table test indicated that there was a significant relationship between varietal designation and the presence or absence of primary branching ($p = 0.001$).

Variety	Primary Branching Present	Primary Branching Absent
var. <i>americanum</i>	1	94
var. <i>latifolium</i>	1	63
var. <i>lineare</i>	2	12
var. <i>pectinatum</i>	0	75

Table 13. Data on the presence or absence of secondary branching (br2). Chi-Squared contingency table test indicated that there was a significant relationship between varietal designation and the presence or absence of secondary branching ($p < 0.001$).

Variety	Secondary Branching Present	Secondary Branching Absent
var. <i>americanum</i>	46	49
var. <i>latifolium</i>	30	34
var. <i>lineare</i>	11	3
var. <i>pectinatum</i>	14	61

Table 14. Data on the presence or absence of tertiary branching (br3). Chi-Squared contingency table test indicated that there was a significant relationship between varietal designation and the presence or absence of tertiary branching ($p < 0.001$).

Variety	Tertiary Branching Present	Tertiary Branching Absent
var. <i>americanum</i>	90	5
var. <i>latifolium</i>	61	3
var. <i>lineare</i>	14	0
var. <i>pectinatum</i>	56	19

Table 15. Data on the number of teeth per side of bract (tpers). Chi-Squared contingency table test indicated that there was a significant relationship between varietal designation and the number of bracteal teeth per side ($p = 0.015$).

Variety	Number of teeth per side					
	0	1	2	3	4	5
var. <i>americanum</i>	15	45	25	8	1	1
var. <i>latifolium</i>	31	17	11	4	1	0
var. <i>lineare</i>	7	6	1	0	0	0
var. <i>pectinatum</i>	20	24	23	7	1	0

Table 16. Data on the maximum number of teeth per side of bract (maxt). Chi-Squared contingency table test indicated that there was a significant relationship between varietal designation and the maximum number of bracteal teeth per side ($p < 0.001$).

Variety	Maximum number of teeth per side					
	0	1	2	3	4	5
var. <i>americanum</i>	3	20	47	22	2	1
var. <i>latifolium</i>	12	14	18	18	2	0
var. <i>lineare</i>	1	9	3	1	0	0
var. <i>pectinatum</i>	0	6	25	33	10	1

Table 17. Data on the habitat moisture levels as recorded on herbarium specimen labels. Chi-Squared contingency table test indicated that there was not a significant relationship between varietal designation and habitat moisture levels ($p = 0.121$).

Variety	Dry Habitat	Moist Habitat
var. <i>americanum</i>	12	2
var. <i>latifolium</i>	12	8
var. <i>lineare</i>	6	3
var. <i>pectinatum</i>	17	2

APPENDIX B: FIGURES

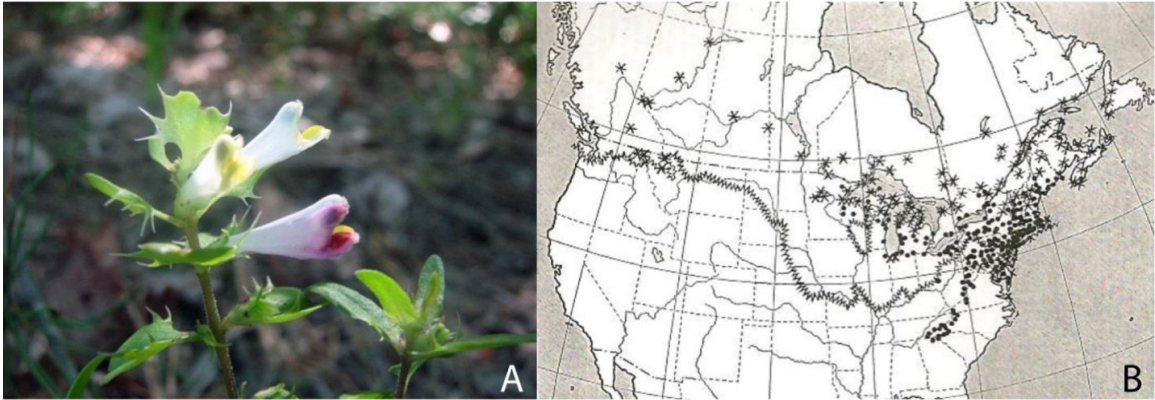


Figure 1. (A) *Melampyrum lineare* in flower, photo courtesy of Renee Brecht, Citizens United: Plants of Southern New Jersey; (B) distribution of *M. lineare* varieties *lineare* (asterisks), *latifolium* (dots) and *pectinatum* (crosses) and the extent of the last Pleistocene glacial maximum (zigzag line)(Pennell, 1935).



Figure 2. (A) *Melampyrum pratense* in flower, Salvan, Switzerland, photo courtesy of Teun Spaans, Wikimedia Commons; (B) *Melampyrum laxum* var. *arcuatum* in flower, photo courtesy of Yahoo Blogs.



Figure 3. Specimen identified in this project as var. *latifolium* from the New York Botanical Garden's herbarium, collected in moist woodlands near Biltmore, North Carolina on May 29, 1897.



Figure 4. Specimen identified in this study as var. *americanum* from the New York Botanical Garden's herbarium, collected in dry woodlands near Biltmore, North Carolina on August 20, 1897.

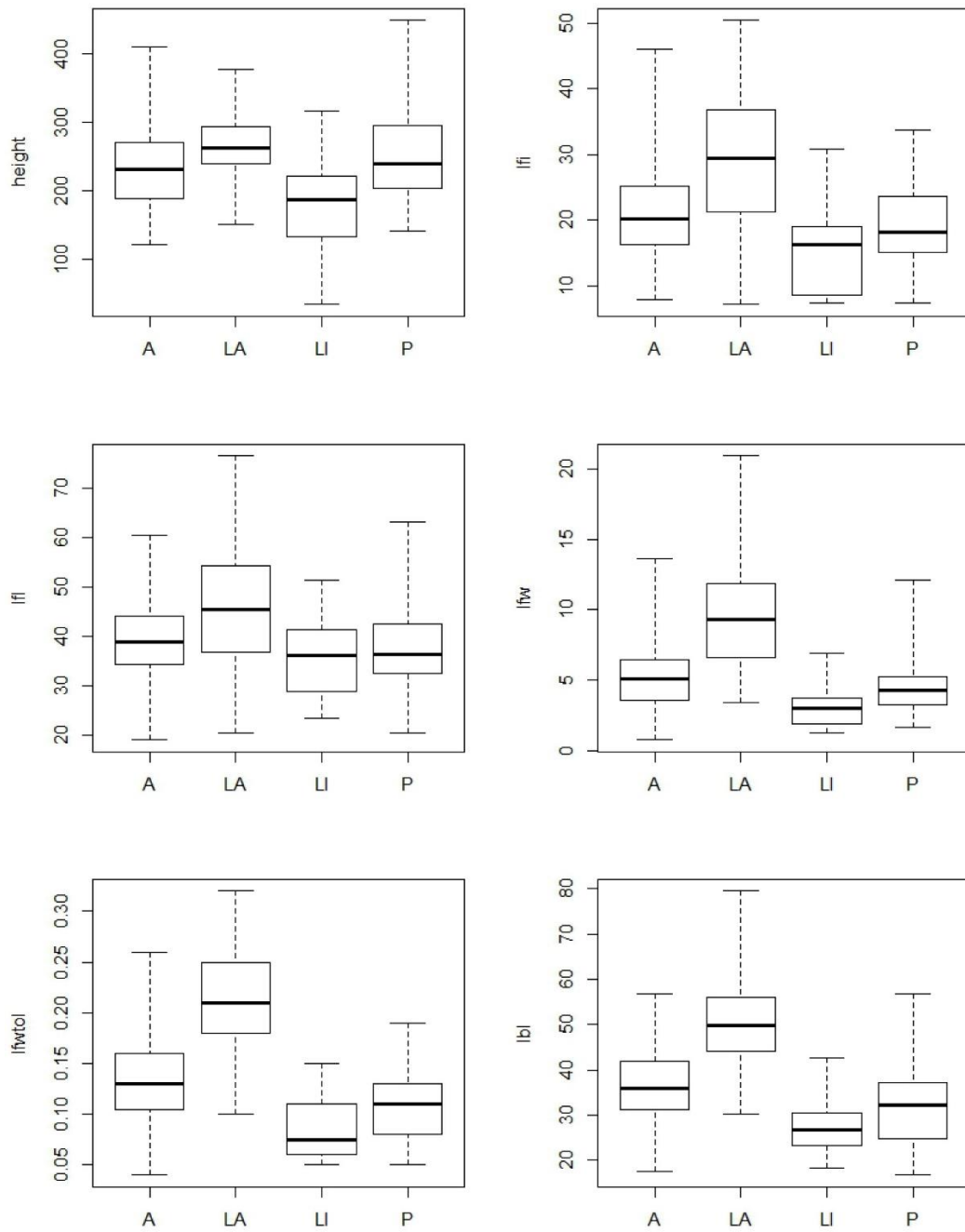


Figure 5. From left to right, top to bottom: boxplots for the continuous variables height, lfi, lfl, lfw, lfwtol, and lbl. On the x-axis, “A” represents var. *americanum*, “LA” represents var. *latifolium*, “LI” represents var. *lineare*, and “P” represents var. *pectinatum*.

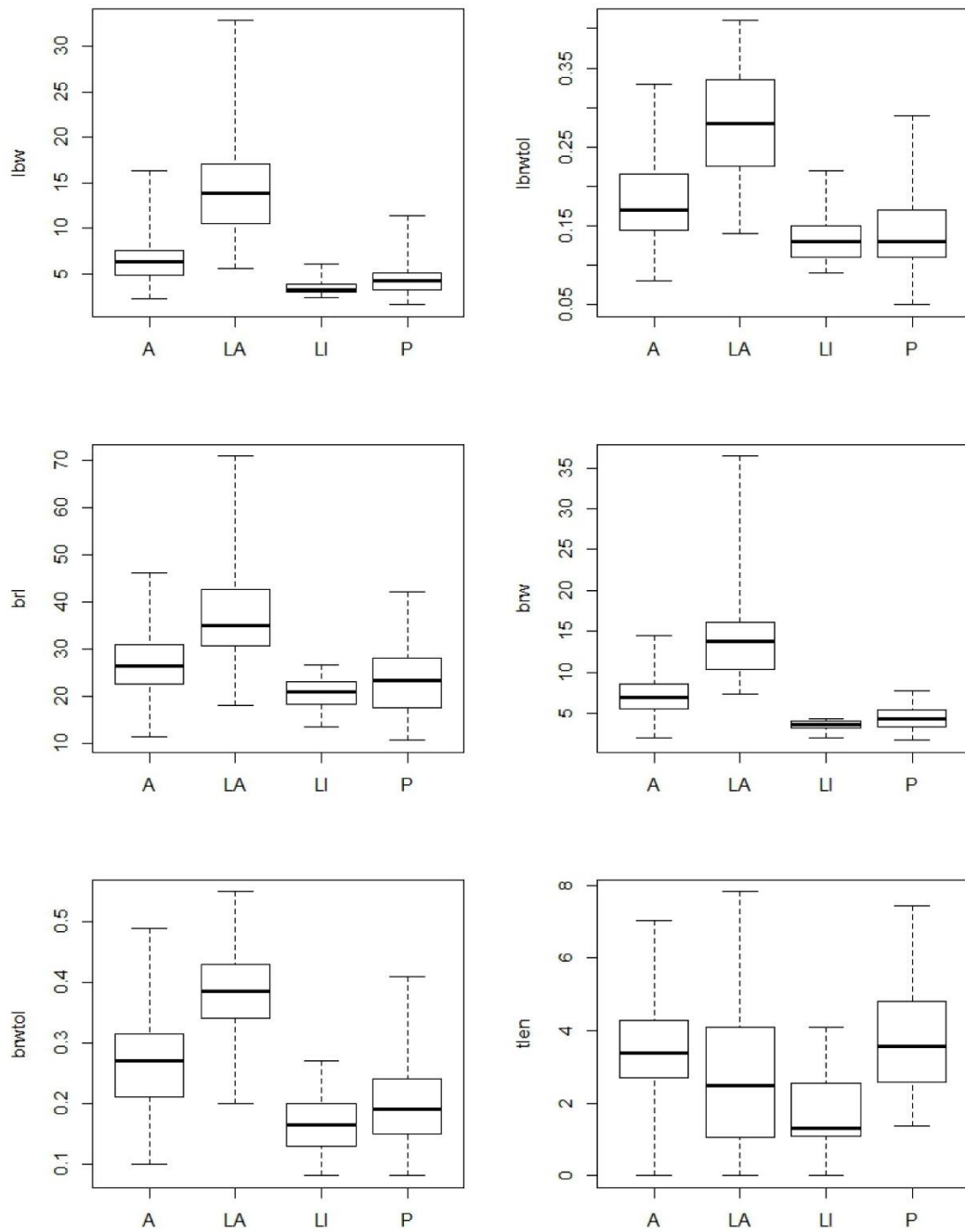


Figure 6. From left to right, top to bottom: boxplots for the continuous variables lbw, lbrwtol, brl, brwtol, tlen. On the x-axis, “A” represents var. *americanum*, “LA” represents var. *latifolium*, “LI” represents var. *lineare*, and “P” represents var. *pectinatum*.

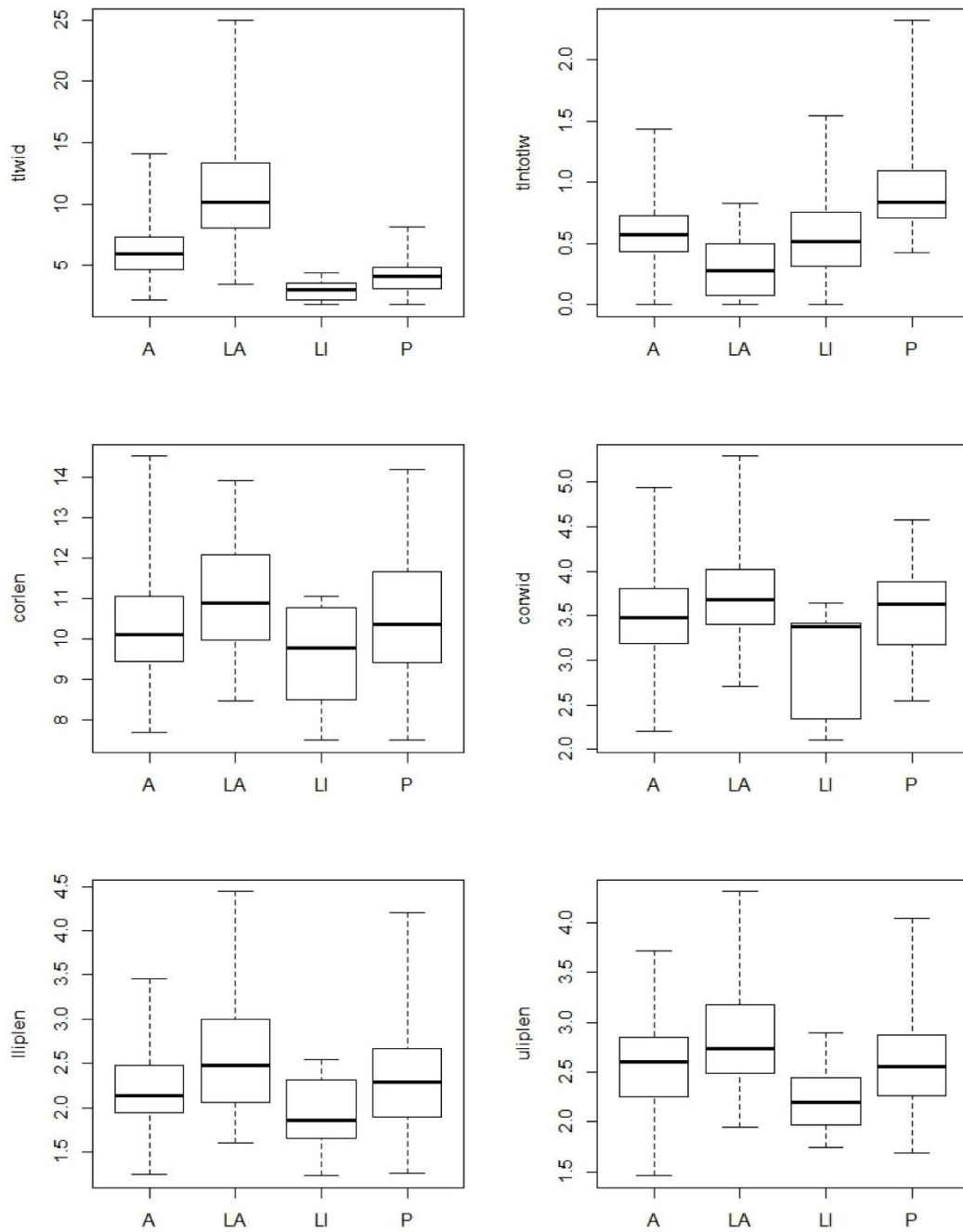


Figure 7. From left to right, top to bottom: boxplots for the continuous variables tlwid, tlntotlw, corlen, corwid, lliplen, uliplen. On the x-axis, “A” represents var. *americanum*, “LA” represents var. *latifolium*, “LI” represents var. *lineare*, and “P” represents var. *pectinatum*.

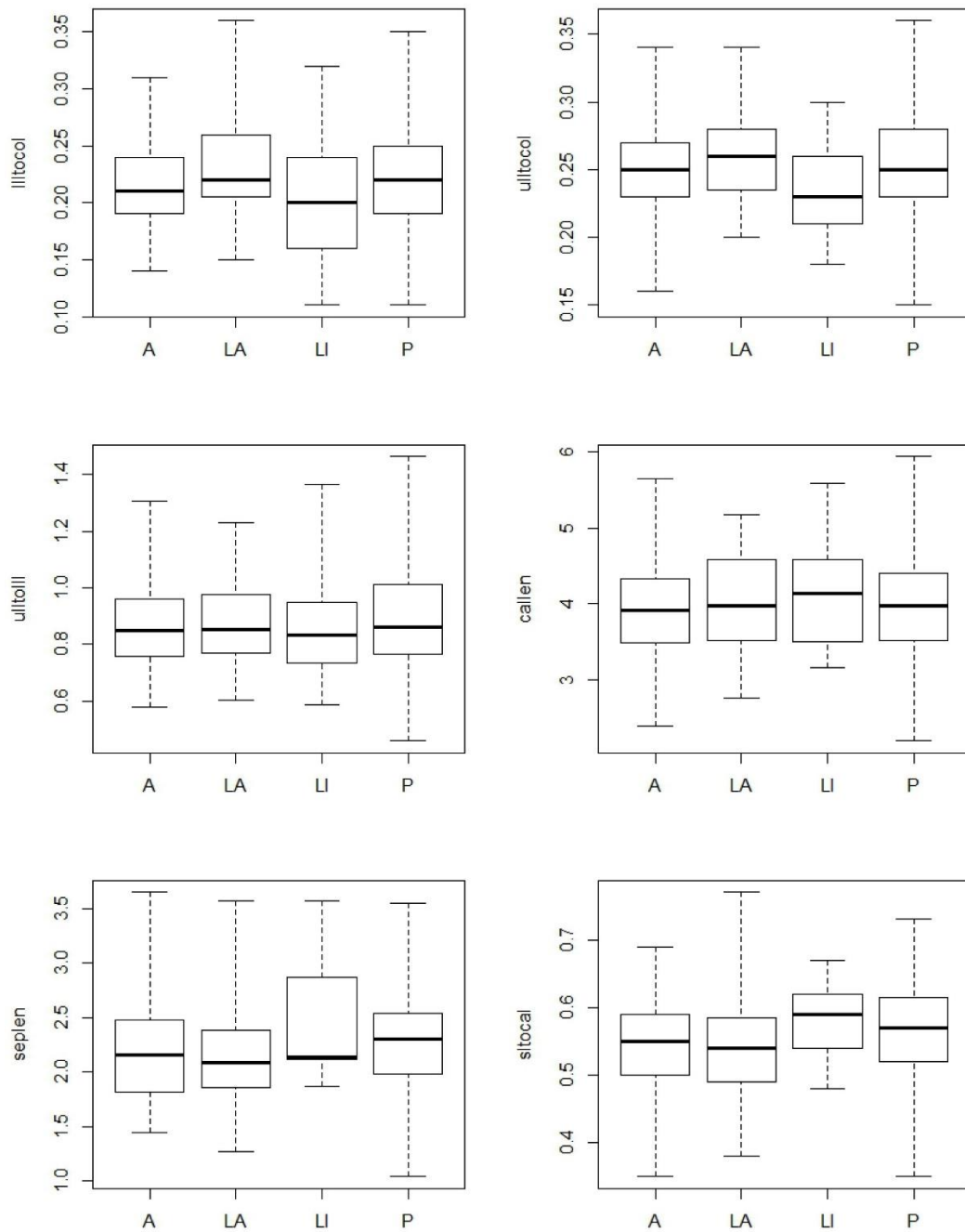


Figure 8. From left to right, top to bottom: boxplots for the continuous variables llltocol, ulltocol, ulltolll, callen, seplen, and sltocal, all measured in millimeters. On the x-axis, “A” represents var. *americanum*, “LA” represents var. *latifolium*, “LI” represents var. *lineare*, and “P” represents var. *pectinatum*.

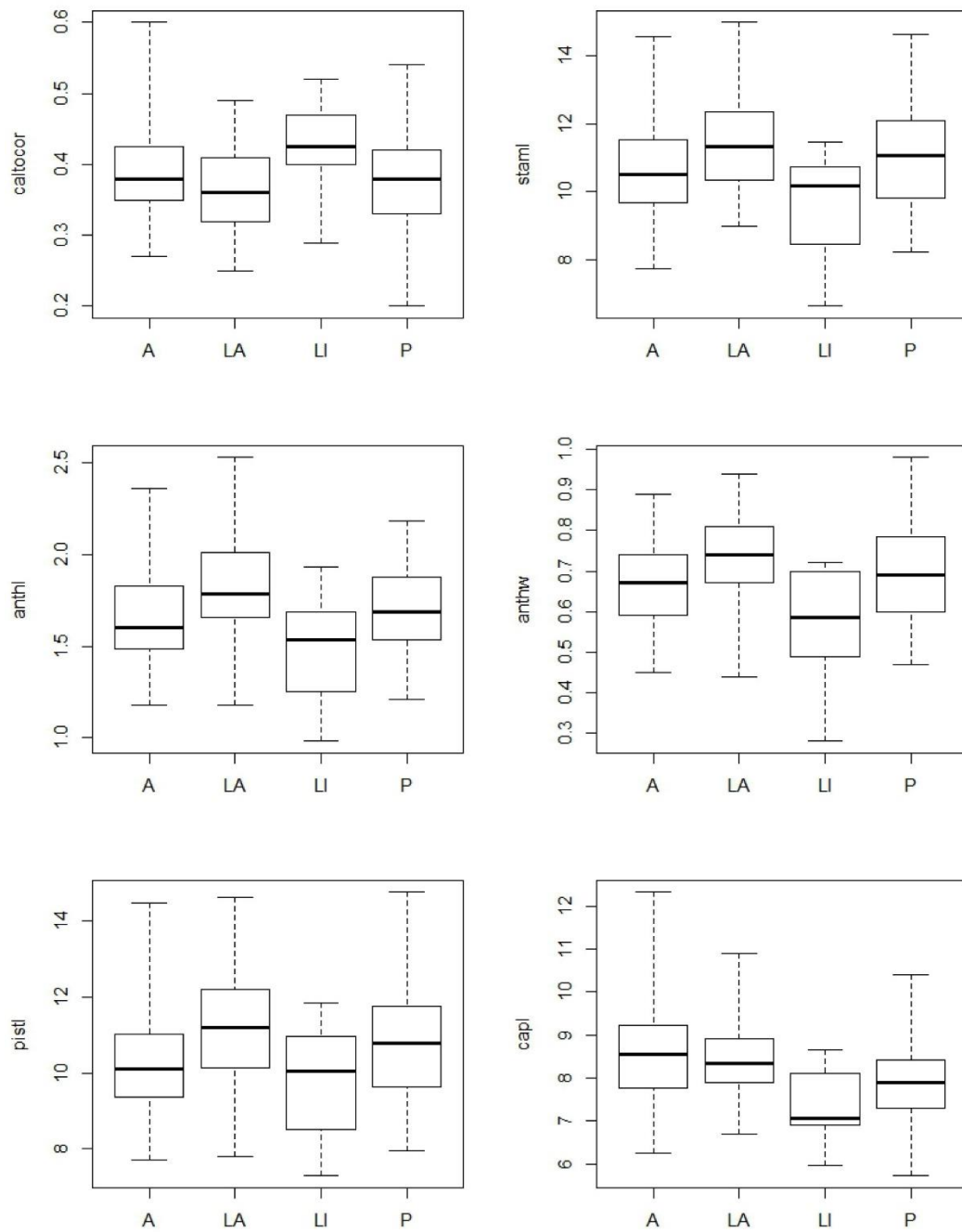


Figure 9. From left to right, top to bottom: boxplots for the continuous variables caltocor, stam1, anth1, anthw, pist1, capl. On the x-axis, “A” represents var. *americanum*, “LA” represents var. *latifolium*, “LI” represents var. *lineare*, and “P” represents var. *pectinatum*.

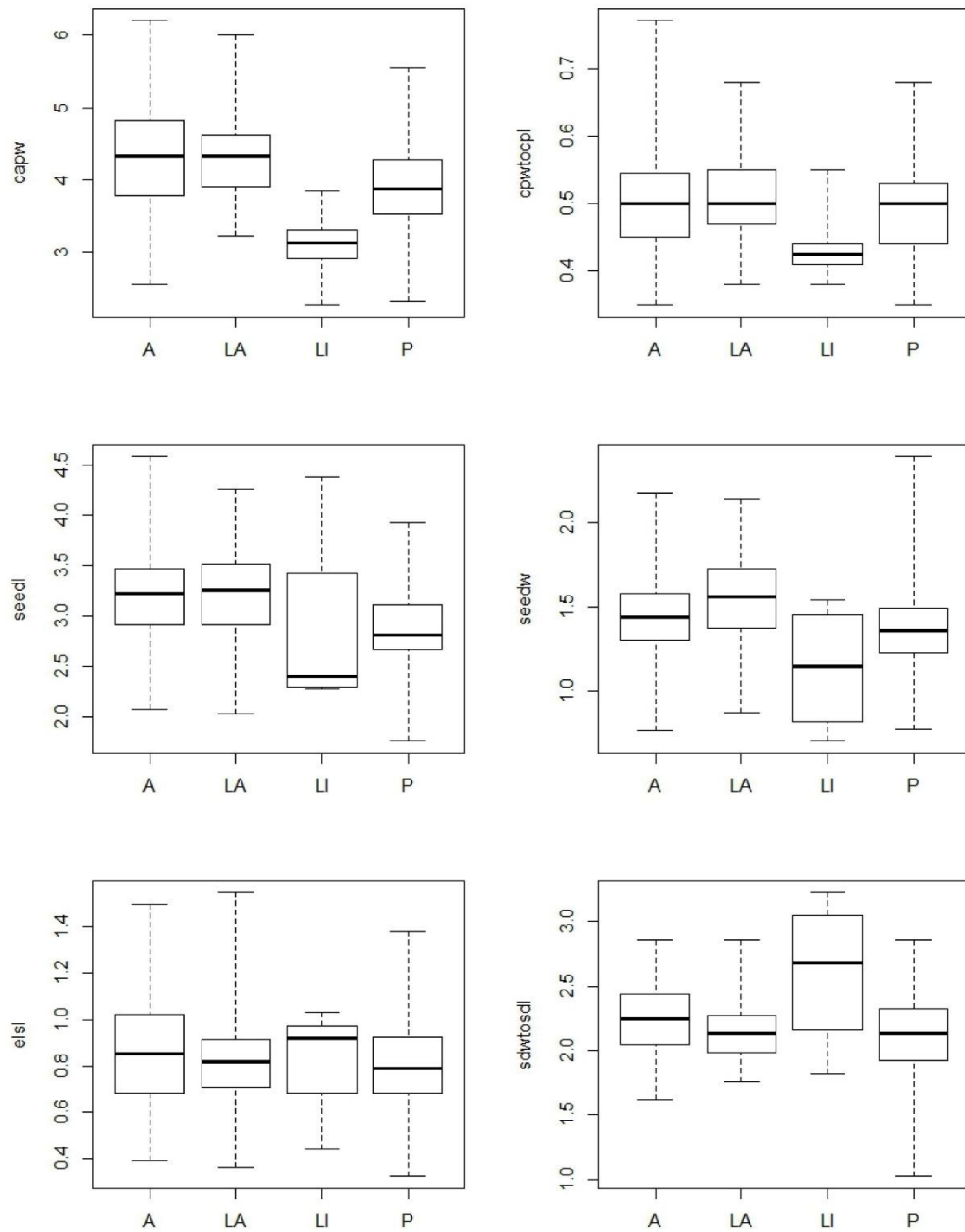


Figure 10. From left to right, top to bottom: boxplots for the continuous variables capw, cpwtocpl, seedl, seedw, elsl, and sdwtosdl. On the x-axis, “A” represents var. *americanum*, “LA” represents var. *latifolium*, “LI” represents var. *lineare*, and “P” represents var. *pectinatum*.

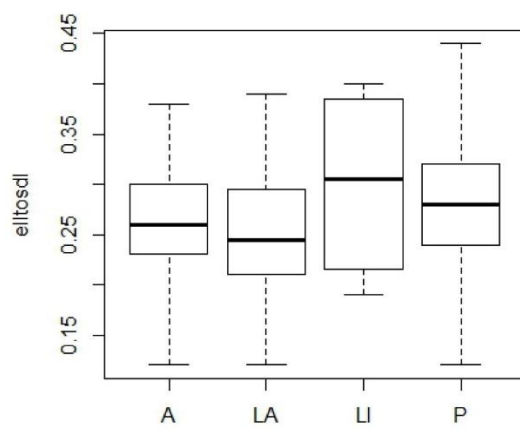


Figure 11. Boxplot for the continuous variable elltosdl. On the x-axis, "A" represents var. *americanum*, "LA" represents var. *latifolium*, "LI" represents var. *lineare*, and "P" represents var. *pectinatum*.

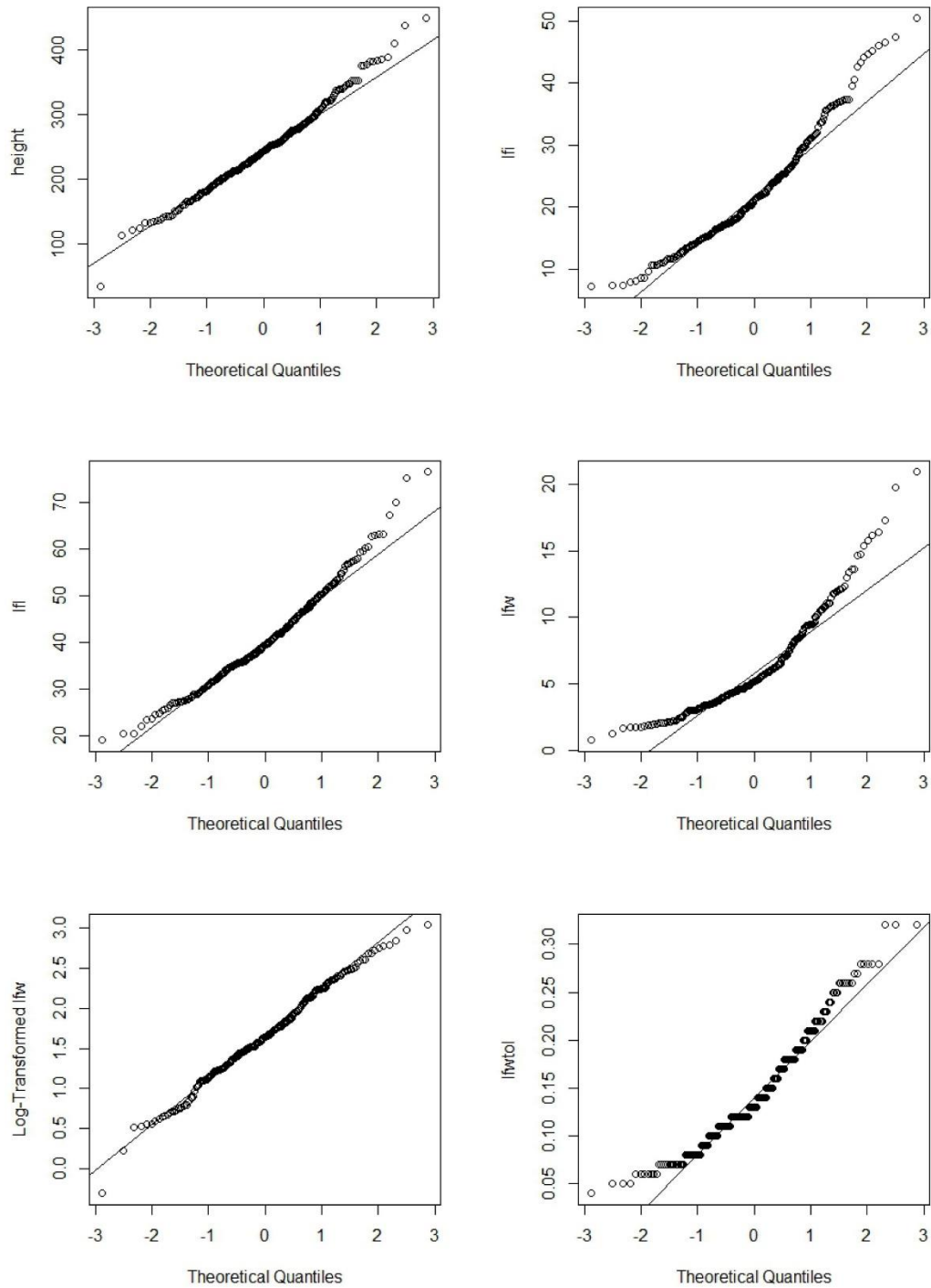


Figure 12. From left to right, top to bottom: univariate Q-Q plots for the continuous variables height, lfi, lfl, lfw, log-transformed lfw, and lfwtol. The diagonal line on each plot represents an ideal normal distribution.

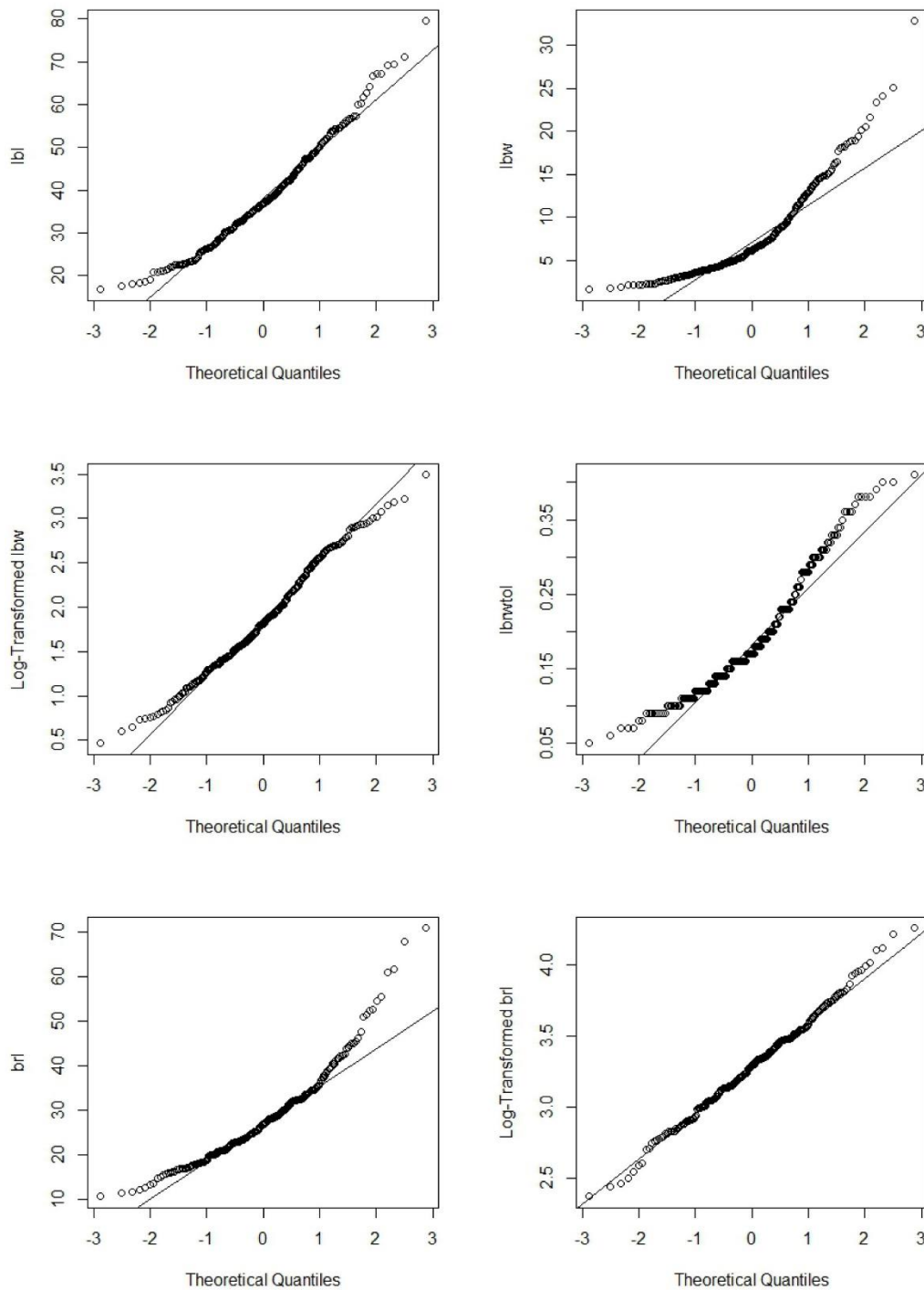


Figure 13. From left to right, top to bottom: univariate Q-Q plots for the continuous variables lbl, lbw, log-transformed lbw, lbwrtol, brl, and log-transformed brl. The diagonal line on each plot represents an ideal normal distribution.

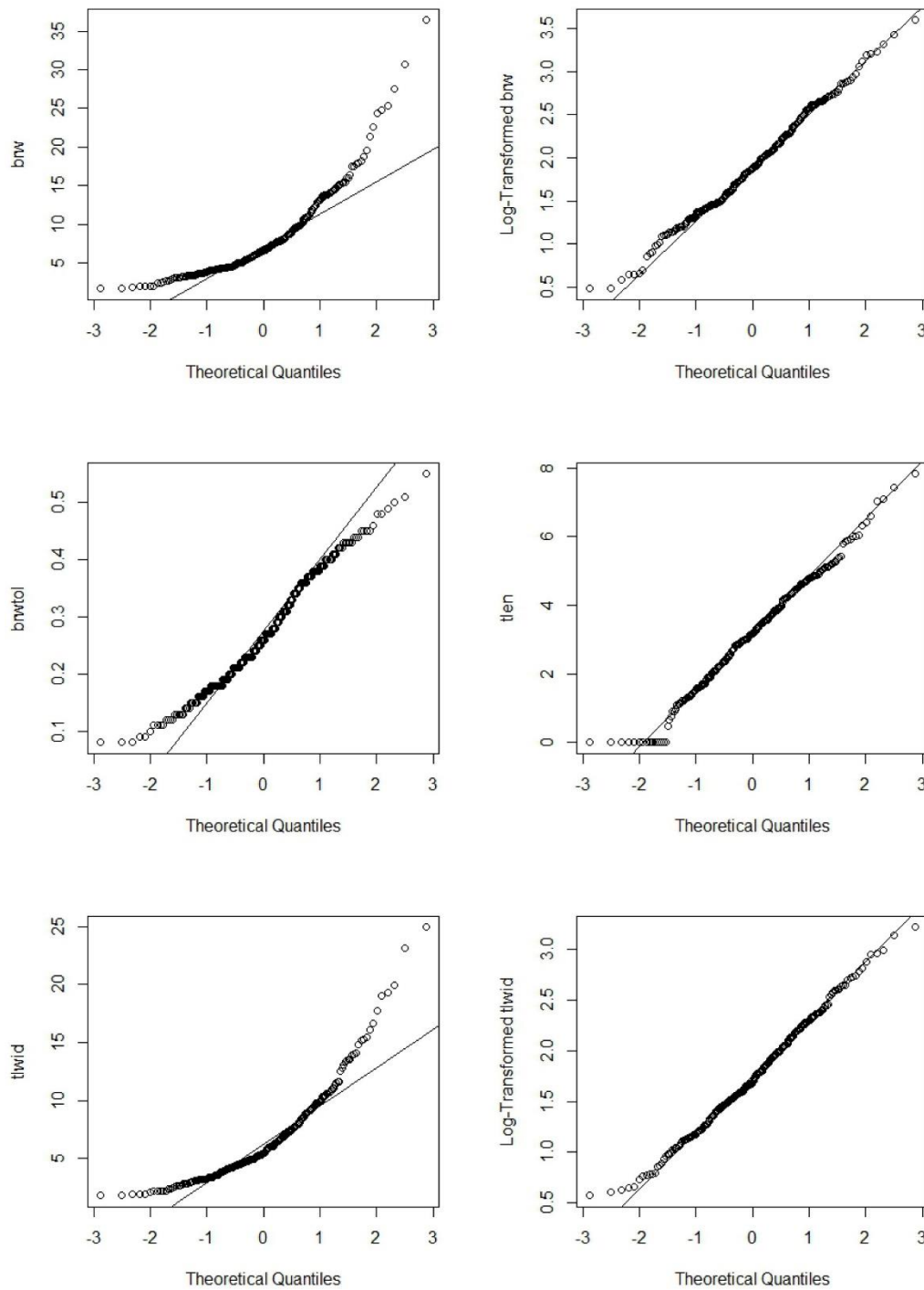


Figure 14. From left to right, top to bottom: univariate Q-Q plots for the continuous variables *brw*, log-transformed *brw*, *brwtol*, *tlen*, *tlwid*, and log-transformed *tlwid*. The diagonal line on each plot represents an ideal normal distribution.

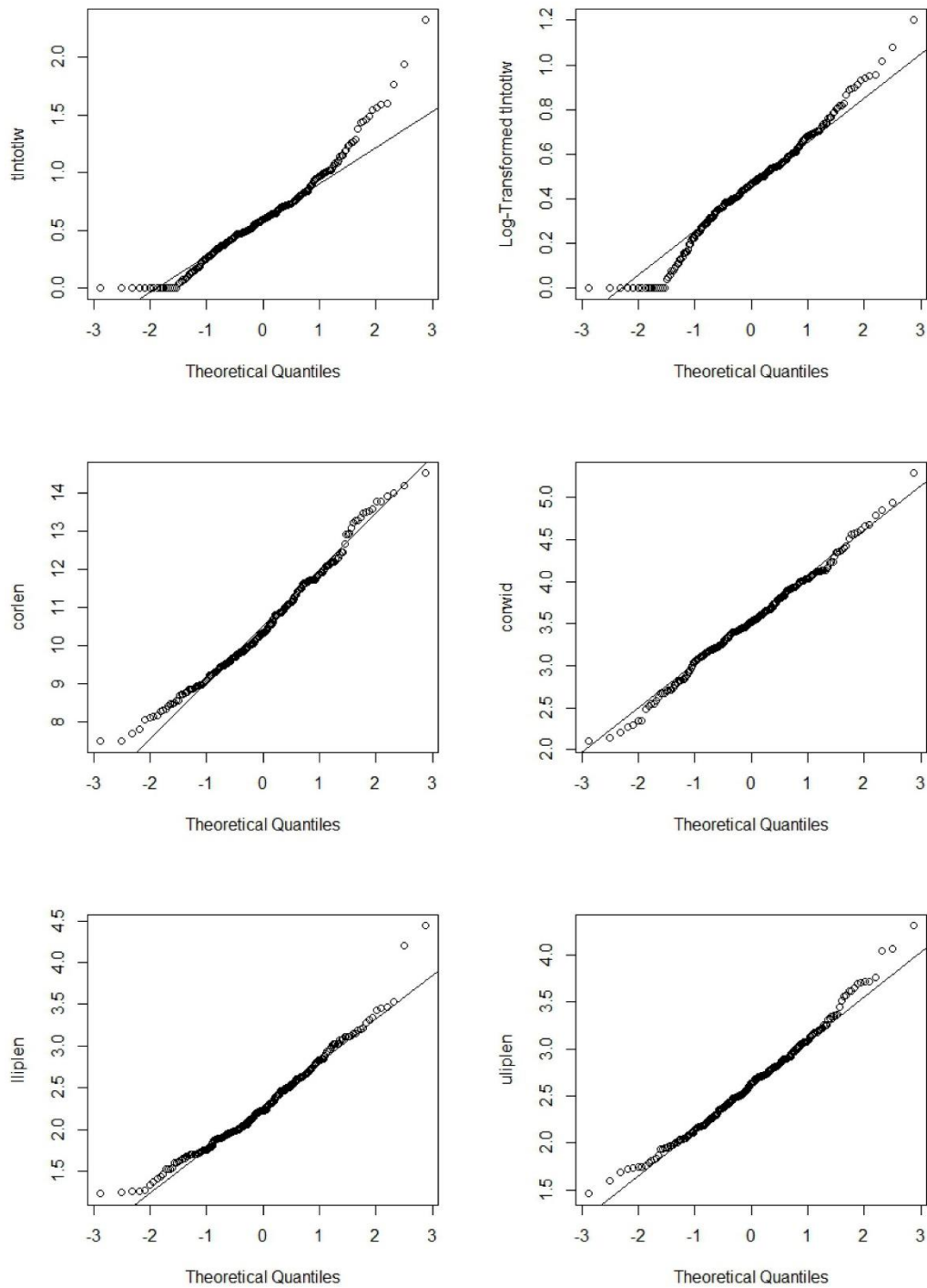


Figure 15. From left to right, top to bottom: univariate Q-Q plots for the continuous variables tntotlw, log-transformed tntotlw, corlen, corwid, lilplen, and uliplen. The diagonal line on each plot represents an ideal normal distribution.

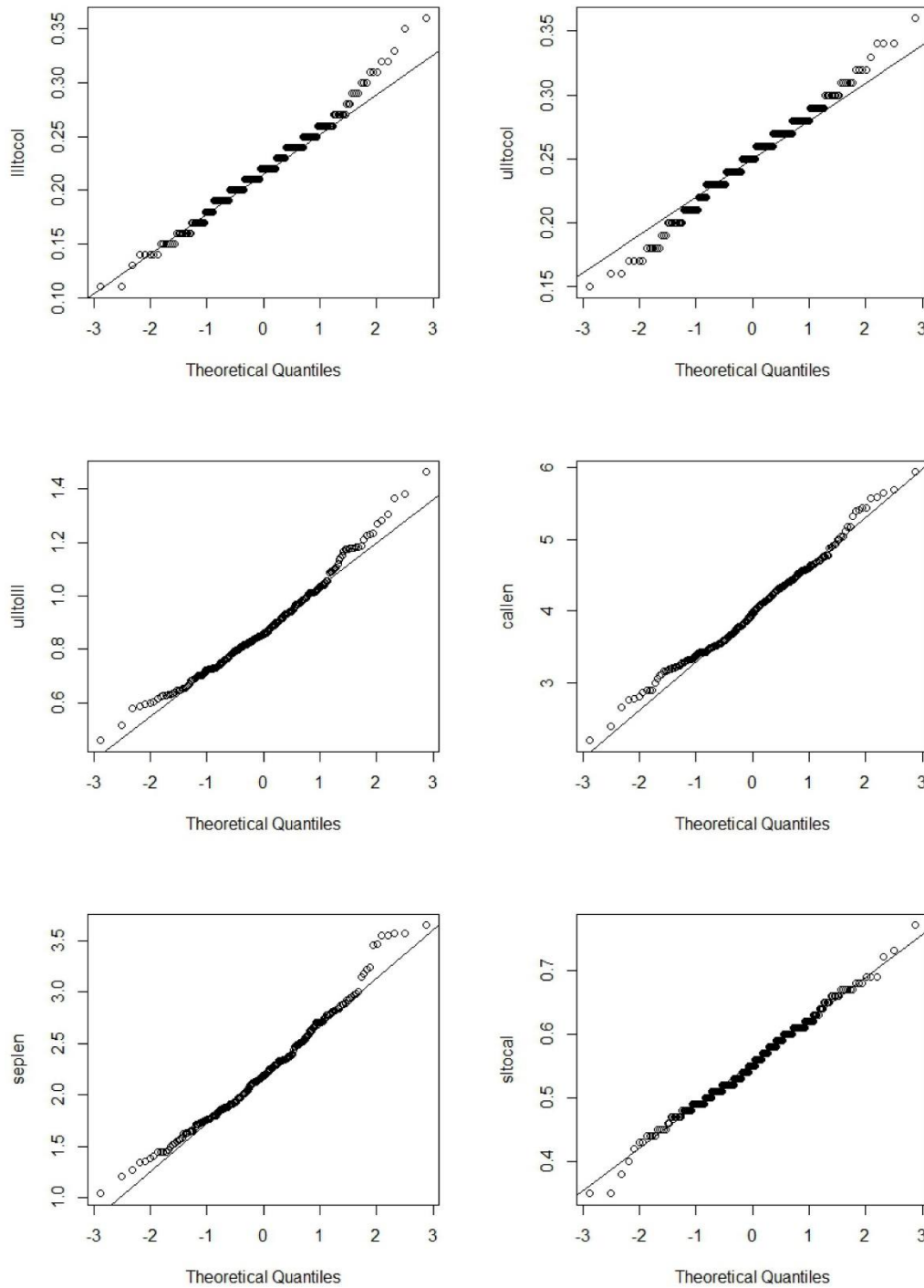


Figure 16. From left to right, top to bottom: univariate Q-Q plots for the continuous variables *lltocol*, *ulltocol*, *ulltolll*, *callen*, *seplen*, and *sltocal*. The diagonal line on each plot represents an ideal normal distribution.

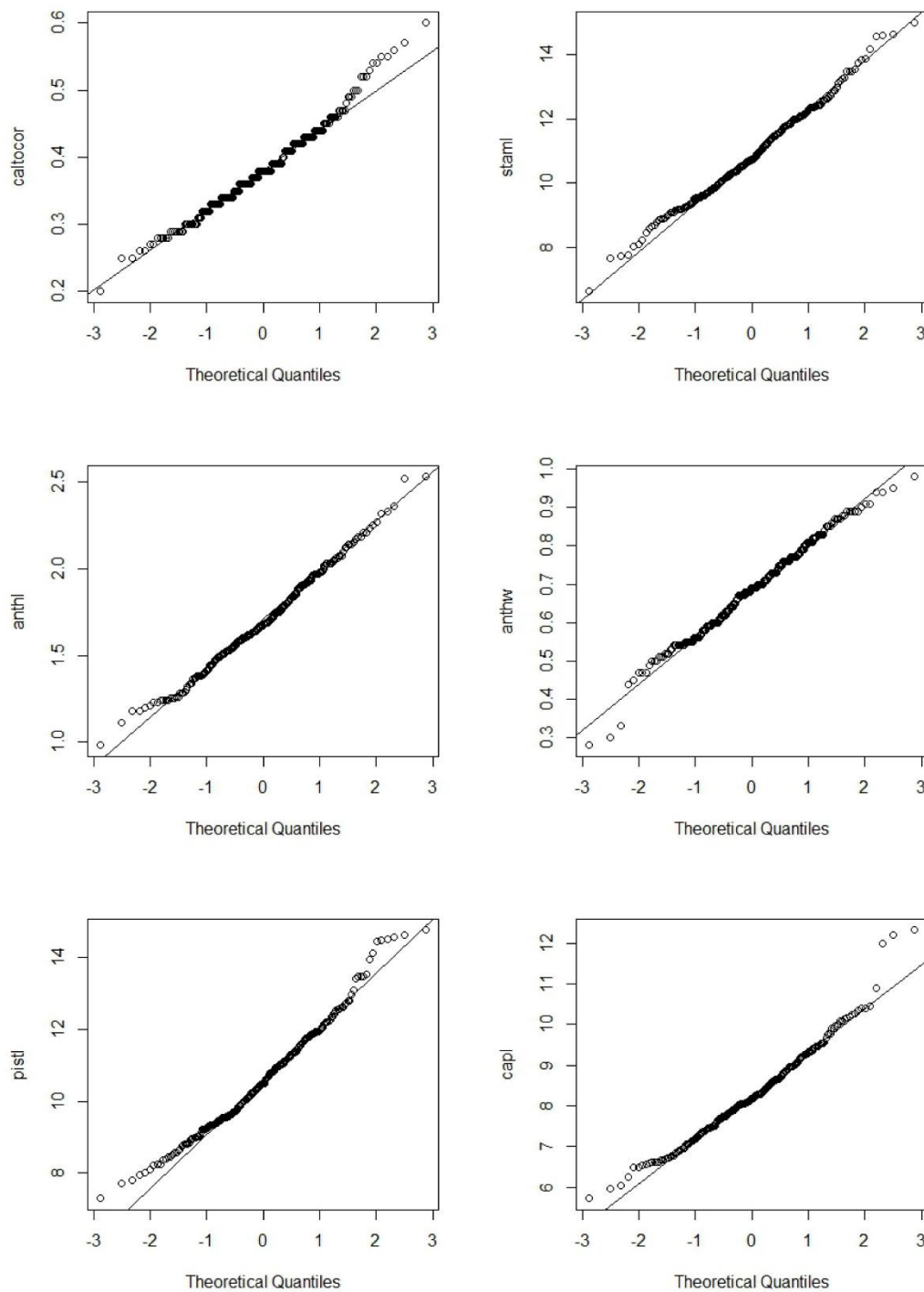


Figure 17. From left to right, top to bottom: univariate Q-Q plots for the continuous variables caltcor, staml, anthl, anthw, pistl, and capl. The diagonal line on each plot represents an ideal normal distribution.

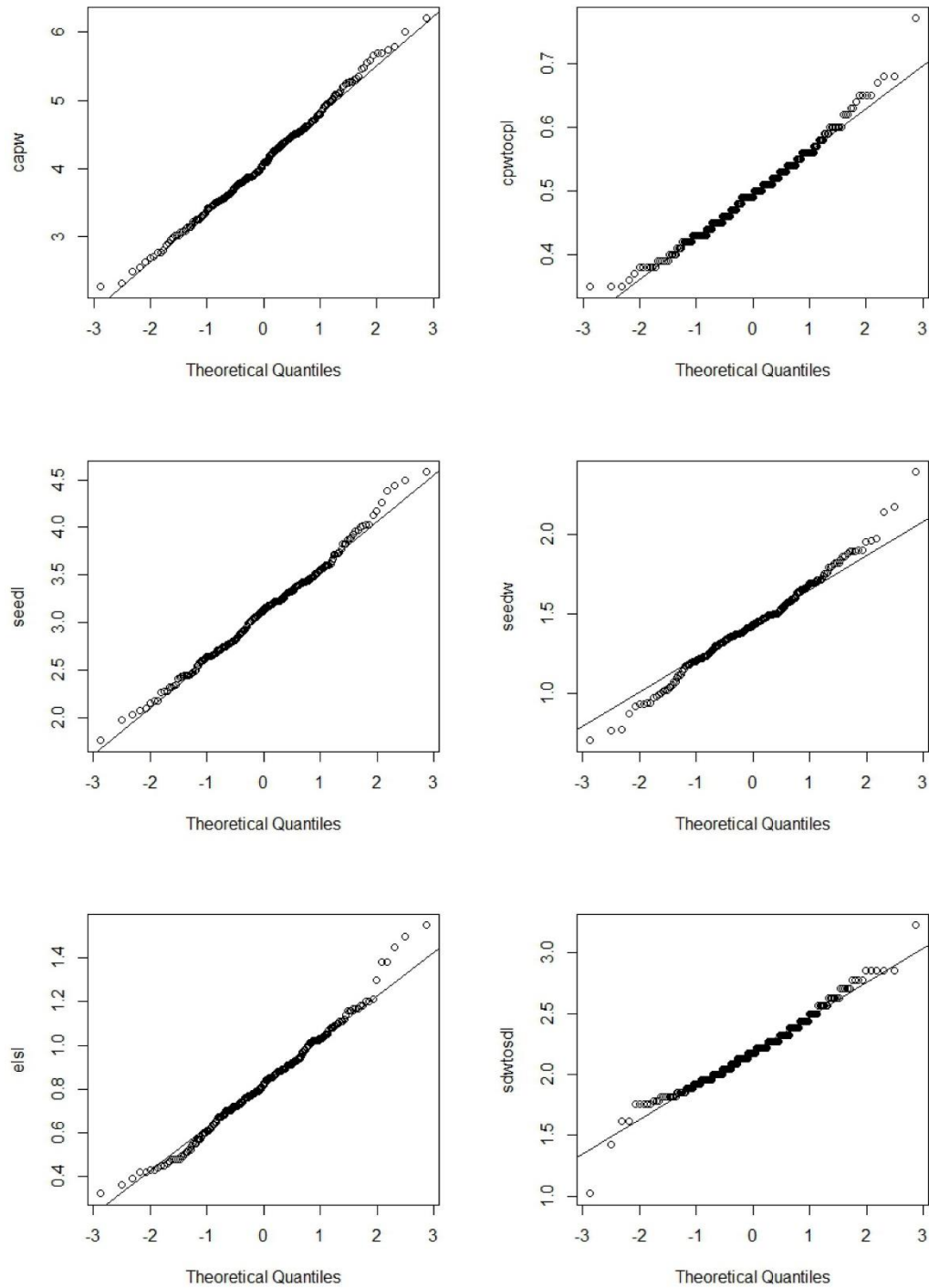


Figure 18. From left to right, top to bottom: univariate Q-Q plots for the continuous variables *capw*, *cpwtocpl*, *seedl*, *seedw*, *elsl*, and *sdwtosdl*. The diagonal line on each plot represents an ideal normal distribution.

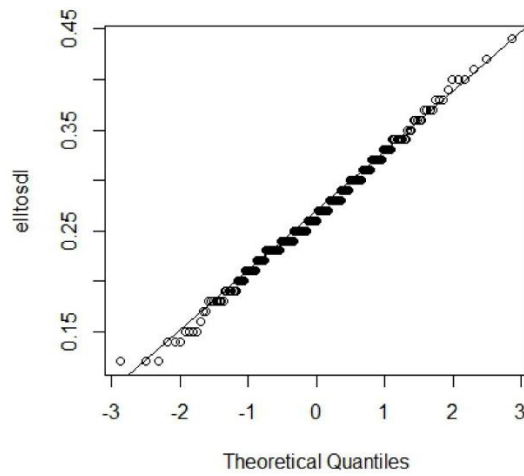


Figure 19. Univariate Q-Q plot for the continuous variable elltosdl. The diagonal line represents an ideal normal distribution.

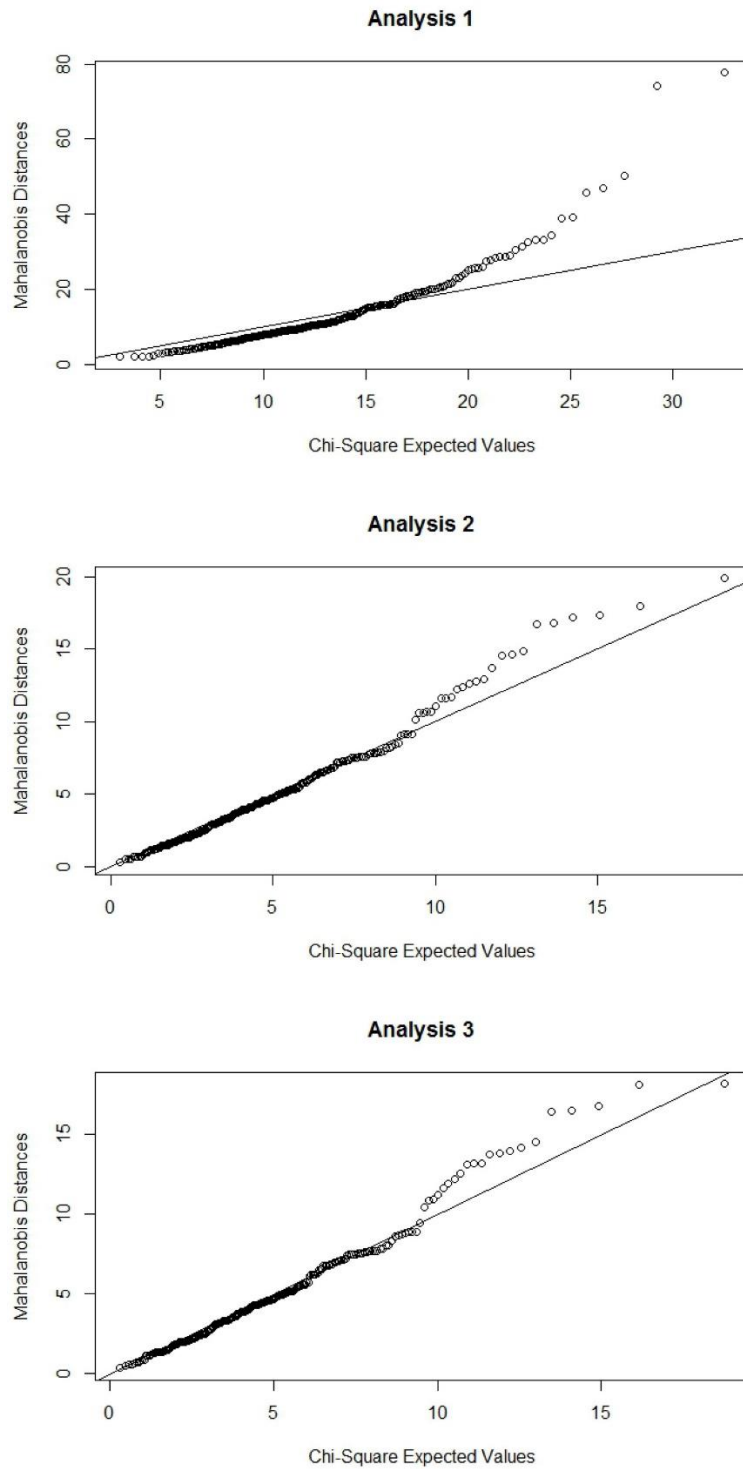


Figure 20. Multivariate Q-Q plots for MANOVA and LDA all vegetative analyses (1-3 in Table 10), with Mahalanobis distances plotted against Chi-Squared expected values. The diagonal line on each plot represents an ideal multivariate normal distribution.

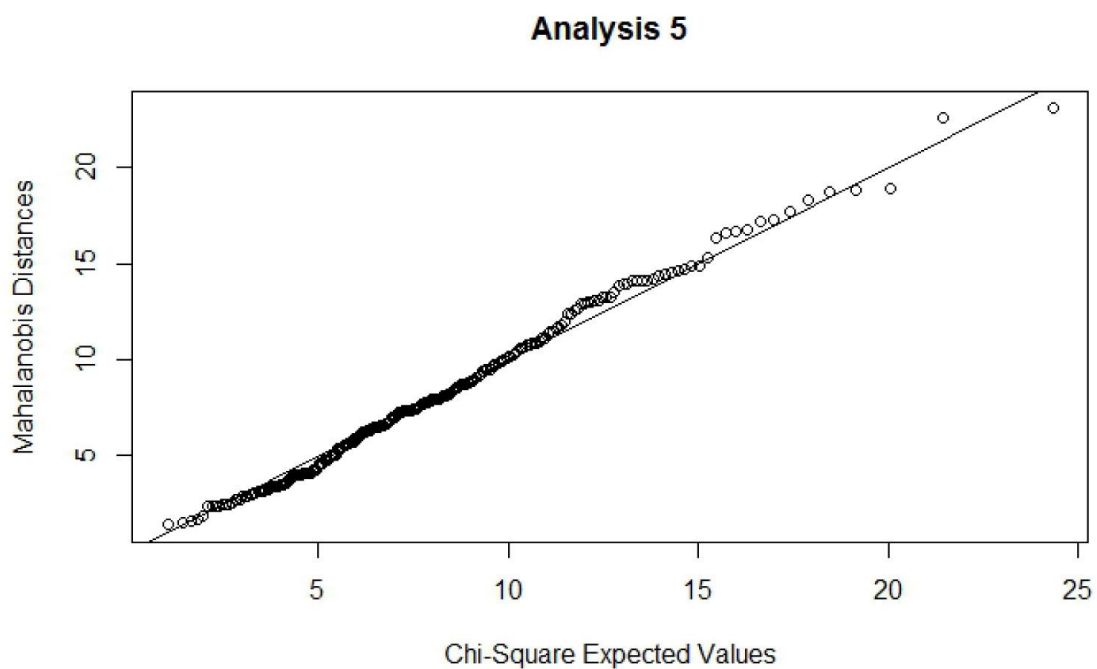
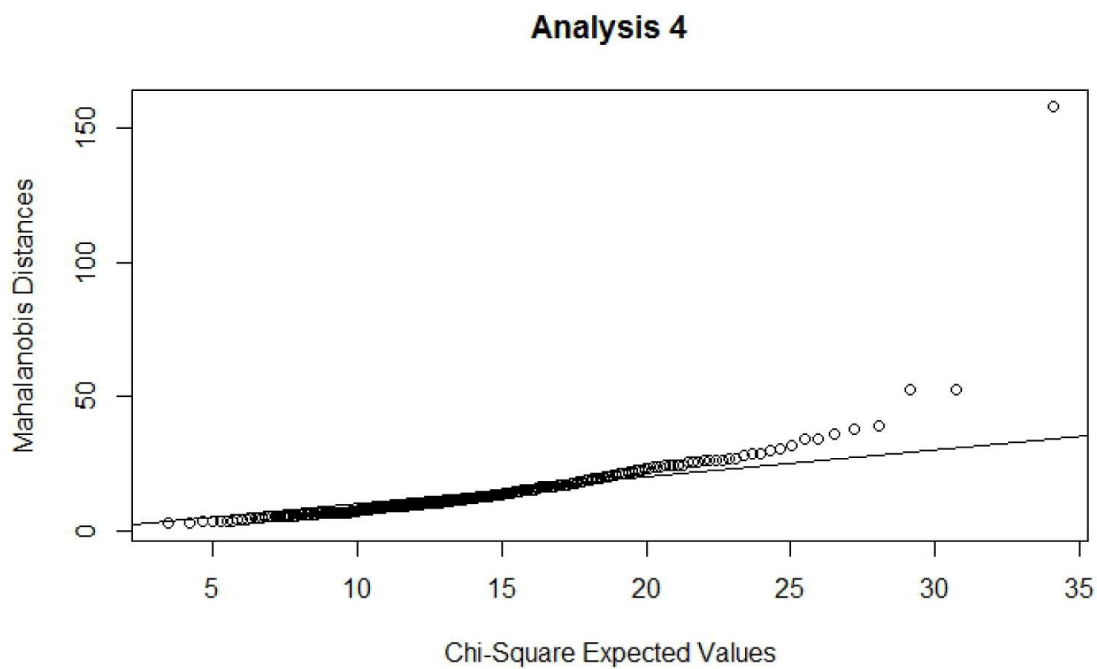


Figure 21. Multivariate Q-Q plots for MANOVA and LDA for all floral analyses (4-5 in Table 10), with Mahalanobis distances plotted against Chi-Squared expected values. The diagonal line on each plot represents an ideal multivariate normal distribution.

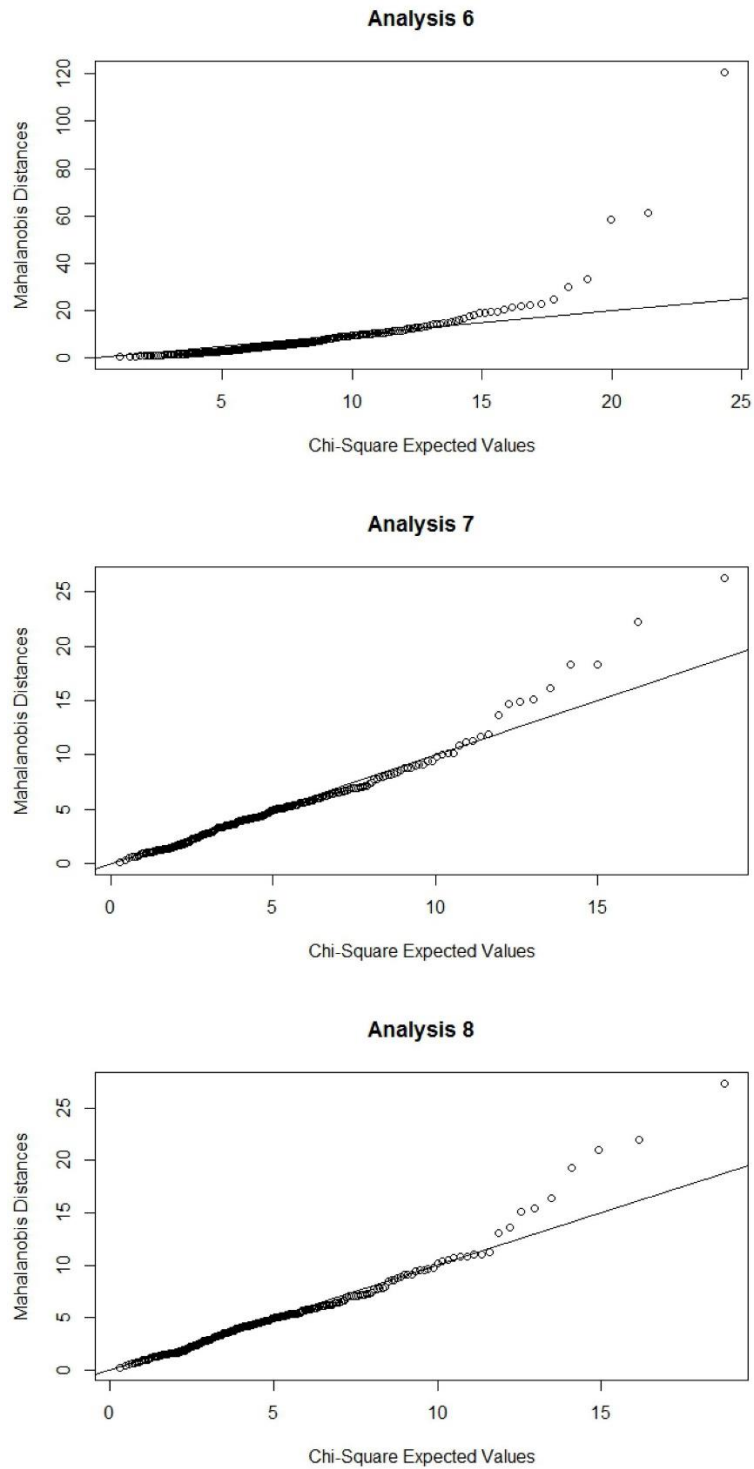


Figure 22. Multivariate Q-Q plots for MANOVA and LDA for all fruit/seed analyses (6-8 in Table 10), with Mahalanobis distances plotted against Chi-Squared expected values. The diagonal line on each plot represents an ideal multivariate normal distribution.

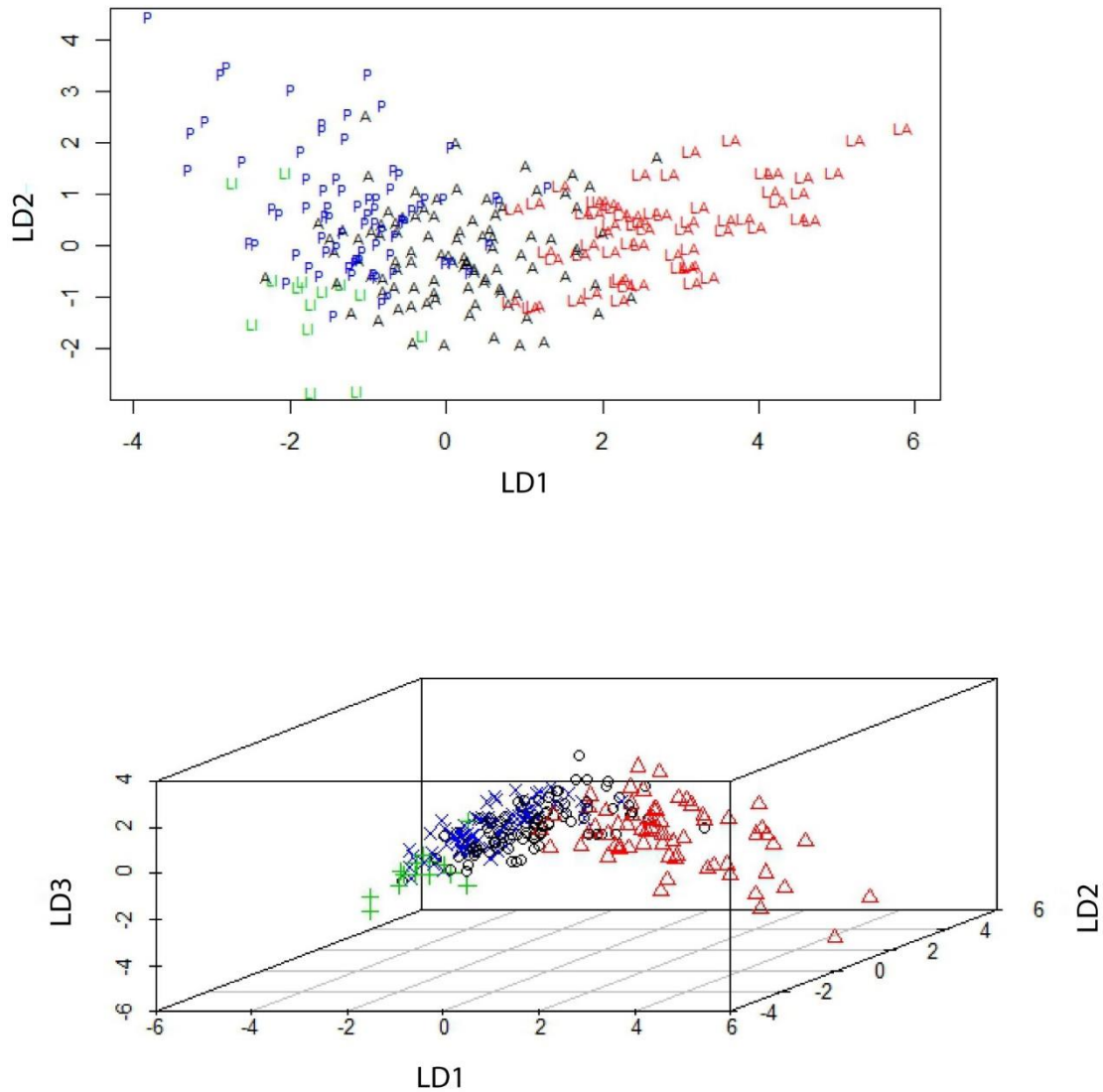


Figure 23. Two- and three-dimensional discriminant function scatterplots for analysis 1. Discriminant function 1 is plotted on the x-axis (LD1), discriminant function 2 is plotted on the y-axis (LD2), and discriminant function 3 is plotted on the z-axis (LD3). Var. *americanum* is represented as “A” on the 2-D plot and circles on the 3-D plot; var. *latifolium* is represented as “LA” on the 2-D plot and triangles on the 3-D plot; var. *lineare* is represented as “L” on the 2-D plot and crosses on the 3-D plot; var. *pectinatum* is represented as “P” on the 2-D plot and X’s on the 3-D plot.

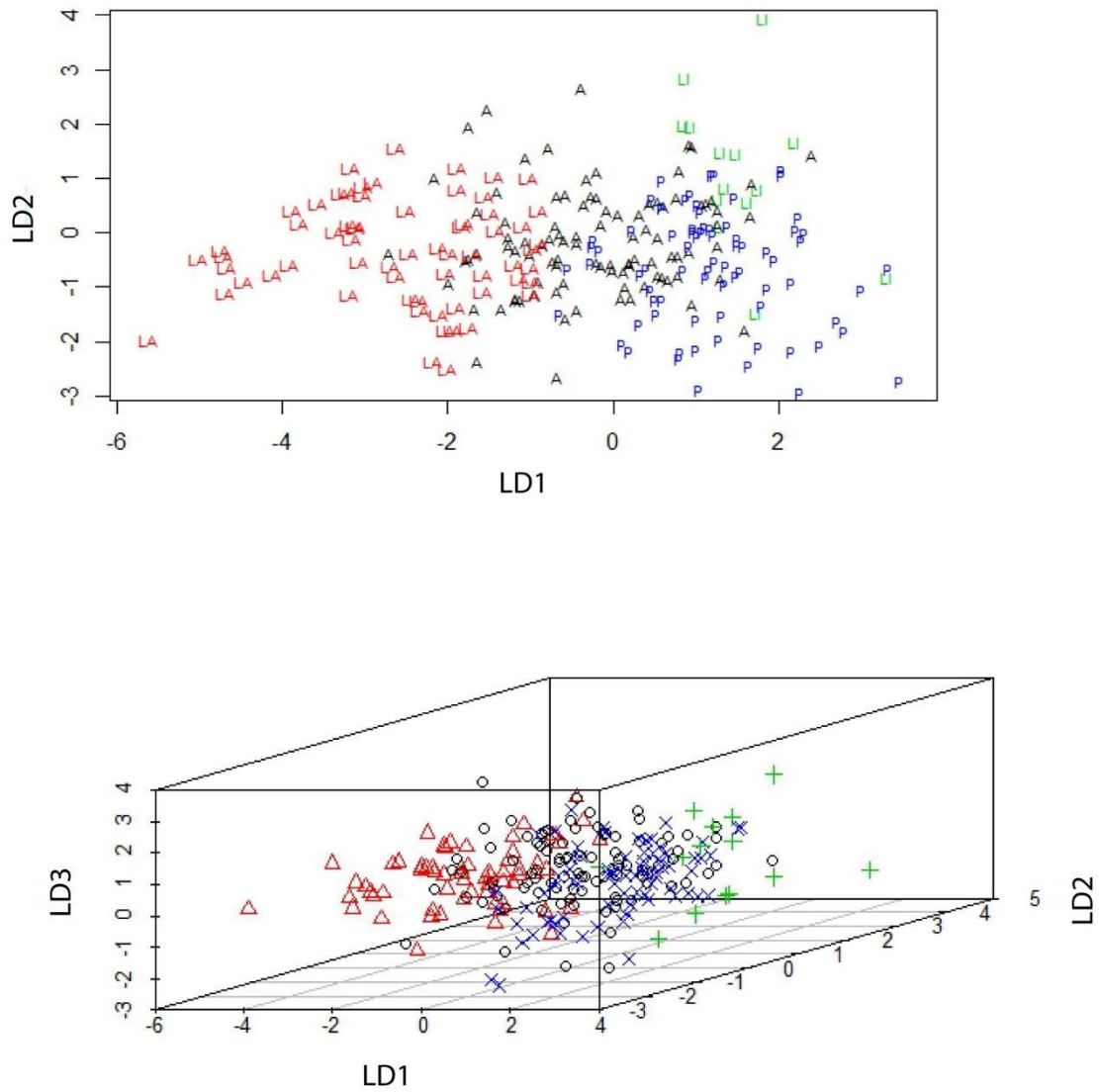


Figure 24. Two- and three-dimensional discriminant function scatterplots for analysis 2. Discriminant function 1 is plotted on the x-axis (LD1), discriminant function 2 is plotted on the y-axis (LD2), and discriminant function 3 is plotted on the z-axis (LD3). Var. *americanum* is represented as “A” on the 2-D plot and circles on the 3-D plot; var. *latifolium* is represented as “LA” on the 2-D plot and triangles on the 3-D plot; var. *lineare* is represented as “LI” on the 2-D plot and crosses on the 3-D plot; var. *pectinatum* is represented as “P” on the 2-D plot and X’s on the 3-D plot.

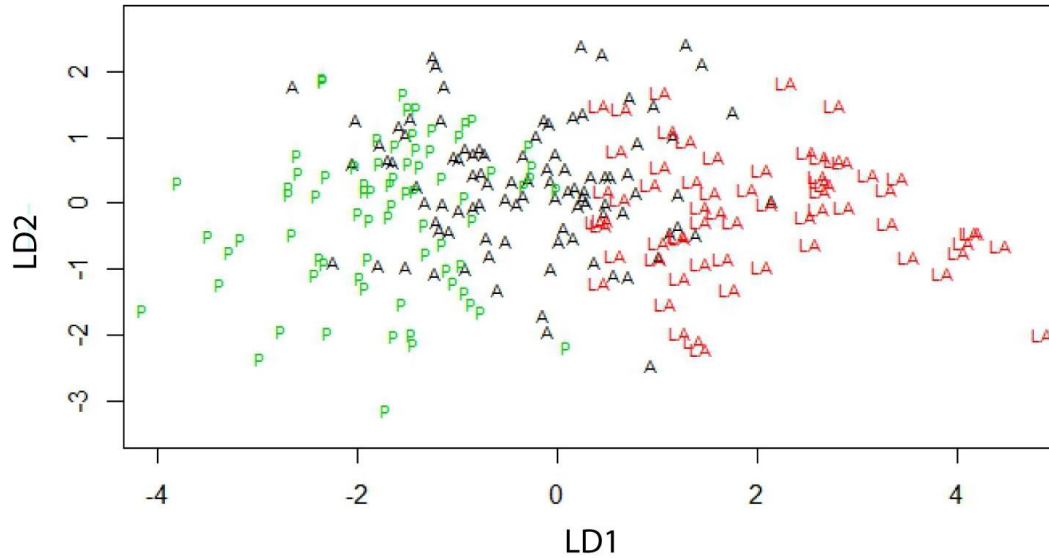


Figure 25. Two-dimensional discriminant function scatterplot for analysis 3. Discriminant function 1 is plotted on the x-axis (LD1), and discriminant function 2 is plotted on the y-axis (LD2). “A” represents var. *americanum*, “LA” represents var. *latifolium*, and “P” represents var. *pectinatum*. Var. *lineare* was excluded from this analysis, so a three-dimensional plot could not be generated.

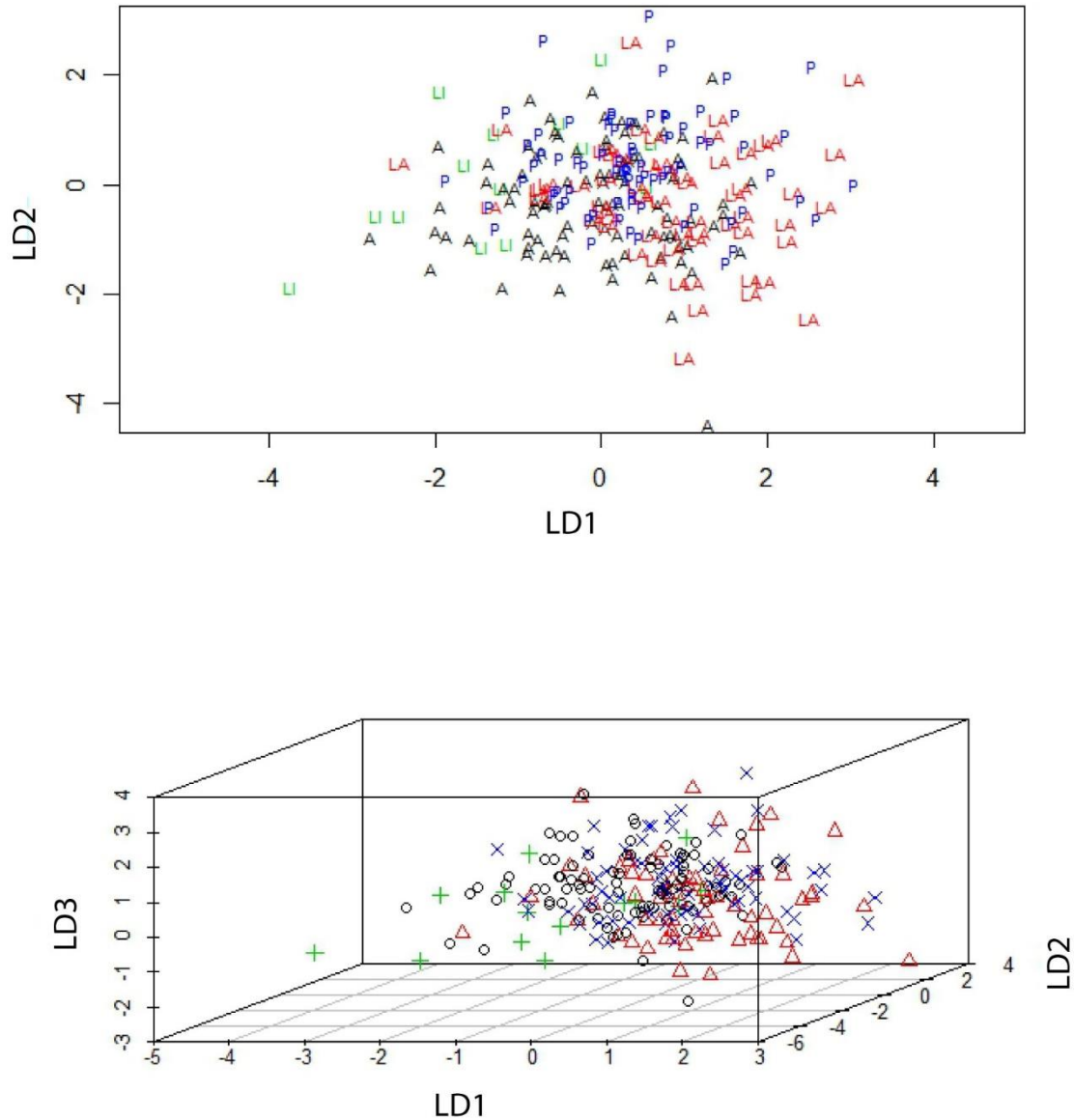


Figure 26. Two- and three-dimensional discriminant function scatterplots for analysis 4. Discriminant function 1 is plotted on the x-axis (LD1), discriminant function 2 is plotted on the y-axis (LD2), and discriminant function 3 is plotted on the z-axis (LD3). Var. *americanum* is represented as “A” on the 2-D plot and circles on the 3-D plot; var. *latifolium* is represented as “LA” on the 2-D plot and triangles on the 3-D plot; var. *lineare* is represented as “LI” on the 2-D plot and crosses on the 3-D plot; var. *pectinatum* is represented as “P” on the 2-D plot and X’s on the 3-D plot.

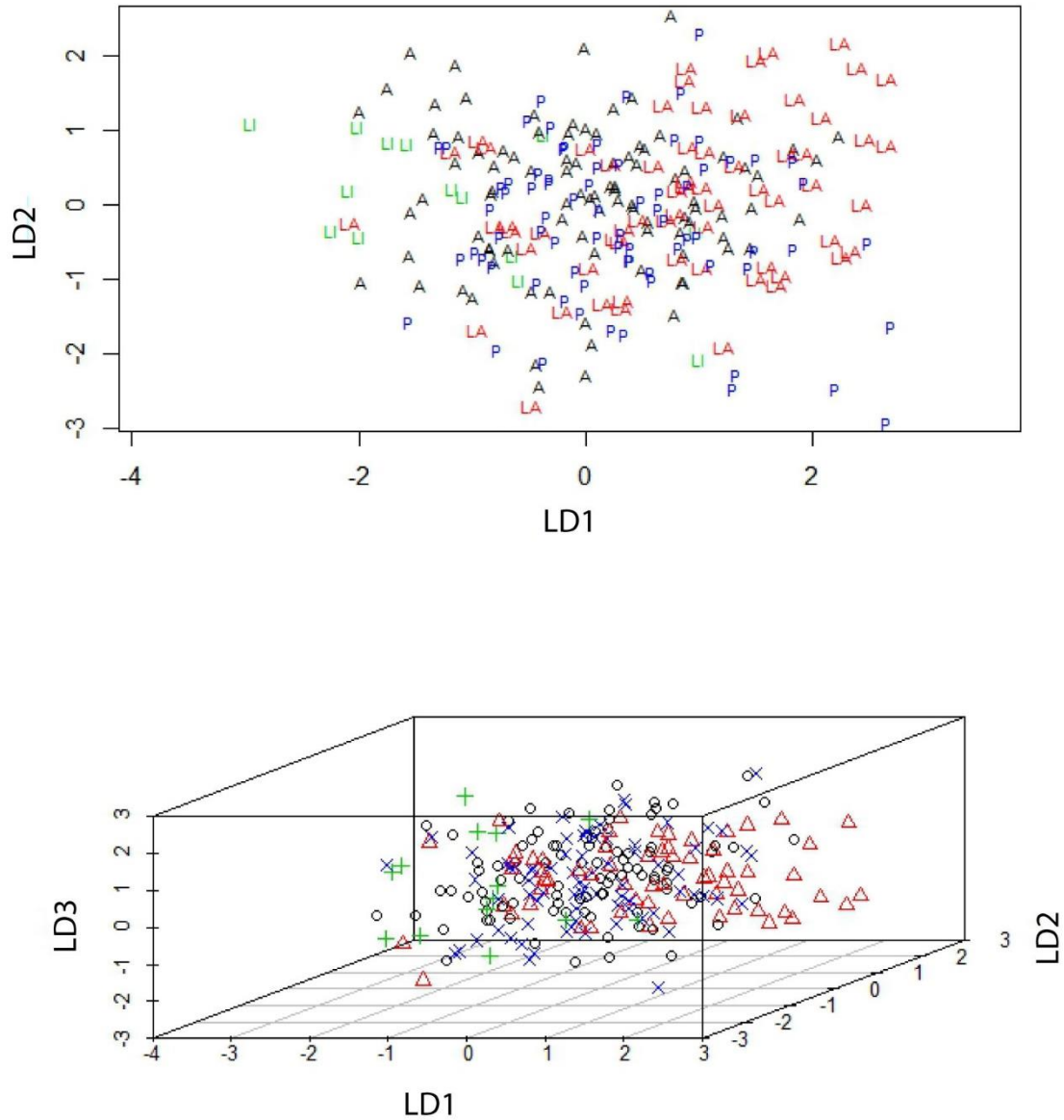


Figure 27. Two- and three-dimensional discriminant function scatterplots for analysis 5. Discriminant function 1 is plotted on the x-axis (LD1), discriminant function 2 is plotted on the y-axis (LD2), and discriminant function 3 is plotted on the z-axis (LD3). Var. *americanum* is represented as “A” on the 2-D plot and circles on the 3-D plot; var. *latifolium* is represented as “LA” on the 2-D plot and triangles on the 3-D plot; var. *lineare* is represented as “LI” on the 2-D plot and crosses on the 3-D plot; var. *pectinatum* is represented as “P” on the 2-D plot and X’s on the 3-D plot.

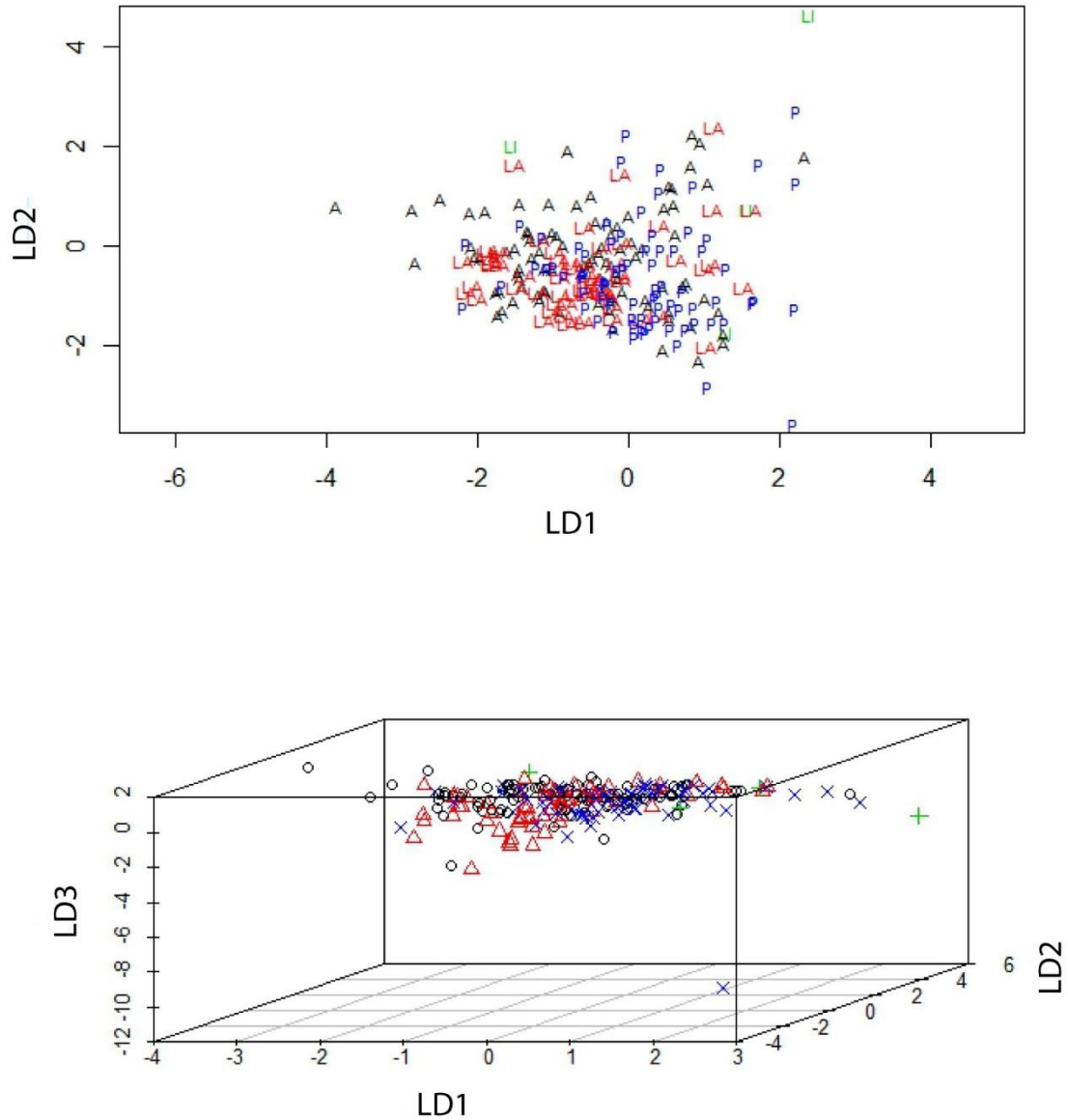


Figure 28. Two- and three-dimensional discriminant function scatterplots for analysis 6. Discriminant function 1 is plotted on the x-axis (LD1), discriminant function 2 is plotted on the y-axis (LD2), and discriminant function 3 is plotted on the z-axis (LD3). Var. *americanum* is represented as “A” on the 2-D plot and circles on the 3-D plot; var. *latifolium* is represented as “LA” on the 2-D plot and triangles on the 3-D plot; var. *lineare* is represented as “LI” on the 2-D plot and crosses on the 3-D plot; var. *pectinatum* is represented as “P” on the 2-D plot and X’s on the 3-D plot.

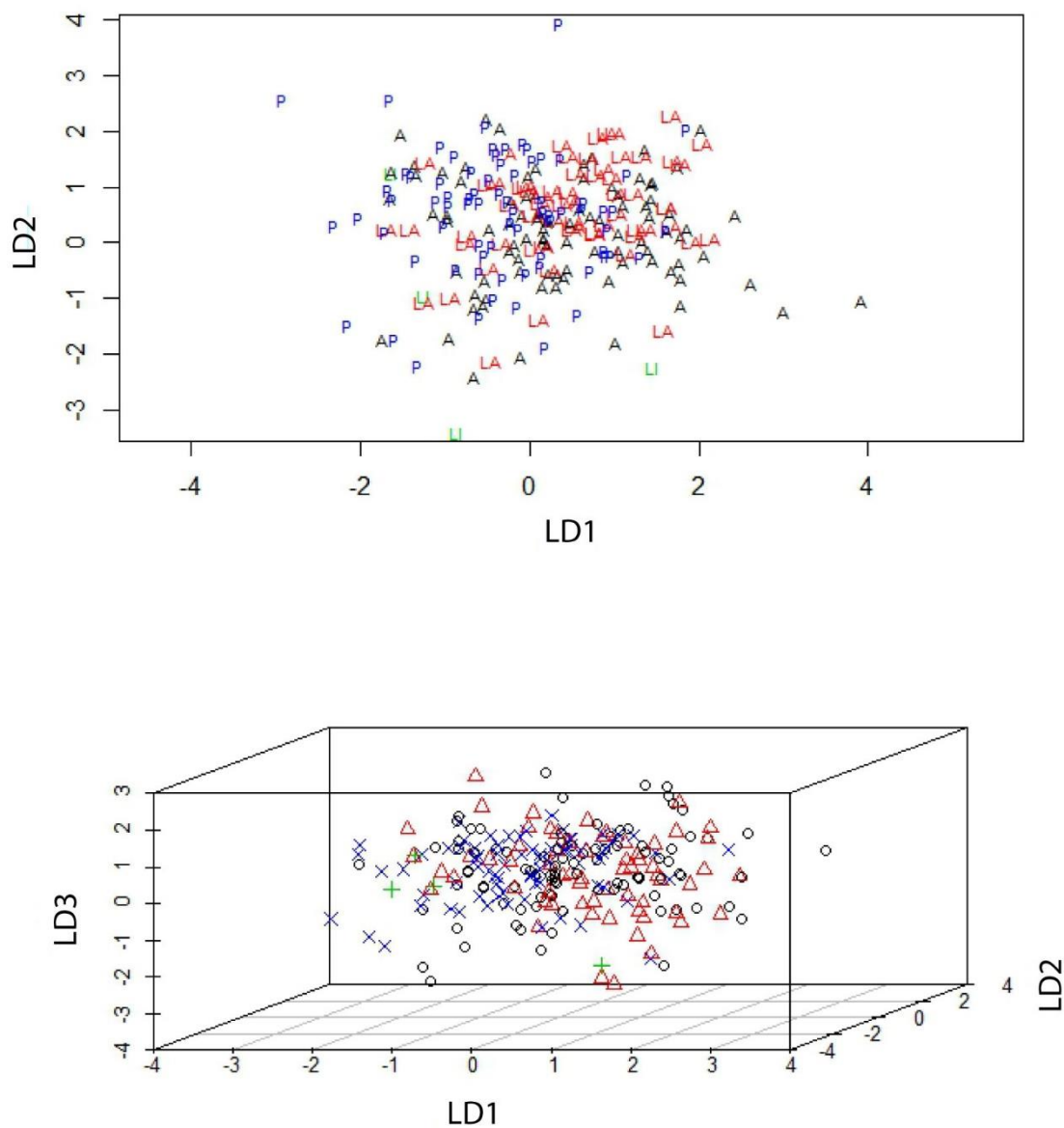


Figure 29. Two- and three-dimensional discriminant function scatterplots for analysis 7. Discriminant function 1 is plotted on the x-axis (LD1), discriminant function 2 is plotted on the y-axis (LD2), and discriminant function 3 is plotted on the z-axis (LD3). Var. *americanum* is represented as “A” on the 2-D plot and circles on the 3-D plot; var. *latifolium* is represented as “LA” on the 2-D plot and triangles on the 3-D plot; var. *lineare* is represented as “LI” on the 2-D plot and crosses on the 3-D plot; var. *pectinatum* is represented as “P” on the 2-D plot and X’s on the 3-D plot.

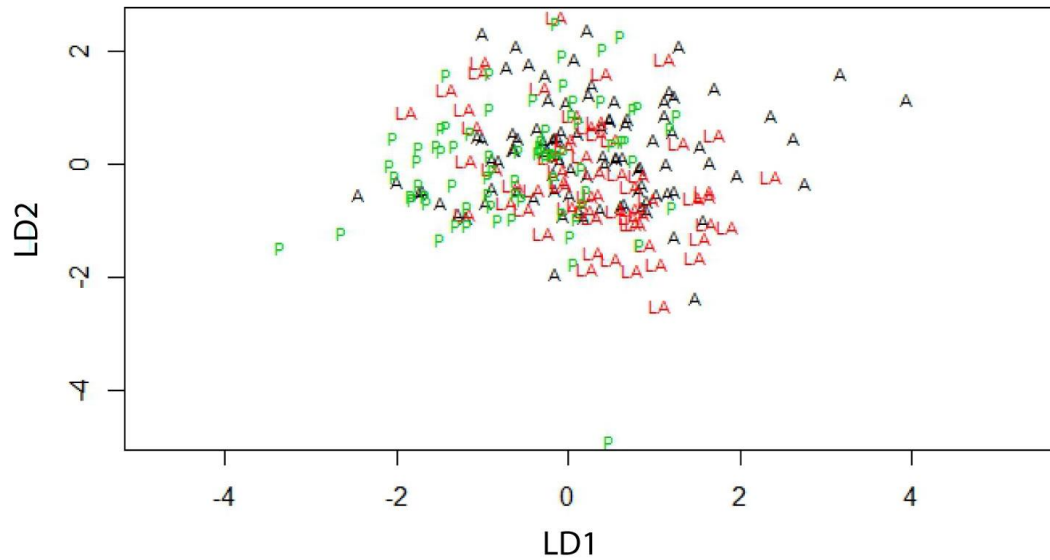


Figure 30. Two-dimensional discriminant function scatterplot for analysis 8. Discriminant function 1 is plotted on the x-axis (LD1), and discriminant function 2 is plotted on the y-axis (LD2). “A” represents var. *americanum*, “LA” represents var. *latifolium*, and “P” represents var. *pectinatum*. Var. *lineare* was excluded from this analysis, so a three-dimensional plot could not be generated.

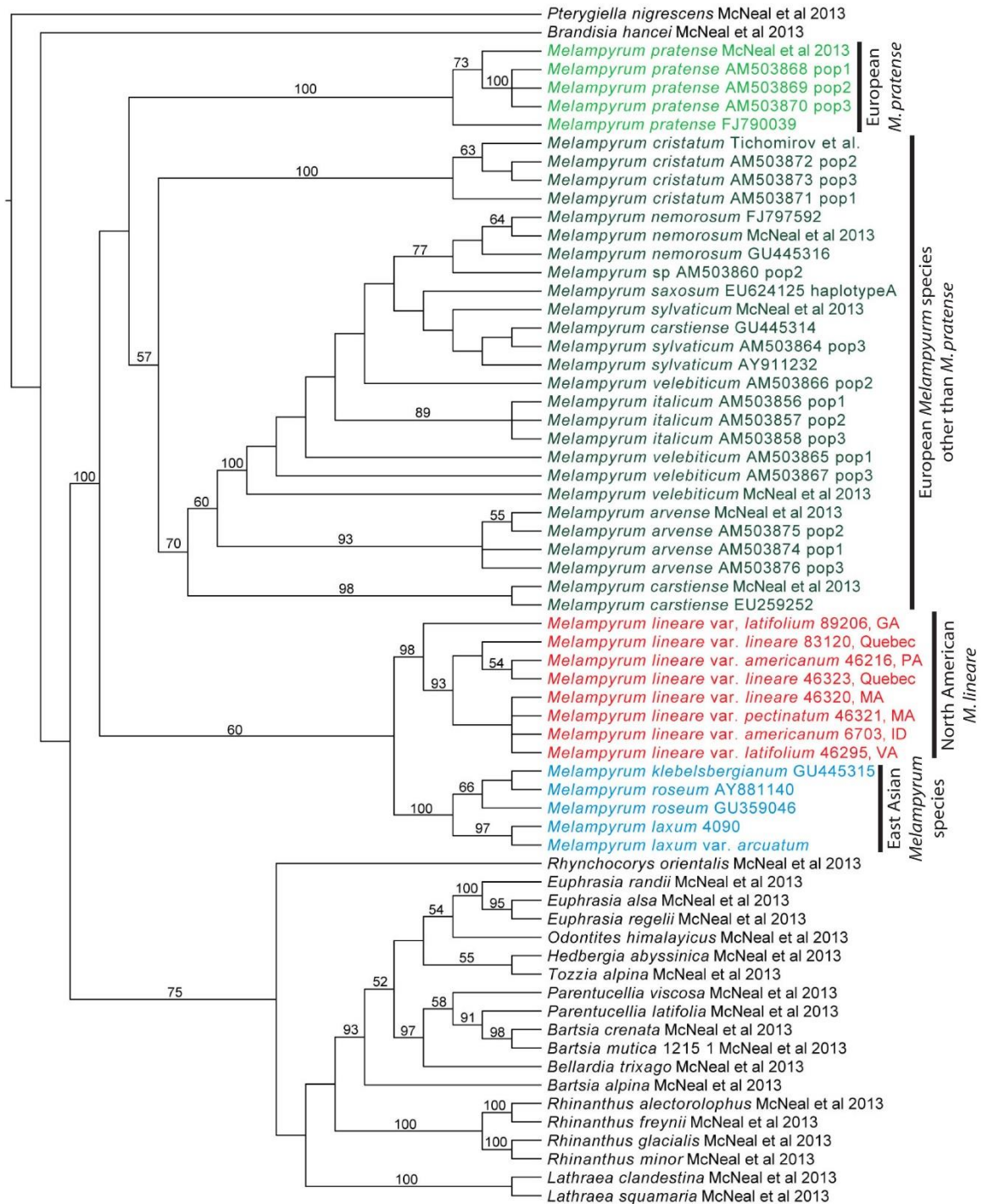


Figure 31. 50% Majority-rule consensus phylogeny of Rhinanthae sensu McNeal et al. (2013) from parsimony analysis of DNA nrITS sequences. Parsimony bootstrap values $\geq 50\%$ are shown above branches; unlabeled branches had bootstrap values $< 50\%$. CI = 0.545, RI = 0.830.

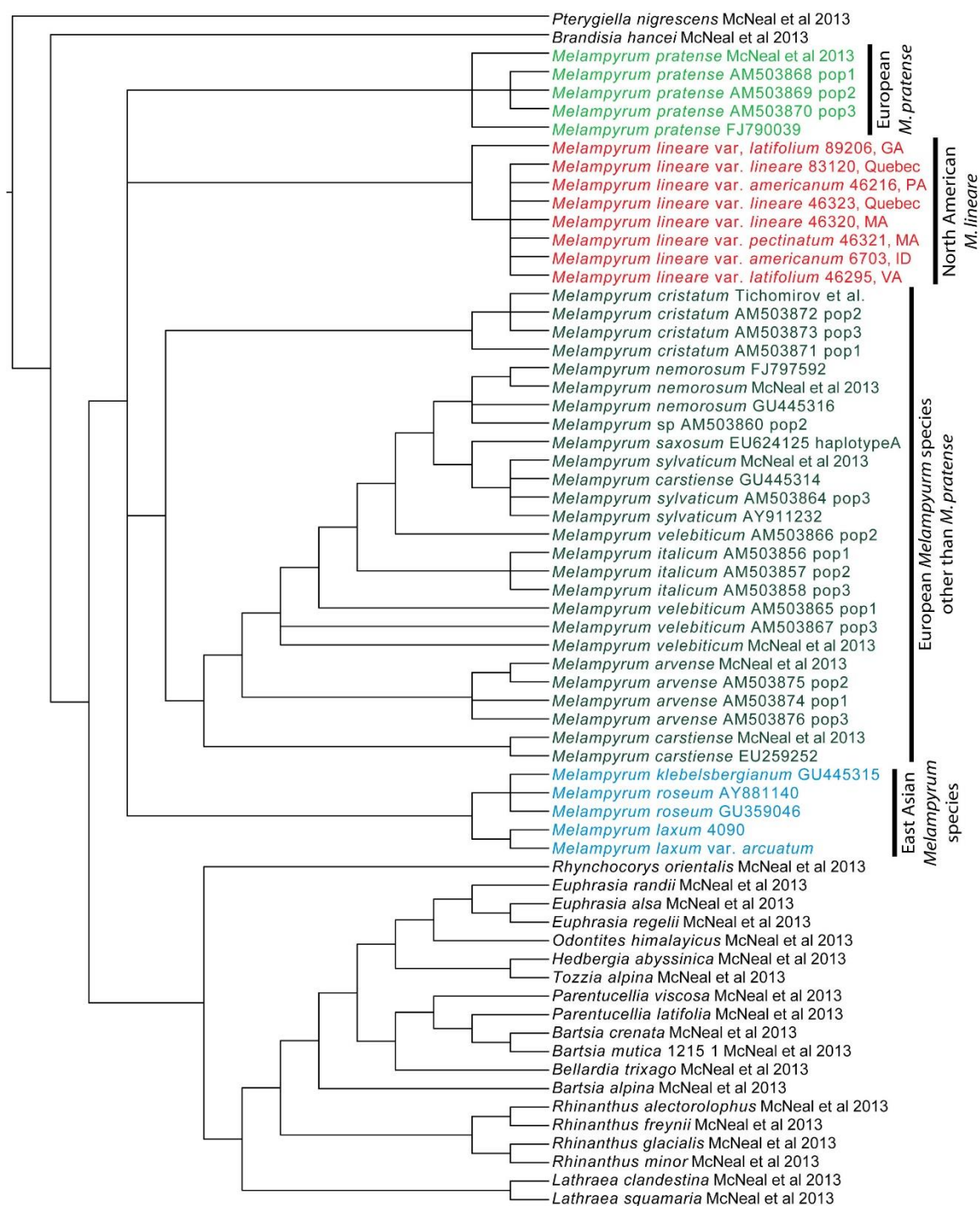


Figure 32. Strict consensus phylogeny of Rhinanthaeae sensu McNeal et al. (2013) from parsimony analysis of DNA nrITS sequences. Parsimony bootstrap values are the same as those in Figure 31.

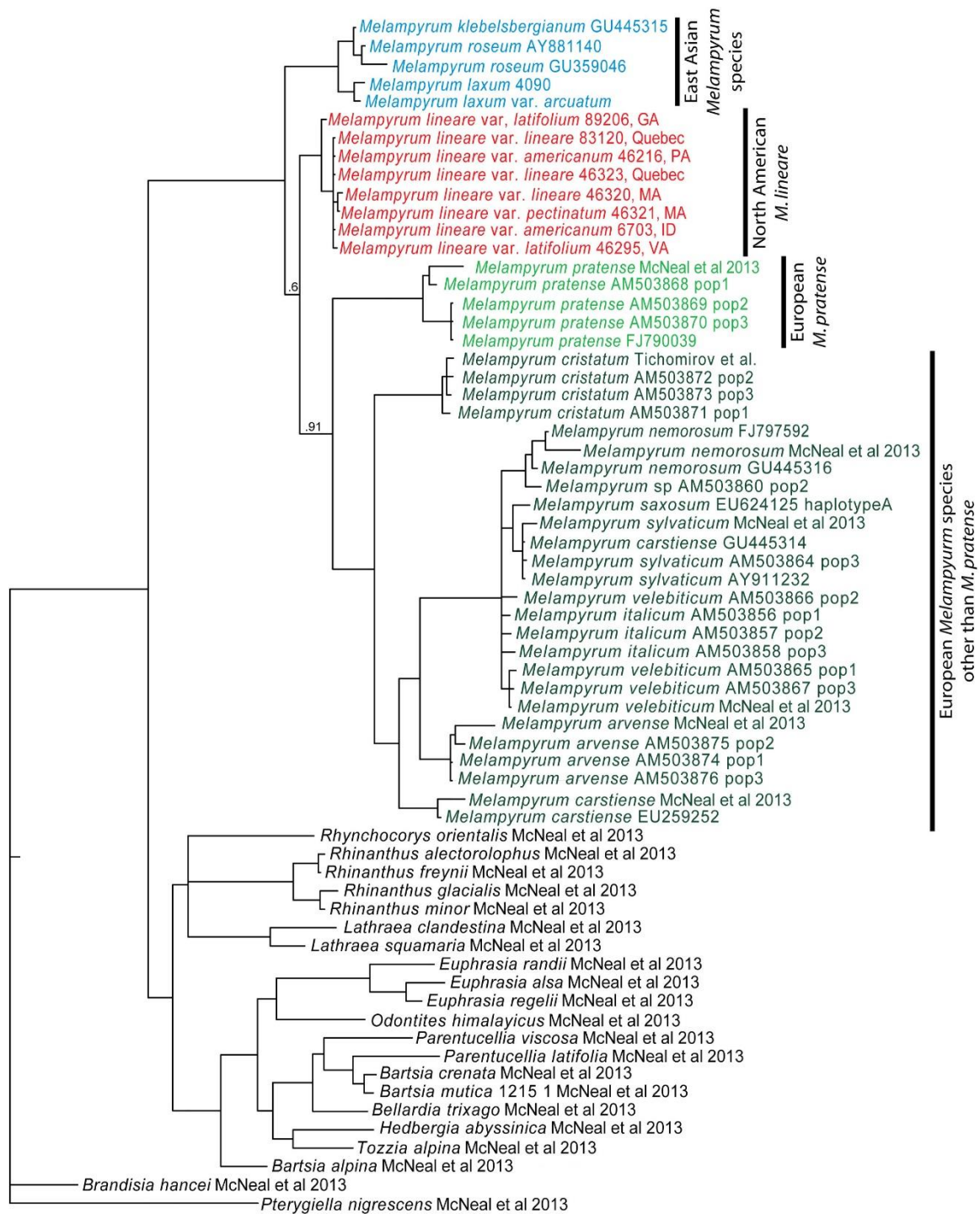


Figure 33. Bayesian 50% majority-rule consensus tree. Only posterior probabilities for the node joining *Melampyrum pratense* to a clade containing other European species of *Melampyrum* (0.91) and the node joining North American *M. lineare* to the European *Melampyrum* clade (0.60) are shown.

APPENDIX C: EXSICCATAE

Variety *americanum*:

Canada:

British Columbia: Summit Lake, 54° N, 122° W, 25 July 1953, *Florian, S.F. 101*, (UBC). Vanderhoof, Open grassy roadside, rest area along Rt. 16, 44 km east of Vanderhoof, BC. 55°55' N, 123°40' W, 18 July 1989, *Straley, G.B. 5622*, (UBC). North of William's Cr. Nr Lakelse Rd, 27 July 1952, *Schmidt, R.L. 8(110-52)*, (UBC). **Manitoba:** Flathead Co. About 2 mi. NE of Coram, W. half of boundary between Sec. 22 & 27, T. 31 N., R. 19 W, 5 August 1954, *Dunn, D.B. 10396*, (MIN); Big Creek of N. Fork of Flathead River, 12 mi. NW of West Glacier, Mount, southwest slope in lodgepole pine, alt. 3800 ft., 3 August 1962, *Stickney, P.F. 721*, (GH). Interlake Region. St. Martin, 25 km north of St. Martin, off hwy #6, 6 August 1982, *Shchepanek, M.J. 4456*, (MIN). Lake Co. Mission Mtns, common in dry, partially cut-over Douglas fir forest on a moderate west-facing slope above Yellow Bay, ca. 4,500 ft., 4 August 1983, *Lesica, P. 2769*, (MONTU). Missoula Co. 1 mi. west of Elbow Lookout, eastern foothills of the Mission Mts., about 80 mi. northeast of Missoula Twp. 19N, R 17W., near-corner 15, 16, 21, and 22, 15 August 1955, *Cronquist, A. 8110*, (GH); Lolo National Forest, Seeley Lake Campground, w. shore Seeley Lake, 31.5 mi. NE of Missoula, Sec. 33 T. 17N, R14W, 22 July 1973, *Stickney, P.F. 2983*, (MONT). Sundown, 60 mi. southeast of Winnipeg, dry jack-pine woods 7 mi. north of town, 12 August 1953, *Scoggan, H.J. 11622*, (MT); 60 mi. southeast of Winnipeg, dry jack-pine woods 7 mi. north of town, 12 August 1953, *Scoggan, H.J. 11622*, (GH). Whiteshell, Fish Hatcheries, S1-T10-R17E, 27 July 1954, *Love, D. 6518*, (UBC). **Nova Scotia:** Halifax. Armdale, dry rocky barrens, 28 July 1921, *Fernald, M.L. 24468*, (PH). **Ontario:** Bruce Peninsula. Cape Croker, 20 August 1926, *Watson, W.R. 3393*, (PH). Lincoln Co. Open woods near Bournique's, 14 July 1897, *McCalla, W.C. 447*, (BH). Thunder Bay District. N. end of Onion Lake, 2 km W. of H. 527. 48°42'N, 89°07'W, 21 July 1987, *Garton, C.E. 23890*, (MICH); Highway 61, 8 mi. east of Pigeon River Bridge, Cooks Township, 31 July 1952, *Garton, C.E. 1970*, (MT); Pukaskwa National Park, point at NW corner of English Fisheries, 18 July 1977, *Garton, C.E. 17798*, (MICH). **Québec:** Baie St. Paul, terrains acides, 17 August 1935, *Frère J.-Laurent sn*, (MT). Cap-aux-Os, Gaspé-Sud, 12 August 1963, *Frère Rolland-Germain 9682*, (MT). Gatineau Co. North of Kazabazua, 1 September 1933, *Pennell, F.W. 16750*, (PH); 14 July 1970, *Gaudieau L. sn*, (MT). Missisquoi Co. Venise-en-Québec, 14 July 1978, *Bernard J.P. B78-323*, (GH). Rimouski Co. Sandstone ridges and barrens, 29 July 1907, *Fernald, M.L. 1163*, (PH); Sandstone ridges and barrens, 29 July 1907, *Fernald, M.L. 1163*, (MT). Rivière à la Martre. 51°15'N, 76°20'W, 31 July 1943, *Dutilly, A. 11361*, (MT). Rivière-du-Loup. St. Antonin, 29 July 1966, *Blouin, J.L. 10901*, (OS). Rouyn-Noranda. Abitibi, 3341 rang Audet, environ 50 m à l'ouest de la cloture latérale (côté ouest du terrain), 9 July 2011, *Bouchard, C. 2011-002*, (MT). Terrebonne Co. Val-Morin, 23 July 1942, *Marie-Victorin 55146a*, (NY); Val-Morin, 23 July 1942, *Marie-Victorin*

55146b, (MT). Vaudreuil-Soulanges Co. Rigaud, 15 July 1934, *Robert, A. 1175a*, (MT). Wolfe Co. Sud de Coleraine, 10 August 1965, *Blais, V. 10698*, (UBC).

United States:

Connecticut: Windham Co. Thompson, Quaddick Pond State Forest, mixed woods between upper and middle ponds, 21 June 1986, *Hayden, W.J. 1485*, (URV). **Delaware:** New Castle Co. Low woods ca. 2 mi. north of Vandyke, 5 July 1940, *Long, B. 54672*, (PH). Sussex Co. Lewes, June 1866, *Canby, M. sn*, (NY). **Idaho:** Boundary Co. Purcell Mountains, Kaniksu National Forest, Meadow Creek CG on Moyie River ca. 7 mi. N of Moyie Springs, T63N R2E S12 sw1/4, elev. Ca. 2300', 11 July 1983, *Henderson, D. 6703*, (ID). **Indiana:** Millers Co. Thickets, 3 July 1914, *Umbach, L.M. 6676*, (NY).

Maryland: Cecil Co. Hill north of bridge over Piney Creek, .75 mi. NE of Elk Neck, 13 June 1939, *Pennell, F.W. 24811*, (PH). Talbot Co. 5 mi. NE of Easton, along road to Queen Anne, margin of woods, 7 June 1936, *Tatnall, R.R. 2965*, (PH). Worcester Co. Pocomoke State Forest, just south of Furnace Road, .5 mi. west of Nassawango Creek crossing, near dirt road about 4.7 air mi. NW of Snow Snow Hill, 14 June 1980, *Reveal, J.L. 5552*, (WVA); Along Mt. Olive Church Road, 1.3 mi. north of Maryland Highway 12 and 1.6 mi. south of Mt. Olive Church at Laws Road, about 4.7 air miles northwest of Snow Snow Hill, 14 June 1980, *Reveal, J.L. 5537*, (WVA). **Massachusetts:** Barnstable Co. Cape Cod, Dennis, Rail Trail, between East-West Dennis Road and Depot Street, 24 June 1999, *Hayden, W.J. 4168*, (URV). Berkshire Co. Along creek, base of Mt. Everett and Mt. Washington, 12 July 1922, *Meredith, H.B. sn*, (PH). Worcester Co. Berlin, open woods, 12 July 1938, *Potter, D. 8705*, (CONN); Berlin, open woods, 28 July 1938, *Potter, D. 8874*, (CONN); Berlin, open woods, 20 July 1938, *Potter, D. 7534*, (CONN). **Michigan:** Benzie Co. Shore of Lowe Herrin Lake, 3 mi. south of Frankfort, 19 August 1927, *R.R.D. 5313*, (PH). Cass Co. Magician Lake, 22 July 1910, *Umbach, L.M. 7141*, (BH). Charlevoix Co. High Island, in woods at south end of old Mormon settlement, T-39N, Sec. 34, 3 August 1958, *Clover, E.U. 89*, (MICH). Delta Co. Along pipeline ROW in moist sand, SW1/4 Sec 24 T41N-R21W, 5 July 1998, *Henson, D. 4286*, (MICH). Grand Rapids, 3 July 1891, *Bailey, H.M. sn*, (MICH). Leelanau Co. Open woods near Bournique's, 9 August 1986, *Hazlett, B.T. 4555*, (MICH). Mecosta Co. Aetna Twp. Sec. 19 NW 1/4, T13N, R10W, Rosy Run Creek at 2-track crossing about 3/4 mi. west of 230th Ave., 5 August 2000, *Ross, S. 732*, (MICH). Montmorency Co. Public fishing site landing on W. shore of Crooked Lake, S. of villate of Atlanta, T30N, R2E, Sec.23,SE1/4, 2 August 1985, *Garlitz 1292*, (MICH). Ogemaw Co. T23N R2E Sec.6, 10 July 1955, *Zimmerman, D.A. 1083*, (MICH). **Minnesota:** Clearwater Co. Jack pine forest above Lasalle Springs, Itasca Park, 10 August 1933, *Lakela, O. 603*, (MIN). Cloquet Co. Vacant lot, 13 August 1940, *Alaspa E. sn*, (MIN). Cook Co. R-6E, T-63N, S-4, trail intersect with U.S. Hwy 61, dry, open woods and talus upslope from Hwy., 3 August 1984, *Monson, P.H. 6067*, (MIN). Hubbard Co. Jack pine forest, 1.5 mi. north of Hubbard, 8 July 1941, *Moore, J.W. 14806*, (MIN). St. Louis Co. Duluth, end of the pine ridge, Sec. 19, 19 August 1936, *Lakela, O. 1746*, (MIN). Wadena Co. Jack pine woods 3 mi. south of Hubbard, 8 July 1941, *Moore, J.W. 14829*, (MIN). **New Jersey:** Camden Co. Atco, 6 August 1903, *Stewart C.C. 2988*, (NY); Essex Co. Montclair Heights, dry open woodland, 9 July 1916, *Pennell, F. 7429*, (NY); Monmouth Co. Wooded roadside bank,

just W. of Walns Mill, 3 August 1947, *Long, B. 66114*, (PH); Wall Township, Hurley Pond Road S. side near the airport, 1/2 mile E of Allaire State park, roadside in sandy soil, 22 July 2003, *Barringer, K. 9993*, (PH). Passaic Co. Great Notch, 6 July 1921, *Denslow, H.M. sn*, (NY); Ringwood, Ringwood Junction, 9 July 1916, *Mackenzie, K.K. 7159*, (NY). Salem Co. Riddleton, dark low woods W of station, 20 June 1926, *Adams, J.W. 305*, (PH). Warren Co. Phillipsburg, 17 July 1896, *Tyler, A.A. sn*, (NY). **New York:** Essex Co. Schroon, 9 August 1881, *I.A.K. sn*, (PH). Ulster Co. Roadside near Wildmere, 13 July 1945, *Githens, T.S. sn*, (PH). **North Carolina:** Cherokee Co. Along old wagon road, above and west of Old Road Gap, elev. 3500 ft., 5 August 1951, *Fox, W.B. 5118*, (GH). Haywood Co. Near top of Mt. Pisgah in red-oak woods, along trail, 8 August 1951, *Fox, W.B. 5203*, (GH); Near top of Mt. Pisgah in red-oak woods, along trail, 8 August 1951, *Fox, W.B. 5203*, (NY). Macon Co. Woodland 5 mi. east of Highlands, 17 July 1932, *Wherry, E.T. sn*, (PH); Summit of Mt. Satulah, intermixed with scrub, rocky openings, 20 July 1951, *Godfrey, R.K. 51436*, (GH); Summit of Mt. Satulah, intermixed with scrub, rocky openings, 20 July 1951, *Godfrey, R.K. 51436*, (MT); Summit of Mt. Satulah, intermixed with scrub, rocky openings, 20 July 1951, *Godfrey, R.K. 51436*, (NY). Watauga Co. Blowing Rock, dry soil on exposed, rocky ledge, 3 August 1922, *Randolph, L.F. 1164*, (BH). **Pennsylvania:** Bedford Co. Half mile E.N.E. of Martin Hill Fire Tower, 5 August 1945, *Berkheimer, D. 6396*, (PH). Bucks Co. Margin of dry woods ca. 1 mi. S. of Mechanicsville, 22 June 1947, *Long, B. 65967*, (PH). **Virginia:** Accomac Co. Chincoteague Island, 14 June 1940, *Gleason, H.A. 8545*, (NY). Alleghany Co. Sandy soil in clearing near Potts Pond, Potts Mt., 23 July 1966, *Harvill, A.M. 14532*, (NCU). Fairfax Co. Fountainhead Park, 29 July 1970, *Weand, D. sn*, (GMUF). Highland Co. Shenandoah Ridge, east of head waters, sandy shale, 25 August 1927, *Wherry, E.T. 13775*, (MIN). Princess Anne Co. Seashore State Park, sandy soil in mixed woods, 14 May 1949, *Mikula, B. 272*, (GMUF). Roanoke Co. Rd. 612 ca. 3200' up north slope Poor Mountain, 3 September 1968, *Uttal, L.J. 9256*, (VPI). **West Virginia:** Greenbrier Co. While Sulfur Springs, 28 July 1931, *WVU Biological Expedition sn*, (WVA). Preston Co. 6 July 1891, *Millsbaugh, C.F. 897*, (WVA). Tucker Co. Dolly Sods, along ridge road, 24 July 1955, *Davis, H.A. 11149*, (WVA). **Wisconsin:** Door Co. Baileys Harbor, T30N, R28E, Sec17, SW1/4, The Ridges, 17 August 1938, *Pohl, R.W. 1136*, (MICH). Douglas Co. Wisconsin Point, about 2.5 mi. out, Superior, WI, 13 August 1936, *Lakela, O. sn*, (MIN); Wisconsin Point, 13 August 1936, *Horton, E.S. sn*, (MIN). Manitowoc Co. Two Rivers, 1 August 1938, *Benke, H.C. 5908*, (PH). Waushara Co. Wild Rose, 31 July 1913, *H.V.O. sn*, (MICH).

Variety *latifolium*:

Canada:

Ontario: Thunder Bay District. 1/8 mi. east of SE corner of Long Point Bay, 4 mi. NE of Black Sturgeon Research Station, 22 July 1969, *Garton, C.E. 12169*, (MT); Dry woods under spruce and birch on old raised bolder beaches, south shore of St. Ignace I. at Locomotive Rock opposite Nest I., 18 August 1959, *Garton, C.E. 7000*, (NCU). York

Co. East Gwillimbury Township, concession IW, lots 111-115, about 1.5 mi NNW of Hollan Landing, east side of Holland River, 22 July 1979, *Reznicek, A.A. 5097*, (MICH).

United States:

Connecticut: Fairfield Co. Bridgeport, 18 July 1892, *Eamer, E.H. 3231*, (CONN); Botsford, sandy roadside bank, 24 July 1955, *Wilkens, H. 8859*, (PH). Litchfield Co. Woodbury, Orenaug Hills, trail up south side, 41°32'38N, 73°12'06W, 4 July 1983, *Mehrhoff, L.J. 8156*, (CONN). New Haven Co. Seymour, pasture, 27 June 1920, *Neuman, C.W. sn*, (CONN). New London Co. Waterford, moist woods, 8 June 1937, *Jansson, K.P. sn*, (CONN); Stonington Dennison-Pequotsepos Nature Center, Pequotsepos Road at junction of Main Trail and south end of Hidden Pond Trail in dry oak-hickory forest, 16 July 1987, *Crossman, T.I. 007*, (CONN); Groton, moist woods, 9 June 1938, *Jansson, K.P. sn*, (CONN). **Georgia:** Rabun Co. Bly Gap - Dicks Creek Gap AT, Towns and Rabun counties, 12 June 1980, *Koth, L. 20*, (NY); At edge of logging road on top of Patterson Gap, 4 June 1966, *Montgomery, F. 518*, (PH). **Indiana:** LaPorte Co. Sandy black oak woods immediately north of Road 20 at junction with Road 35, 5 mi. east of Michigan City, 14 June 1946, *Friesner, R.C. 204683*, (GH). **Maine:** Lincoln Co. South Bristol, McFarland Cove Road, woods along roadside, 43°51'50N, 69°33'40W, 5 August 1996, *Mehrhoff, L.J. 18968*, (CONN). **Massachusetts:** Barnstable Co. Woods Hole, 24 June 1908, *Lewis, W.H. sn*, (NY). Dukes Co. Martha's Vineyard, North Tisbury, West Tisbury, 30 June 1916, *Seymour, F.C. sn*, (NY). Hampshire Co. Woodland, Horse Mt., Hatfield, 29 June 1973, *Ahles, H.E. 77672*, (WTU). Plymouth Co. Manomet, along sandy edge of public pond, Bartlett Road, 30 July 2000, *Schori, M. 2000-24*, (BH). Worcester Co. Open woods, East Brookfield, 4 July 1938, *Potter, D. 86697*, (CONN); Sutton, 21 July 1933, *Hodge, W.H. 2468*, (CON); Open woods, Lunenburg, 28 July 1938, *Snide, J.H. 9026*, (CONN). **Michigan:** Kent Co. 3 August 1893, *Cesle, E. 40893*, (MICH). **Minnesota:** Crow Wing Co. Along the Pine River 4 mi. SSE of the town of Cross Lake, T136N R27W NW 1/4 of NW1/4 of SE1/4 Sec. 4, SW1/4 of Ne1/4 Sec.9, 30 July 1998, *Lee M.D. 2329*, (MIN). **New Jersey:** Cumberland Co. Dry sterile wood, Long Branch Run, 12 June 1932, *Long, B. 37667*, (PH). Essex Co. Dry woodland, Montclair Heights, 9 July 1916, *Pennell, F.W. 7429*, (PH). **New York:** Warren Co. Dry woods, 5 mi. north of Belton Landing, 21 July 1945, *House, H.D. 29952*, (BH). Washington Co. Fred Vaughan's woods, S.Beaver creek, north of Kingsbury St., 29 July 1917, *Burnham, S.H. sn*, (PH). **North Carolina:** Buncombe Co. Moist woodlands near Biltmore, North Carolina, 29 May 1897, *Biltmore Herbarium 636*, (NY); Moist woodlands near Biltmore, North Carolina, 29 May 1897, *Biltmore Herbarium 636*, (MICH). Haywood Co. Oak woods, Cove Creek Gap, 15 June 1931, *Wiegand, K.M. sn*, (BH). Henderson Co. One mile southeast of Mills River, 24 May 1957, *Freeman, O.M. 57336*, (NCU). Macon Co. Appalachian Trail, Wallace Gap to Albert Mt., 10 June 1975, *Weiss, T. 150*, (NY); 9 mi west on Coweeta Road from its jct. with US 441, then .3 mi south on Norton Road to Stewart Road, 25 May 1990, *Jones, S. 5002*, (NY); North side Mt. Satula, Highlands, 28 July 1934, *Harbison, T.J. 1262*, (TENN); Ridge west of Cliffside Lake and east of Long Branch, Northwest of Highlands, 35°5'N, 83°14'30W, 6 June 1975, *Boufford, D.E. 16163*, (NCU). Madison Co. Woods near Hot Springs, 5 June 1956, *Freeman, O.M. 56313*, (NCU). Transylvania Co. Deciduous woods 2.5 mi. north of jct. US 276 and US 64, 5

June 1963, *Bradley, T.* 747, (GMUF); Edge of Faith Memorial Chapel Road, dry pine woods, Cedar Mountain, 11 June 1956, *Horton, J.H.* 225, (NCU). **Ohio:** Geauga Co. Thompson Quad, in semishade on tops of sandstone ledges, Thompson Ledges Park, south of Thompson Rd and east of St. Rt. 528 at Thompson, Thompson twp., 25 June 1991, *Cusick, A.W.* 29685, (NY); Thompson Quad, in semishade on tops of sandstone ledges, Thompson Ledges Park, south of Thompson Rd and east of St. Rt. 528 at Thompson, Thompson twp., 25 June 1991, *Cusick, A.W.* 29685, (MICH). Lorain Co. Oak-hickory woods, Ohio 2 rest area just west of Vermilion River, 1 July 1979, *Jones, G.T.* 79-7-1-240, (OS). Summit Co. Twinsburg Twp., rocky woods, top of south-facing slope, east side Cannon Road, 1 mile east of Route 91, 25 June 1956, *Herrick, E.M.* 1736, (OS). **Pennsylvania:** Berks Co. Detunks Bridge 1 mi. NW of Kutltown, dry soil on wooded slope, 26 June 1967, *Brumbach, W.C.* 5866, (PH). Bucks Co. Bensalem Twp, dry woods, Torresdale Manor, Delaware River, Andalusia, 22 June 1924, *Long, B.* 31009, (PH). Lancaster Co. Churchtown Road, dry woods on Welsh Mountain, 17 June 1909, *Long, B. sn.* (PH). Lehigh Co. Woods, .5 mi. north of Werleys Corner, 17 August 1950, *Schaeffer, R.L.* 34320, (PH). Perry Co. Acid rocky soil near base, wooded slope, Run Gap, 2 mi. NW of Ickesburg, 7 July 1947, *Adams, J.W.* 47-52, (PH). Philadelphia. York Furnace, 4 July 1904, *Crawford, J. sn.* (PH). Warren Co. In dense woodland, North Warren, 1 August 1943, *Moldenke, H.N.* 15446, (PH). **Tennessee:** Blount Co. Great Smoky Mountains National Park, Panther Creek Watershed, Calderwood 7.5 minute quadrangle, Hannah Mountain Trail, Elevation about 3500 feet, 35°31'48N, 83°52'40W, 9 July 2001, *Busemeyer, D.T.* 658, (TENN). Cooke Co. Rich low woods near Wolf Creek, May 1893, *Kearney, T.H. sn.* (OS). Grundy Co. Sandy acidic woods by I-24, just west of Monteagle near escarpment, 25 May 1974, *Kral, R.* 52889, (TENN). Monroe Co. Along Forest Service Road 217-1 between Rattlesnake Rock and Beech Gap, Cherokee National Forest, 25 June 1979, *Wofford, B.E.* 79-172, (TENN). Polk Co. Near highest pass, Little Frog Mt., road from Reliance to Ducktown, 12 June 1933, *Wherry, E.T. sn.* (PH). Sevier Co. Elkmont, in well drained woods soil, border of woods along trail on Fightin' Creek Gap, 29 May 1935, *Jennison, H.M.* 235, (PH). Unicoi Co. Along roadside to Pleasant Gardens, from Erwin Tennessee, 9 June 1971, *Odenwelder, J.C.* 71-127, (TENN). **Virginia:** Accomac Co. Pine woods 2.7 mi. north of Oak Hall, Route 13, 1 June 1935, *Tatnall, R.R.* 2601, (PH). Avery Co. Blue Ridge Parkway M.P. 308.3 R., Flat Rock, woods, 9 June 1967, *Crandall, D.L.* 10232, (LYN). City of Chesapeake. Mesic beech-oak forest on upland island in swamps of the Northwest River, ca. 1.0 mi. south of jct. Indian Creek Road and Gallbush Road, 1.5 mi. NW of Northwest, 30 May 1996, *Fleming, G.P.* 11663, (GMUF); Mesic beech-oak forest on upland island in swamps of the Northwest River, ca. 1.0 mi. south of jct. Indian Creek Road and Gallbush Road, 1.5 mi. NW of Northwest, 30 May 1996, *Fleming, G.P.* 11663, (WILLI). Giles Co., Moist, acid chestnut oak-heath woods on slope along Dismal Creek, .3 mi. ENE of Walnut Flats Campground, 2.5 mi. WNW of White Gate, elev. 2400 ft., 21 June 1995, *Fleming, G.P.* 10628, (GMUF). Patrick Co. Rocky Knob Fire Road, Blue Ridge Parkway M.P. 174 L., 25 June 1966, *Crandall, D.L.* 10231, (LYN). Pittsylvania Co. Rich woodlands, shady, 15 June 1963, *Hathaway, W.* 15650, (LYN). Pulaski Co. Along Max Creek Road to Powhatan Boy Scout Camp, 4 July 1972, *Uttal, L.J.* 8861, (VPI).

Variety *lineare*

Canada:

British Columbia: Kishinena Valley, 40°N, 114°W, 22 July 1970, *Beamish, K.I.* 760, (UBC). **New Brunswick:** St. John Co. Saint John, Lancaster, 45°15'N, 66°06'W, 6 July 1878, *Hay, G.U.* VP-10695, (MT). **Newfoundland:** Port aux Basques. Cape Ray, dry peaty barrens among the gneise hills, 21 July 1924, *Fernald, M.L.* 27026, (PH). **Nova Scotia:** Shelburne Co. Boggy barrens, upper Wood's Harbor, 14 July 1921, *Fernald, M.L.* 24466, (PH). **Québec:** Anse aux Canards. Dry stony barren, east of Anse aux Canards (east of Newport), 11 July 1931, *Fernald, M.L.* 587, (CONN); Dry stony barren, east of Anse aux Canards (east of Newport), 11 July 1931, *Fernald, M.L.* 587, (MIN); Dry stony barren, east of Anse aux Canards (east of Newport), 11 July 1931, *Fernald, M.L.* 587, (MICH). Peribonka Region. Northeast of Lake St. John, around Lac Alex, Patrick West R., and vicinity, 28 July 1948, *Hustich, I.* 361, (MT). **Saskatchewan:** Lac Ile-a-la-Crosse. South Bay, 1 mi. NE of Mile 82.5, Hwy 155, .5 mi. NW of Lac Ile-a-la-Crosse Provincial Campsite, T73, R13, W3rd, 55°22'N, 107°51'W, 18 July 1971, *Harms, V.* 17999, (GH).

United States:

Maine: Lincoln Co. South Bristol, Pemaquid Point, Pemaquid beach along west shore of Point, west of new Harbor, 43°49'53N, 69°30'56W, 11 August 1977, *Mehrhoff, L.J.* 2228, (CONN). York Co. Saco heath bog, 43°33'10N, 70°28'13W, 20 July 1995, *Mehrhoff, L.J.* 18235, (CONN). **Maryland:** Charles Co. Just north of Smallwood Rd. and 1 mi. south of MD 228, E side of Hwy 301, Waldorf, 31 May 1980, *Hill, S.R.* 9161, (NY). **Minnesota:** Cook Co. Swamp east of Grand Portage, 22 August 1929, *Bemm, F.R.* 678, (MIN). **New York:** Hamilton Co. Sunny dry slope, Indian Lake, 22 July 1929, *Lambert, B.B.* 26a, (PH).

Variety *pectinatum*:

Canada:

British Columbia: Bridesville. 7.5 km on Mt. Baldy Road, north of Rt. 3 in Bridesville, 7 August 1994, *Straley, G.B.* 7920, (UBC). **Ontario:** Bruce Peninsula. Edge of clearing, acid woods, near Indian Harbour, 3 August 1959, *Sherk, L.C.* sn, (BH). Fort William. Sphagnum bogs, 24 July 1912, *Williamson, C.S.* 1724, (PH). Manitoulin Island. On hillside below Battery Bluff, 9 August 1932, *Grassl, C.O.* 3716, (MICH). Thunder Bay District. 1 km. south of Mattawin Rd. of Great Lakes Forest Products limits just east of Nelson Creek, 19 July 1981, *Garton, C.E.* 20358, (MICH). **Québec:** Lac St-Jean. Dolbeau, 29 July 1937, *Morin, C.* 714, (MT). Rimouski Co. Dry woods, 12 July 1910, *Bartram, E.B.* 457, (PH). Rouyn-Noranda. Réserve d'Aiguebelle, Québec Forêt, extrémité sud-est de la réserve, près de la jonction pour la route de Destor, 29 July 1978, *Bergeron, Y.* 78-242a, (MT). Saguenay. Ilets Jeremie, terrain sablonneux, bord du St-Laurent, 10 August 1961, *Brisson, F.S.* 64, (MT). Trois-Rivieres, August 1929, *Stanislas, F.* 686, (MT). **Saskatchewan:** Meadow Lake Provincial Park. North shore of Little Raspberry Lake, ESE of Kimball Lake, 54°24'N, 108°49'W, T62, R19, W3rd, 18 July 1978, *Harms,*

V.L. 24702, (GH); North shore of Little Raspberry Lake, ESE of Kimball Lake, 54°24'N, 108°49'W, T62, R19, W3rd, 18 July 1978, *Harms, V.L. 24702*, (MT).

United States:

Connecticut: Hartford Co. Granby, Manitook Mountain, woods on Traprock ridge, The Notch, 41°59' 04"N, 72°46'44"W, 28 July 1981, *Mehrhoff, L.J. 4594*, (CONN). **Idaho:** Boundary Co. Moyie River Valley, ca., 75 mi. south of Eastport (Canadian line), T65N R2E S15 NE4NE4, 24 July 1990, *Moseley, B. 1801*, (NY). **Kentucky:** Wolfe Co. Sky Bridge, 18 June 1944, *McFarland, F.T. 76*, (BH); Sky Bridge, 18 June 1944, *McFarland, F.T. 76*, (NY). **Massachusetts:** Franklin Co. In oak woods on granitic substrate on the south ridge of Jerusalem Hill ca. 2 km NW of Ruggles Pond, Town of Wendell, Elevation ca. 200 m, 3 August 1975, *Cronquist, A. 11330*, (NY). Plymouth Co. Dry oak and pine scrub in gravelly soil, Massachusetts State Forest, SW of Plymouth, 28 August 1932, *Weatherby, C.A. 588*, (MIN); Dry oak and pine scrub in gravelly soil, Massachusetts State Forest, SW of Plymouth, 28 August 1932, *Weatherby, C.A. 588*, (MONT); Dry oak and pine scrub in gravelly soil, Massachusetts State Forest, SW of Plymouth, 28 August 1932, *Weatherby, C.A. 588*, (CONN). Worcester Co. Shrewsbury, 1 August 1937, *Dodge, R.B. 6416*, (CONN). **Michigan:** Bay View, 10 August 1896, *Allan, C.W. sn*, (MIN). Cheboygan Co. Herb in the aspens near the Biological Station, 26 July 1927, *Gates, F.C. 14838*, (MIN); T25N R1W Sec. 5, 27 June 1951, *Zimmerman, D.A. 39*, (MICH); Jack pine plains, 5 miles south of Indian River, 20 July 1924, *Erlanson, C.O. 451*, (MICH); Jack pine plains south of Burt Lake, 21 August 1920, *Ehlers, J.H. 1239*, (MICH). Houghton Co. Edge of spruce-birch woods near shore of Lake Superior, Calumet Water Works, 4 mi. NW of Calumet, 24 July 1936, *Hermann, F.J. 8267*, (PH). Iosco Co. Open grassy dunes 2 miles south of Oscoda, 11 August 1951, *McVaugh, R. 12503*, (MICH); 200 ft. north of the end of Rhea Road at intersection with Bissonette Road, Oscoda Twp. T24N, R8E, Sec. 26 NE1/4, 1 August 1985, *Garlitz, D. 109*, (MICH). Iron Co. Along US hwy 2/US hwy 41, 5 mi. south of its jct. in Crystal Falls with State hwy 69, 4 August 2004, *Lammers, T.G. 11877*, (NY). Kent Co. Damp rich woods, 1 July 1893, *Cole, E.J. sn*, (PH). Mackinac Co. Pte. Aux Pins, Bois Blanc Island, 5 August 1932, *Ehlers, J.H. 5715*, (MICH). Mason Co. Ludington State Park, 25 July 1937, *Bartlett, H.H. sn*, (OS); Ludington State Park, 25 July 1937, *Bartlett, H.H. sn*, (MICH). Roscommon Co. Cedar swamp, Prudentville, 31 July 1927, *Dreisbach, R.R. 5194*, (PH). **Minnesota:** Ithaca Co. Pine barrens, August 1891, *Sandberg, J.H. sn*, (MIN). **New Jersey:** Burlington Co. Herb on small mound at edge of bog, Wharton State Forest, Medford Lakes, 27 August 1974, *Bazzolo, T. 342*, (CONN); Chatsworth, 20 August 1948, *Lawrence, G.H.M. 511*, (BH). Monmouth Co. Dry, sandy thickets along railroad, 4 mi. SW of Eatontown, 26 September 1937, *Fogg, J.M. 13914*, (PH); Dry, sandy soil in pine-barrens, South Belmar, 5 September 1942, *Thorne, R.F. 1051*, (BH). Ocean Co. In pine woods along inlet back of Point Pleasant, 17 August 1940, *Bright, J. 17848*, (MONT); 5.9 mi. SE of state road 72 on state road 539, 15 August 1975, *Boufford, D.E. 17957*, (GH). **New York:** Albany Co. Pine woods west of Albany, 23 August 1940, *House, H.D. 27626*, (GH); Pine woods west of Albany, 23 August 1940, *House, H.D. 27626*, (BH); Sand plains west of Albany, 28 August 1934, *Muenschner, W.C. 4722*, (BH); Pine woods west of Albany, 23 August 1940, *House, H.D. 27626*, (NY). Tompkins Co. Dry bluff, south side Shurger's Glen,

Lansing, 8 August 1917, *Gershoy, A. 8771*, (BH). **North Carolina:** Henderson Co. 16 August 1885, *Memminger, E.R. sn*, (NCU). Macon Co. Summit of Mt. Satulah, intermixed with scrub, rocky openings, 20 July 1951, *Godfrey, R.K. 51436*, (BH); Summit of Mt. Satulah, intermixed with scrub, rocky openings, 20 July 1951, *Godfrey, R.K. 51436*, (WVA). **Ohio:** Hocking Co. Deep Woods Farm, ca. 1.5 mi. SE of South Bloomingville, along SR 56, 39°24.489'N, 82°34.523'W, frequent at border of woods along ridgetop prairie above house, 16 June 2010, *Rose, J. 10-459*, (OS). **Pennsylvania:** Berks Co. Mt. Neversink, open woods on north slope above 12th St., 13 August 1946, *Wilkens, H. 7839*, (PH); Rocky woods, alt. 900 ft. 1 mi. east of Landis Store, 29 August 1955, *Berkheimer, D. 17221*, (PH). Centre Co. Oak barrens, 2 mi. west of State College, 11 August 1937, *Wahl, H.A. sn*, (PH); Oak barrens, 2 mi. west of State College, 11 August 1937, *Wahl, H.A. sn*, (PH); Dry hillside clearing 3.5 mi. NW of Bellefonte, 15 August 1945, *Wahl, H.A. sn*, (BH). Elk Co. Open hillside path south of borough of Ridgeway, 22 August 1943, *Rood, A.N. 203*, (PH). Monroe Co. In woods along trail, Buck Hill Falls, 21 July 1926, *Modenke, H.N. 2980*, (NY). Pike Co. Sandy humus in open woods, Twin Lakes, 15 July 1960, *Wherry, T.E. sn*, (PH). **Tennessee:** Hampton Co. Floodplain and slopes of Laurel Forks Road on N edge of Laurel Forks Wildlife Management Area, 12 September 1973, *Gonsoulin, G. 3406*, (VPI). Sevier Co. Bullhead Trail, Gatlinburg, 5 August 1935, *Jennison, H.M. 556*, (TENN). **Virginia:** Augusta Co. Oak woods, 3 mi. SE of Moscow, 24 August 1927, *Wherry, E.T. 3356*, (PH). Caroline Co. Roadside, just NE of main entrance to Fort A.P. Hill on U.S. 301, near pond, 17 July 1984, *Bradley, T. 20783*, (FARM). Grayson Co. Four miles NE of Konnarock, steep slope in mixed woods, about 3000 ft., 31 August 1949, *Mikula, B. 4115*, (WILLI). Rockbridge Co. Appalachian Trail between hickory stand and Matts Creek, 16 August 1976, *Ramsey, G.W. 22589*, (VPI); Appalachian Trail between hickory stand and Matts Creek, 16 August 1976, *Ramsey, G.W. 22589*, (LYN). Rockingham Co. Along the trail on Hone Quarry Ridge near the Big Hollow Trail intersect. George Washington National Forest, 25 July 1974, *Roe, G.F. 549*, (WILLI). Spotsylvania Co. Roadside, .1 mi NE of Rte. 624 on Rte. 612, NE of Payne's Store, 26 July 1979, *Bradley, T. 15917*, (FARM). Stafford Co. Near Aquia Station at top of gravelly bluff, 28 July 1973, *Stevens, C.E. 7514*, (VPI). **Washington, DC:** Along Blair Road near Lamond, D.C., 29 July 1917, *VanEseltine, G.P. 1453*, (BH); 11 August 1897, *Steele, E.S. sn*, (MIN). **West Virginia:** Preston Co. On piles of sand at edge of dry woods at old sandstone quarry, .05-.15 mi. SSW, St. Rt. 7 and .2 mi. NW BM 2565, ca. 1 mi. E of Hopemont, Oakland Quad., 18 July 1985, *Cusick, A.W. 24493*, (NCU). Roanoke Co. Sandy soil near summit of Poor Mt., ca. 4.5 mi. SW of Singer P.O., 18 July 1946, *Wood, C.E. 6262*, (GH). Smyth Co. On slopes and summit of the Iron Mountains at Skull Gap, 11 August 1892, *Small, J.K. sn*, (NY). **Wisconsin:** Jackson Co. Dry upland woods bordering east branch of Indian Grave Creek, T21N R3W Sec. 22, Brockway Twp., 20 July 1958, *Hartley, T.G. 4762*, (MIN).

REFERENCES

- Alexander, S. N., Hayek, L.-A. C., & Weeks, A. (2012). A subspecific revision of North American saltmarsh mallow *Kosteletzkya pentacarpos* (L.) Ledeb. (Malvaceae). *Castanea*, 77(1), 106–122.
- Altekar, G., Dwarkadas, S., Huelsenbeck, J. P., & Ronquist, F. (2004). Parallel Metropolis-coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics*, 20, 407–415.
- Ash, R. B. (2013). The Multivariate Normal Distribution. In *Lectures on Statistics*. Urbana, IL: University of Illinois, Dept. of Mathematics. Retrieved from <http://www.math.uiuc.edu/~r-ash/Stat/StatLec21-25.pdf>
- Baldwin, B. G. (1992). Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. *Molecular Phylogenetics and Evolution* 1, 1(1), 3–16.
- Bennett, J. R., & Mathews, S. (2006). Phylogeny of the parasitic plant family Orobanchaceae inferred from Phytochrome A. *American Journal of Botany*, 93(7), 1039–1051.
- Brown, M. B., & Forsythe, A. B. (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association* 1, 69, 264–267.
- Brunell, M. S., & Whitkus, R. (1998). Assessment of morphological variation in *Eriastrum densifolium* (Polemoniaceae): implications for subspecific delimitation and conservation. *Systematic Botany*, 23(3), 351–368.
- Cantlon, J. E., Curtis, E. J. C., & W.M., M. (1963). Studies of *Melampyrum lineare*. *Ecology*, 44(3), 466–474.
- Chmielewski, J. G. (1997). A taxonomic revision of the *Antennaria media* (Asteraceae: Inuleae) polyploid species complex in western North America. *Brittonia*, 49(3), 309–327.
- Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest2: more models, new heuristics and parallel computing. *Nature Methods*, 9(8), 772.
- Dugan, K. A., Lawrence, H. S., Hares, D. R., Fisher, C. L., & Budowle, B. (2002). An improved method for post-PCR purification for mtDNA sequence analysis. *Journal of Forensic Sciences*, 47, 1–8.

- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792–1797.
- Enquist, B. J. (2002). Universal scaling in tree and vascular plant allometry: toward a general quantitative theory linking plant form and function from cells to ecosystems. *Tree Physiology*, 22, 1045–1064. <http://doi.org/10.1093/treephys/22.15-16.1045>
- Fritsch, P. W., & Lucas, S. D. (2000). Clinal variation in the *Halesia carolina* complex (Styracaceae). *Systematic Botany*, 25, 197–210.
- Gastwirth, J. L., Gel, Y. R., Wallace Hui, W. L., Lyubchich, V., Miao, W., & Noguchi, K. (2013). *lawstat: An R package for biostatistics, public policy, and law* (R package). Retrieved from <http://cran.r-project.org/package=lawstat>
- Gene Codes Corporation. (n.d.) Sequencher®. Ann Arbor, MI: Gene Codes Corporation. Retrieved from <http://genecodes.com>
- Gibson, W. (1993a). Selective advantages to hemi-parasitic annuals, genus *Melampyrum*, of a seed-dispersal mechanism involving ants: I. Favorable nest sites. *Oikos*, 67, 334–344.
- Gibson, W. (1993b). Selective advantages to hemi-parasitic annuals, genus *Melampyrum*, of a seed-dispersal mechanism involving ants: II. Seed-predator avoidance. *Oikos*, 67, 345–350.
- Glass, G. V., & Hopkins, K. D. (1996). *Statistical Methods in Education and Psychology*. Needham Heights, MA: Allyn & Bacon.
- Gleason, H. A., & Cronquist, A. (1991). *Manual of Vascular Plants of Northeastern United States and Adjacent Canada* (2nd Ed., p. 487). Bronx, NY: The New York Botanical Garden.
- Guindon, S., & Gascuel, O. (2003). A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Systematic Biology*, 52, 696–704.
- Hill, T., & Lewicki, P. (2007). *Statistics: Methods and Application* (Electronic). Tulsa, OK: Statsoft, Inc. Retrieved from <http://www.statsoft.com/textbook/>
- Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, 90, 773–795.
- Kephart, S., Bledsoe, K., & Sturgeon, K. (1999). Varietal relationships in *Silene douglasii* (Caryophyllaceae): morphological variability at the population level. *Systematic Botany*, 24, 529–544.

- Li, M., Wunder, J., Bissoli, G., Scarponi, E., Gazzani, S., Barbaro, E., ... Varotto, C. (2008). Development of COS genes as universally amplifiable markers for phylogenetic reconstructions of closely related plant species. *Cladistics*, 24(5), 727–745. <http://doi.org/10.1111/j.1096-0031.2008.00207.x>
- Ligges, U., & Machler, M. (2003). Scatterplot3d: an R package for visualizing multivariate data. *Journal of Statistical Software*, 8(11), 1–20.
- Lindman, H. R. (1974). *Analyses of Variance in Complex Experimental Designs*. San Francisco, CA: W.H. Freeman & Co.
- Maddison, W. P., & Maddison, D. R. (2011). Mesquite: a modular system for evolutionary analysis. Retrieved from <http://mesquiteproject.org>
- Magallon, S., Crane, P. R., Herendeen, P. S., & Magalln, S. (1999). Phylogenetic pattern, diversity, and diversification of Eudicots. *Annals of the Missouri Botanical Garden*, 86(2), 297–372.
- McNeal, J. R., Bennett, J. R., Wolfe, A. D., & Mathews, S. (2013). Phylogeny and origins of holoparasitism in Orobanchaceae. *American Journal of Botany*, 100(5), 971–983. <http://doi.org/10.3732/ajb.1200448>
- Miller, R. G. (1974). The jackknife -- A review. *Biometrika*, 61(1), 1–15.
- Minkin, J. P., & Eshbaugh, W. H. (1989). Pollen morphology of the Orobanchaceae and rhinanthoid Scrophulariaceae. *Grana*, 28(1), 1–18. <http://doi.org/10.1080/00173138909431007>
- Mwanyambo, M. L. (1996). Intraspecific variation in the African genus *Margaretta* (Asclepiadaceae: Asclepiadoideae). *Kew Bulletin*, 51, 717–728.
- Olmstead, R. G., DePamphilis, C. W., Wolfe, A. D., Young, N. D., Elisons, W. J., & Reeves, P. A. (2001). Disintegration of the Scrophulariaceae. *American Journal of Botany*, 88(2), 348–361.
- Pennell, F. W. (1935). *The Scrophulariaceae of Eastern Temperate North America* (pp. 505–516). Lancaster, PA: Wickersham Printing Company.
- Penny, K. I. (1996). Appropriate critical values when testing for a single multivariate outlier by using the Mahalanobis distance. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 45(1), 73–81.

- R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.r-project.org/>
- Rodrigo da Silva, A. (2014). *biotools: Tools for biometry and applied statistics in agricultural science* (R package). Retrieved from <http://cran.r-project.org/package=biotools>
- Schneeweiss, G. M., Colwell, A., Park, J.-M., Jang, C.-G., & Stuessy, T. F. (2004). Phylogeny of holoparasitic *Orobanche* (Orobanchaceae) inferred from nuclear ITS sequences. *Molecular Phylogenetics and Evolution*, 30(2), 465–478. [http://doi.org/10.1016/S1055-7903\(03\)00210-0](http://doi.org/10.1016/S1055-7903(03)00210-0)
- Schori, M. (2001). *The taxonomic status of the varieties of Melampyrum lineare Desr.* Cornell University.
- Swofford, D. L. (2002). PAUP*. Sunderland, Massachusetts: Sinauer Associates, Inc.
- Thorne, R. F. (1972). Major disjunctions in the geographic ranges of seed plants. *The Quarterly Review of Biology*, 47(4), 365–411.
- Tiffney, B. H., & Manchester, S. R. (2001). The Use of Geological and Paleontological Evidence in Evaluating Plant Phylogeographic Hypotheses in the Northern Hemisphere. *International Journal of Plant Sciences*, 162(S6), S3–S17.
- Tutin, T. G., Heywood, V. H., Burges, N. A., Moore, D. M., Valentine, D. H., Walters, S. M., & Webb, D. A. (Eds.). (1972). *Flora Europaea Volume 3: Diapensiaceae to Myoporaceae* (pp. 253–257). London: Cambridge University Press.
- Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (4th Edition). New York: Springer.
- Weakley, A. S. (2012). *Flora of the Southern and Mid-Atlantic States* (p. 965). Unpublished.
- Weihs, C., Ligges, U., Karsten, L., & Raabe, N. (2005). klaR Analyzing German Business Cycles. In D. Baier, R. Decker, & L. Schmidt-Thieme (Eds.), *Data Analysis and Decision Support* (pp. 335–343). Berlin: Springer-Verlag. Retrieved from <http://cran.r-project.org/web/packages/klaR/klaR.pdf>
- Wen, J., & Zimmer, E. a. (1996). Phylogeny and biogeography of *Panax* L. (the ginseng genus, Araliaceae): inferences from ITS sequences of nuclear ribosomal DNA. *Molecular Phylogenetics and Evolution*, 6(2), 167–77. <http://doi.org/10.1006/mpev.1996.0069>

- Wikström, N., Savolainen, V., & Chase, M. W. (2001). Evolution of the angiosperms: calibrating the family tree. *Proceedings of the Royal Society*, 268(1482), 2211–20. <http://doi.org/10.1098/rspb.2001.1782>
- Wolfe, A. D., & Randle, C. P. (2001). Relationships within and among Species of the Holoparasitic Genus *Hyobanche* (Orobanchaceae) Inferred from ISSR Banding Patterns and Nucleotide Sequences. *Systematic Botany*, 26(1), 120–130.
- Wolfe, A. D., Randle, C. P., Liu, L., & Steiner, K. E. (2005). Phylogeny and biogeography of Orobanchaceae. *Folia Geobotanica*, 40, 115–134.

BIOGRAPHY

Karoline Oldham received her Bachelor of Science in 2011 from Virginia Tech, where she first became interested in studying and working with plants. Since then, she has completed an internship with the Smithsonian Natural History Museum's Botany Department and with NatureServe's Botany Department, and she looks forward to continuing her work in the botanical world after she finishes her Master's of Science at George Mason University in 2015.