1-1-2012

Insights into the Taxonomy and Evolution within an Orchid, Platanthera Dilatata, based on Orphometrics and Molecular Markers

Binaya Adhikari

Insights into the taxonomy and evolution within an orchid, *Platanthera dilatata*, based on morphometrics and molecular markers

By

Binaya Adhikari

A Thesis
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Masters of Science
in Biological Sciences
in the Department of Biological Sciences

Mississippi State, Mississippi

August 2012

Insights into the taxonomy and evolution within an orchid, *Platanthera dilatata*, based on morphometrics and molecular markers

By

Binaya Adhikari

Approved:

Lisa E. Wallace Assistant Professor of Biological Sciences (Director of Thesis)	Brian Counterman Assistant Professor of Biological Sciences (Committee Member)
Mark E. Welch Assistant Professor of Biological Sciences (Committee Member)	Gary N. Ervin Professor of Biological Sciences (Graduate Coordinator)
R. Gregory Dunaway Professor and Interim Dean College of Arts & Sciences	

Name: Binaya Adhikari

Date of Degree: August 11, 2012

Institution: Mississippi State University

Major Field: Biological Sciences

Major Professor: Lisa E. Wallace

Title of Study: Insights into the taxonomy and evolution within an orchid, *Platanthera dilatata*, based on morphometrics and molecular markers

Pages in Study: 61

Candidate for Degree of Masters of Science

Documenting biodiversity, at and below the species level, is a persistently challenging task for biologists. Poor understanding of biodiversity may lead to incorrect interpretations of observed variation. The underlying basis of variation can be understood by quantifying multiple sources of information. Nine morphometric characters and plastid DNA sequences (2511 bps) were quantified in a highly variable orchid species, *Platanthera dilatata*, to evaluate taxonomy of the three named varieties and to understand patterns of evolution. Three morphological groups, identified in a cluster analysis, were distinct in multiple floral traits. Additionally, the three clusters were consistently genetically divergent as indicated by infrequent haplotype sharing, significantly different haplotype frequencies, and significant values of the genealogical sorting index. This level of genetic divergence suggests three species rather than varieties in this complex. The divergent floral morphologies suggest that pollinator-mediated selection may be a driving factor for speciation in this complex.

DEDICATION

To my daughter Nayana Adhikari and my wife Bipana Adhikari who have compromised in several ways to provide me time and environment for my study, and my parents, Bijaya P. and Sabitra Adhikari, who always inspire me.

ACKNOWLEDGEMENTS

I would like to thank Dr. Lisa Wallace, my major advisor, for providing me an opportunity to work in her lab and closely supervising me throughout the entire process of this research. She provided me help and suggestions every time I needed them. She also provided plant samples used in this study. I would like to thank Dr. Mark Welch and Dr. Brian Counterman, my committee members, for providing valuable suggestions on my research and the thesis. I would like to thank my friends and lab mates, Christopher Doffitt, Giuliano Colosimo, Gregory Wheeler, Hanna Dorman and Lavanya Challagundla for their cooperation and help during my lab work. I also thank Dr. Shannon Datwyler for providing samples from California. I am grateful to the Department of Biological Sciences for providing me with a Teaching Assistantship. This research was funded by Department of Biological Sciences and Research Initiation Program of Office of the Research, Mississippi State University to my advisor. Last but not the least; I will always remember my brothers, Bikal and Bipin Adhikari, and my sister, Sapana Pandit for their inspiration, and their support in my difficulties.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Recognizing species in the earliest stages of evolution is crucial to conserve evolutionarily significant units (Moritz, 1994; Ryder, 1986; Waples, 1991) and aids in our understanding of speciation (Wiens, 2004). Poor understanding of biological diversity may lead to erroneous taxonomy, which is detrimental to conservation efforts (Isaac et al., 2004; Haig et al., 2006). Phenotypic variation is the primary means by which species have been identified and named (Cronquist, 1978) and is expected to reflect evolutionary changes associated with speciation. However, delimiting species solely based on phenotypic variation may be problematic because phenotypic variation can be the result of different processes, such as selection, drift, and plasticity (Grant, 1963). For example, plastic changes that are believed to not have stable genetic basis are not expected to track speciation. Similarly, stochastic variation due to genetic drift may be confounded with variation associated with directional changes capable of leading to speciation (Coyne and Orr, 2004), although in some cases genetic drift may lead to the evolution of new species (Coyne and Orr, 2004; Tremblay et al., 2005). The underlying basis of phenotypic variation within a system may be understood through analyses of additional independent sources of information, such as molecular variation (e. g. Hodges and Arnold, 1995; Rieseberg et al., 2003). Such studies provide a basis for evaluating existing taxonomic hypotheses.

Orchids with tremendous phenotypic diversity and adaptive capability (Dressler, 2005), provide ample opportunity to study complex pathways of diversification in plants. In this study, morphological and genetic variability were quantified and compared in an integrative way within a morphologically diverse and taxonomically challenging orchid, *Platanthera dilatata* (Pursh) Lindl. *ex* Beck, to evaluate the taxonomy proposed by Luer (1975) and to understand patterns of evolution within this system. This orchid is formally documented from western and northeastern North America, while being absent or sparse in the Midwestern U. S. (Sheviak, 2002). These plants were, earlier, described as members of genus *Habenaria* (e. g. Ames, 1910) or *Limnorchis* (Rydberg, 1901) before *Platanthera* was formally recognized as a distinct (Dressler, 1993; Smith, 1993) monophyletic (Hapeman and Inoue, 1997) genus. Currently, *P. dilatata* is classified within section *Limnorchis*, one of the five formally recognized sections in the genus *Platanthera* (Hapeman and Inoue, 1997). Section *Limnorchis* is monophyletic (Hapeman and Inoue, 1997) and comprises small white and green-flowered species. All white-flowered forms are included under *P. dilatata*. Other identifying features of *P. dilatata* are basally dilated lip, and long, slender nectar spur.

Floral morphology among the white flowered members of section *Limnorchis* is sufficiently variable that this has led to an unstable taxonomy (e. g. Ames, 1910; Luer, 1975; Rydbergh, 1901). Currently, one species, *P. dilatata,* with three varieties: *albiflora* (Cham.) Ledeberg, *dilatata* (Pursh) Lindl. *ex* Beck, and *leucostachys* (Lindl.) Luer, is recognized (Luer, 1975; Schrenk, 1978; Sheviak, 2002; Wallace, 2003). The varieties are usually identified by having short (var. *albiflora*), medium (var. *dilatata*), or long (var. *leucostachys*) spurred flowers. In some populations the varieties can be difficult to

diagnose because morphological traits intergrade (Sheviak, 2002). Additionally, Wallace (2003) showed genetic divergence among the named varieties but did not identify fixed genetic markers diagnostic of the varieties. Although Luer (1975) mentioned some variation in the geographic distributions of the varieties, their ranges overlap in western North America (Sheviak, 2002). Thus, the taxonomy of this complex is still debated and hypotheses for the observed morphological and genetic divergence remain untested.

Similar to *P. dilatata*, high morphological variability occurs in other species of *Platanthera*. Thus, this genus is an excellent study system for exploring evolution of floral diversity and its relationship to speciation patterns (Hapeman and Inoue, 1997). Floral radiation within *Platanthera* is hypothesized to be the result of pollinator-mediated selection (van der Pijl and Dowson, 1966). In particular, spur length is believed to evolve in response to pollinator morphology (Maad and Alexandersson, 2004; Robertson and Wyatt, 1990) because the length of the nectar spur determines whether a particular pollinator can effectively transfer pollen among flowers with a particular morphology (see Hapeman and Inoue, 1997; Nilsson, 1988). Floral morphological variants noticed in *P. dilatata* may, thus, be specialized to utilize different sets of effective pollinators (*sensu* Stebbins et al., 1970) given that different pollinators have been recorded among different morphological forms within *P. dilatata* (Boland, 1993; Kipping, 1971).The main aim of this study was to evaluate intraspecific taxonomy of *P. dilatata* by quantifying morphological and molecular variation. If morphological variation reflects evolutionary divergence, then I expect to find evidence of concordant genetic divergence. The resulting data were also used to develop a better understanding of population variability across the geographic range of the species and to develop hypotheses of possible

evolutionary pathways that could have operated within this complex to produce the observed morphological and genetic variation.

References

Ames, O. 1910. Orchidaceae: Illustrations and studies of the family Orchidaceae, Vol. IV: The Genus *Habenaria* in North America. Merrymount Press, Boston, MA.

Boland, J. T. 1993. The floral biology of *Platanthera dilatata* (Pursh.) Lindl. (Orchidaceae). M. S. Thesis, Memorial University of Newfoundland, St. Johns, Newfoundland.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Inc., Sunderland, MA.

Cronquist, A. 1978. Once again, what is a species? *In* L. Kuntson [ed.], Biosystematics in agriculture, 3-20. Alleheld Osmun, Montclair, NJ.

Darwin, C. R. 1862. On the Various Contrivances by which British and Foreign Orchids are Fertilised by Insects. John Murray, London, England.

Darwin, C. R. 1877. The different forms of flowers on plants of the same species. John Murray, London, England.

Dressler, R. L. 1993. Phylogeny and Classification of the Orchid Family. Dioscorides Press, Portland, OR.

Dressler, R. L. 2005. How many orchid species. *Selbyana* 26:155-158.

Grant, V. 1963. The origin of adaptations. Columbia University Press, New York and London.

Haig, S. M., Beever, E. A., Chambers, S. M., Draheim, H. M., Dugger, B. D, et al. 2006. Taxonomic considerations in listing subspecies under the U. S. Endangered Species Act. *Conservation Biology* 20(6):1584-1594.

Hapeman, J. R., and K. Inoue. 1997. Plant-pollinator interactions and floral radiation in *Platanthera* (Orchidaceae). *In* T. J. Givnish, and K. J. Sytsma [eds], Molecular Evolution and Adaptive Radiation, 433-454. Cambridge University Press, Cambridge, England.

Hodges, S. A., and M. L. Arnold. 1995. Spurring plant diversification: are floral nectar spurs a key innovation? *Proceedings of the Royal Society of London, Series B* 262:343-348.

Isaac, N. J. B., J. Mallet, and G. M. Mace. 2004. Taxonomic inflation: its influence on macroecology and conservation. *Trends in Ecology and Evolution* 19:464-469.

Kipping, J. L. 1971. Pollination studies of native orchids. M.S. Thesis, San Francisco State College, San Francisco, CA.

Luer, C. A. 1975. The Native Orchids of the United States and Canada, Excluding Florida. New York Botanical Garden, New York, NY.

Maad, J., and R. Alexandersson. 2004. Variable selection in *Platanthera bifolia* (Orchidaceae): phenotypic selection differed between sex functions in a drought year. *Journal of Evolutionary Biology* 17:642-650.

Mallet, J. 1995. A species definition for modern synthesis. *Trends in Ecology and Evolution* 10:294-299.

Moritz, C. 1994. Applications of mitochondrial DNA analysis in conservation: a critical review. *Molecular Ecology* 3:401-411.

Nilsson, L. A. 1978. Pollination ecology of *Epipactis palustris* (Orchidaceae). *Botaniska Notiser* 131: 355-368.

Nilsson, L. A. 1983. Mimesis of bellflower (*Campanula*) by the red helleborine orchid *Cephalanthera rubra. Nature* 305:799-800.

Nilsson, L. A. 1988. The evolution of flowers with deep corolla tubes. *Nature* 334:147-149.

Rieseberg, L. H., S. A. Church, and C. L. Morjan. 2003. Integration of populations and differentiation of species. *New Phytologist* 161:59-69.

Robertson, J. L, and R Wyatt 1990. Evidence for pollination ecotypes in the yellow-fringed orchid, *Platanthera ciliaris. Evolution* 44:121-133.

Rydberg, P. A. 1901. The American species of *Limnorchis* and *Piperia*, north of Mexico. *Bulletin of the Torrey Botanical Club* 28:605-643.

Ryder, O. A. 1986. Conservation and systematics: The delimitation of sub-species. *Trends in Ecology and Evolution*. 1:9-10.

Schrenk, W. J. 1978. North American platantheras: evolution in the making. *American Orchid Society Bulletin* 47: 429-437.

Sheviak, C. J. 2002. *Platanthera*. *In* Flora of North America Editorial Committee [ed.], Flora of North America North of Mexico, vol. 26, 551-571. Oxford University Press, Oxford, England.

Sokal, R. R., and P. H. A. Sneath. 1963. Principles of Numerical Taxonomy. San Francisco: Freeman.

Sokal, R. R., and J. Crovello 1970. The biological species concept: a critical evaluation. *American Naturalist* 104:127-153.

Stebbins Jr, G. L. 1970. Adaptive radiation of reproductive characteristics in Angiosperms, I: Pollination mechanisms. *Annual Review of Ecology and Systematics* 1:307-326.

Wallace, L. E. 2003. An evaluation of taxonomic boundaries in *Platanthera dilatata* (Orchidaceae) based on morphological and molecular variation. *Rhodora* 105:322-336.

Waples, R. S. 1991. Pacific salmon, *Oncorynchus* spp., and the definition of a "species" under the Endangered Speceis Act. *Marine Fisheries Review* 53:11-22.

Wiens, J. J. 2004. What is speciation and how should we study it? *American Naturalist* 163:914-923.

CHAPTER II

INSIGHTS INTO THE TAXONOMY AND EVOLUTION WITHIN AN ORCHID, *PLATANTHERA DILATATA*, BASED ON MORPHOMETRICS AND MOLECULAR MARKERS

Introduction

Variations among and within species are valuable components of biodiversity and also reflect their evolutionary potential. Documenting such variability and defining formal species has been a subject of debate among systematists, evolutionists and conservationists because of the existence of numerous concepts used to define and diagnose species (see de Queiroz, 1998, 2007; Mayden, 1997; McDade, 1995; Naomi, 2011) although no single accepted species concept exists. Correct interpretation of observed variations, as well as objectivity in defining species, are both essential to ensure proper documentation and conservation of biological diversity (Haig et. al., 2006; Hey et al., 2003; Isaac et al., 2004; Sites and Marshall, 2004). While taxonomic exaggeration could potentially bias conservation attention (Isaac et al., 2004), recognizing incipient species is important to preserve crucial evolutionary units (ESUs; Moritz, 1994; Ryder, 1986; Waples, 1991) and ultimately aids in our understanding of speciation (Wiens, 2004).

Numerous concepts proposed to define species (see Mayden, 1997; Naomi, 2011) create confusion among systematists about what species are and how they arise (de

Queiroz, 2007; Sites and Marshall, 2004). However, multiple species concepts may essentially highlight various aspects of a common species concept, which is that species are 'evolutionary lineages' (de Queiroz, 2007). Species concepts, in essence, differ on whether the focus is on the processes or patterns within species (Helbig et al., 2002). For example, while the Biological Species Concept (Mayr, 2000) emphasizes the process of reproductive isolation; the phylogenetic species concept (e. g. Donoghue, 1985) focuses on the pattern of reciprocal monophyly. Similarly, the Phenetic species concept (e. g. Sokal and Sneath, 1963; Sokal and Crovello, 1970) and the Genotypic cluster concept (Mallet, 1995) advocate phenotypic and genotypic distinctiveness, respectively. Given the diversity of living organisms and multitude of possible evolutionary processes within them, a single concept may not work universally (Dayrat, 2005; Puorto et al., 2001; Sites and Marshall, 2004).

Analogous to species concepts debates, confusions also exist about the reliability and efficiency of data types (e. g. morphological and molecular) used in delineating species (e. g. Blaxter, 2004; Hebert et al., 2003; Gaston and O'Neil, 2004; Valdecasas et al., 2008). The primary data used for biological taxonomy has traditionally been phenotypic (Coyne, 1994; Hennig, 1966; see review by McDade, 1995; Sattler and Rutishauser, 1997; Wiley, 1981) because phenotypic variation is expected to reflect evolutionary changes within species (e. g. Stuessy, 2009). Phenotypic variation is also exposed to natural selection, and species may arise by differences in selective pressures (Darwin, 1859). Selective pressures may then enhance pre-zygotic isolation, thus restricting gene exchange among subgroups (see Coyne and Orr, 2004). Thus, it is expected that phenotypic differences reflect species. However, sole dependence on

phenotypic variation in taxonomy has limitations (e. g. Dayrat, 2005) because multiple and unrelated processes, such as natural selection, drift, plasticity and adaptive convergence may result in similar phenotypes (e. g. Coyne and Orr, 2004; Grant, 1963; Wiens et al., 2003) without reflecting speciation. Given the multiple pathways by which phenotypic variation may be formed, evidence from additional independent sources are necessary to evaluate species hypotheses obtained from phenotypic data (e. g. Duminil and Michele, 2009; Hodges and Arnold, 1995; Dayrat, 2005; Padial et al., 2012; Shaffer and Thomson, 2007).

Molecular markers, although widely adopted and appreciated in systematics (e. g. Blaxter, 2004; Hebert et al., 2003; Gaston and O'Neil, 2004), have limits (e. g. Thorpe et al., 1996). First, different regions of the genome evolve at different rates (e. g. Shaw et al., 2007; Wolfe et al., 1987); thus choice of markers may heavily affect genetic clustering or gene tree topologies, which may or may not reflect species status and evolutionary relationship among lineages. Also, organeller genomes, which are the most favored markers for ease of use, are uniparentally inherited and non-recombining. The effectiveness of all molecular markers may be limited by processes like introgressive hybridization and homoplasy, which can potentially conceal the evolutionary/taxonomic signals (Garcia et al., 2009; van Oppen et al., 2000). Thus, no such data type, so far, has been found to be universally superior (Valdecasas et al., 2008), and consequently, systematists recommend using multiple independent lines of evidence to test their hypotheses (e. g. Dayrat, 2005; Padial et al., 2010).

In the current study, morphological and molecular variation was investigated within a morphologically diverse, reward-providing orchid species, *Platanthera dilatata*

(Pursh) Lindl. *ex.* Beck, to evaluate intraspecific taxonomy and evolutionary divergence. *Platanthera dilatata*, also known as the 'white bog orchid', is a white-flowered species of section *Limnorchis*, which is one of the five sections of the large and diverse genus *Platanthera* (Hapeman and Inoue, 1997). *Platanthera dilatata* is a diploid (2n=42; Sheviak, 2002; Sheviak and Bracht, 1998) perennial herb and provides nectar as reward to its pollinators. The species is distributed in western and northeastern North America (Sheviak, 2002) and is sparse or absent from the Midwestern United States and adjacent Canadian provinces, possibly due to low abundance of open mesic habitats (Sheviak, 2002). *Platanthera dilatata* has fleshy roots, a hollow stem with few to several leaves, and a long raceme with sparsely to densely crowded flowers, which are individually ca. 6-12 mm long and ca. 10-20 mm wide. The flower typically has a basally dilated lip and a slender nectar spur extending from the back of the flower. In addition to nectar, the flowers also produce fragrance.

Taxonomic treatment of the white-flowered plants within section *Limnorchis* has long been controversial (Ames, 1910; Luer, 1975; Rydberg, 1901; Schrenk, 1978; Sheviak, 2002) owing to the tremendous morphological variability that mainly occurs in the size and shape of the floral parts, and even in the floral fragrance (Sheviak, 2002; Wallace, 2003). However, no character than spur length is known to vary so drastically with estimates ranging from just 2 mm to 20 mm (Sheviak, 2002). This variation has been treated variously by past workers who proposed different numbers of specific or intraspecific taxa based exclusively on morphological characters (see Ames, 1910; Luer, 1975; Rydberg, 1901; Schrenk, 1978; Sheviak, 2002). For example, Rydberg (1901) considered white-flowered members in two groups, *Dilatatae* and *Leucostachyae*, and

recognized nine species: Group *Dilatatae* being described as having "spur length equaling or slightly exceeding the lip" while group *Leucostachyae* as "spur from one-quarter to two-thirds longer than the lip". On the other hand, Luer (1975) combined all white-flowered members into a single species, *Platanthera dilatata*, and recognized three varieties, *albiflora* (Cham.) Ledeberg, *dilatata* (Pursh) Lindl. *ex* Beck, and *leucostachys* (Lindl.) Luer, based on the relative lengths of spurs and lips: var. *albiflora* with spur shorter than the lip, var. *dilatata* with spur nearly equaling the lip and var. *leucostachys* with spur from one and a half to two times longer than the lip. Luer's (1975) treatment is adopted by Sheviak (2002) in the Flora of North America. Although Luer (1975) and Sheviak (2002) have both noted some variation in the geographic distributions of the varieties, their ranges overlap in some areas in western North America, thus making it unclear if geographic boundaries really exist in this complex. Attempts have also been made to quantify variation across the range of this species. In a recent study based on populations from the northern Rockies and eastern North America, Wallace (2003) compared morphological and genetic variation within *P. dilatata* and suggested that there was genetic divergence among the varieties, although some overlap was found among them, thus, making it difficult to identify the varieties as genetically distinct. Also, all previous studies have described spur and lip lengths to be the most important identifying characters of the varieties. However, the documented sizes of spur and lip lengths overlap among varieties, and are not uniformly described across studies because variation is continuous rather than discrete (Table 1.1). Thus, previous studies have not sufficiently determined if there are three distinct lineages within *P. dilatata*, and if evolution has occurred in a particular direction, and in association with certain forces.

The major aim of this study was to assess the taxonomic hypothesis proposed by Luer (1975) and Sheviak (2002) by quantifying morphological and genetic variation using a wide geographic sampling of populations. The study addressed the following questions: i) Are there three distinct species within *P. dilatata* that can be defined morphologically and genetically? ii) Is morphological and/or genetic variation within this species concordant with each other and with geographic distance? iii) What are the evolutionary relationships and direction of evolution among the morphologically different units within *P. dilatata?*

To evaluate a species hypothesis a species concept/criteria that would be most appropriate for the system needs to be determined (Sites and Marshall, 2004). Previous studies (Luer, 1975; Sheviak, 2002; Wallace, 2003) in *P. dilatata* have indicated that the three named varieties overlap in morphology (Sheviak, 2002; Wallace, 2003), geography (Luer, 1975; Sheviak, 2002) and molecular markers (Wallace, 2003) indicating young divergence. In cases of recent divergence, stringent criteria such as complete reproductive isolation (Biological Species Concept; Mayr, 2000) or reciprocal monophyly (Phylogenetic Species Concept; *sensu* Donoghue, 1985) may not effectively capture incipient species (Coyne and Orr, 2004) because these patterns may be expected only in older divergences and towards the final stages of speciation (de Queiroz, 2007). Additionally, these species concepts are criticized for being impractical and for making unnecessary assumptions about species and the process of speciation (Mallet, 1995). For example, the Phylogenetic Species Concept assumes that gene trees equal species trees, which may not always be true (Maddison 1997; Carstens and Knowles 2007). Also, the

Biological Species Concept requires strict reproductive isolation which is not only hard to achieve in natural populations but is also difficult to test (Mallet, 1995).

In order to document species more practically, alternative concepts, such as the Genotypic Cluster Criterion (GCC), have been proposed (Mallet, 1995). GCC is considered to be an extension of Phenetic species concept (PSC, Sokal and Crovello, 1970) the latter being based on distinctive phenotypic clustering. Moreover, GCC also emphasizes on genetic distinctiveness and defined species as: "a morphologically and genetically identifiable clusters of individuals that can co-exist with other similar clusters with a few or no intermediates" (Mallet, 1995). GCC actually aims to identify species based on morphological and genetic gaps and not by the processes (e. g. reproductive isolation, phylogeny, cohesion) that govern these gaps. The gaps are characterized by no or low frequency of hybrids which experience strong selection (Mallet, 1995). Moreover, GCC can accommodate gene flow, selection, mutation and genetic drift (Mallet, 1995). Thus, under the definition of GCC, if the groups of individuals/populations have similar distributions but still maintain morphological and genetic identity, then they can be considered distinct species.

Mallet (1995) claims that the GCC is one of the widely practiced species criteria with its history being extended back to Darwin (1859). This criterion has also been successfully employed by many recent studies (e. g. Noble et al., 2010; Pettengill and Neel, 2011; Reeves and Richards, 2011; Verbruggen et al., 2005) and was adopted in the present study. The hypothesis was that the three previously proposed varieties within *P. dilatata* would actually form three species. Thus, under the criterion of GCC, it was

expected that three groups consistently identifiable morphologically and genetically would be obtained within *P. dilatata*.

Materials and methods

Morphological analyses

A total of 92 individuals from 24 populations covering the distribution of *P. dilatata*, including 15 populations previously examined by Wallace (2003) (Table 2.2; Fig. 2.1), were used to quantify morphological variation. Voucher specimens for populations are deposited in the herbarium at The Ohio State University Herbarium (OS) or Mississippi State University (MISSA) (Table 2.2). Measurements were made on flowers preserved in FAA (45% ethyl alcohol, 45% water, 5% glacial acetic acid, and 5% formalin). All the flowers used for measurements were fully matured and were taken from the middle of the inflorescence to avoid potential age and placement biases. Measurements were made on at least five individuals per population when available. A mean value based on three flowers per inflorescence was used for each individual in the data set. A total of 9 floral morphological characters that were found to be significantly different among the varieties in Wallace (2003) were considered in this study. These include: lengths and widths of the dorsal sepal, lateral sepal, lateral petal and lip, and spur length. Although anther width was also found to vary among the varieties by Wallace (2003), this character could not reliably be measured in all samples and was not included here. All measurements were made manually under a dissecting microscope using a miniscale (BioQuip, Rancho Dominquez, CA, USA). Lengths of the sepals, lateral petals, and lip were measured form the point of attachment to the tip. Widths were

measured at the points with maximum width. Spur length was measured from the opening to the tip along the full curvature.

Pair-wise correlation analyses of all the morphological characters were done using SPSS v 18.0 (IBM SPSS Company, NY, USA) to determine if these characters were variable in a correlated manner ($p \leq 0.05$). A UPGMA (Unweighted Pair-Group Method using Arithmetic average) hierarchical clustering analysis was used to identify clusters of related individuals based on all morphological characters. A pairwise morphological distance matrix was generated using the Euclidean distance method in PASSaGE v 2 (Rosenberg and Anderson, 2011). UPGMA was conducted using this distance matrix in PAUP* v 4.0b10 (Swofford, 2002). The tree was visualized in FigTree v1.3.1 (Rambaut, 2012). Three clusters identified in the UPGMA (hereafter, cluster I, cluster II, cluster III) were considered as natural morphological groupings within the species and served as the basis for subsequent morphological and genetic analyses. Two individuals not clearly grouped with any of the clusters were excluded from further analyses. Kurskal-Wallis non-parametric tests (Zar, 1996) were done using SPSS v 18.0 (IBM SPSS Company, NY) to determine which of the morphological characters varied significantly ($p \leq 0.05$) among the three morphological groups identified in the cluster analysis. Characters found to be significantly variable ($p \leq 0.05$) among clusters were further analyzed using Dunn's multiple comparisons tests (Zar, 1996).

Molecular analyses

Fresh leaf samples were collected from a total of 78 individuals belonging to 26 populations. These samples are derived from the same populations used in the morphological analysis plus two additional populations from Oregon and California

(Table 2.2; Fig. 2.1). From 1-5 individuals per population were used in molecular-genetic analyses. Two individuals from one population of *Platanthera aquilonis* were sampled as outgroups. Leaf samples were preserved on ice in the field and stored at -80°C in the lab or stored in silica gel before DNA was extracted. Total DNA was extracted from leaf samples using either a CTAB protocol (Doyle and Doyle, 1987) or the DNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA).

A total of six chloroplast markers were used in the study: the intron of *rpL16,* the intergenic spacers of *psaI-accD* and *trnV-ndhC,* and three microsatellite regions (cpSSRs) contained in the intergenic spacers of *atpF-atpH, psbA-trnK,* and *psbC-trnS* (Table 2.3). The PCR reactions for *rpL16, trnV-ndhC,* and *psaI-accD* were done in 25 μL volume. Each PCR contained 2 μL template DNA, 0.2 μM of each primer, 2.0 mM $MgCl_2$, 160 μM dNTP, 1X GoTaq® Flexi buffer, 0.5 U colorless GoTaq® DNA polymerase (Promega, Madison, WI, USA), 0.1X BSA and 12.65 μL sterile distilled water. The thermal cycler program followed Shaw et al., (2007) and consisted of denaturation at 80^{o}C for 5 min; and 30 cycles of denaturation at 95^{o}C for 1 min, annealing at 50^{o}C for 1 min, a ramp of 0.3^{o}C/s to 65^{o}C, extension at 65^{o}C for 4 min; and a final extension at 65^{o}C for 5 min. Primers for the cpSSRs were developed from the chloroplast genome of *Phalaenopsis aphrodite* subsp. *formosa* (GenBank Accession AY916449). The chloroplast sequence of *P. aphrodite* was examined for perfect repeats of at least 10 nucleotides long (mononucleotide repeats) or 14 nucleotides long (dinucleotide repeats) using Microsatellite Repeats Finder (Bikandi, 2010). Primers expected to amplify a fragment of 100-600 bp in length and that were anchored in exon regions were developed for 12 loci using Primer3 (Rozen and Skaletsky, 2000) and

subsequently tested in a sample of 10 individuals. From the 12 cpSSR loci, six regions were found to be variable. The three most variable loci were used for this study (Table 2.3). The PCR reactions for cpSSR regions were done in 10 uL volume. Each PCR contained 1.0 μL template DNA, 0.25 μM each primer, 2.5 mM $MgCl_2$, 160 μM dNTP, 1X GoTaq® Flexi buffer, 0.5 U colorless GoTaq® DNA polymerase (Promega, Madison, WI, USA) and 4.6 μL sterile distilled water. The thermal cycler program consisted of denaturation at 94°C for 2 min; 30 cycles of denaturation at 94°C for 1 min, annealing at 52°C for 1 min, extension at 72°C for 1 min; and a final extension at 72°C for 7 min. A negative control, lacking only template DNA, was included with each set of reactions to check for contamination.

Successful amplification of each PCR product was determined by running a small amount of the product on 1.5% agarose TBE gels. PCR products were cleaned using an enzyme mixture [0.25 μL each of Antarctic phosphatase, Antarctic phosphatase buffer (10X) and Exonuclease I (New England BioLabs® Inc.), and 3.25 μL of distilled water] to remove excess primers and dNTPs. Four μL of this mixture was added to each PCR product, and this reaction was incubated at 37°C for 15 min and then at 80°C for 15 min. All individuals were sequenced at all six loci using the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA). For *rpL16, trnV-ndhC,* and *psaI-accD* sequencing was done in both the forward and the reverse directions while for the microsatellite regions (*atpF-atpH, psbA-trnK,* and *psbC-trnS*) sequencing was done only in the forward direction because they were less than 300 bp in size and full sequence could be obtained with a single primer. Each 10 μL sequencing reaction contained 1.0 μL of PCR product, 0.3 μM of each primer, 0.875 X sequencing buffer, 0.5

μL Big Dye (v. 3.1) and 4.4 μL sterile distilled water. The thermal cycler program consisted of a denaturation step at 96°C for 1 min; 40 cycles at 96°C for 0.10 min, 50°C for 0.05 min, 60°C for 3 min and 96°C for 0.10 min. The sequencing reactions were cleaned using Sephadex™ G-50 Fine (GE HealthCare Bio-Sciences AB, Uppsala, Sweden) columns. The cleaned products were dried and sent to the DNA Lab at Arizona State University (Tempe, AZ, USA) for sequencing by capillary electrophoresis.

The raw sequences were edited in Sequencher v. 4.1 (Gene Codes, Ann Arbor, MI, USA) and aligned manually using SE-Al v. 2.0 (Rambaut, 2010). Only unambiguously aligned sequences were used in analyses. A small portion (i.e., 165bp) of sequence in *trnV-ndhC* and a region after a long T repeat in *psbC-trnS* (i.e., 131bp) were removed from the data set due to an inability to obtain high quality sequence of this region across most samples. Samples that did not amplify (two samples each for *trnV-ndhC* and *psbC-trnS*, and four samples for *psbA-trnK*) were treated as missing data. The chloroplast genome is usually non-recombining and thus functionally a single locus in most plants (Olmstead and Palmer, 1994). Thus, sequences from the six chloroplast regions were concatenated into a single cpDNA haplotype for each individual to combine information contained in all the regions for genetic analyses. Since alignment gaps (i.e., indels) in cpDNA sequences are suggested to contain important phylogenetic information at lower taxonomic levels (e. g. Simmons et al., 2001; Kelchner, 2000; Ingvarsson et al., 2003), these were coded as simple (simple indel coding, SIC) or complex (complex indel coding, CIC) indels (Simmons and Ochoterena, 2000) depending upon the compatibility with a particular analysis software. Both of these methods treat gaps as presence/absence characters; however, while SIC considers all gap positions with different 5' and 3'

termini as different presence/absence characters, CIC may treat gaps with different 5' and 3' termini as a single character if they had similar sequential evolution (Simmons and Ochoterena, 2000). SeqState v 1.4 (Muller, 2005) was used for coding gaps.

The three clusters identified in the UPGMA cluster analysis were tested for genetic divergence using a variety of analyses. Since individuals in the morphological analyses were not perfect matches to the genetically sampled individuals, populations were assigned to a cluster based on the primary morphological pattern in the population. Individuals of six populations (AK-1, CAN-1, MT-1, MT-3, VT-1, WY-1; see Table 2.2 for codes) were split between two clusters, so population assignment was based on the cluster that contained most of the individuals for that population. Basic genetic diversity metrics (i.e., number of unique haplotypes, haplotype diversity, and nucleotide diversity) were estimated for *P. dilatata* and each of the three morphological clusters in Arlequin v 3.5 (Excoffier and Lischer, 2010). Here, haplotype diversity is a measure of uniqueness of a haplotype within a group (Nei and Tajima, 1987) while nucleotide diversity is the average number of differences in nucleotides per base position between the pairs of sequences (Nei, 1987). These metrics were calculated based on concatenated sequences (Table 2.3) as well as for each of the six cpDNA regions separately (Appendix A). jModelTest 0.1.1 (Posada, 2008) was used to find the best fitting model of DNA substitution for the combined sequence data based on the corrected Akaike information criterion (AIC_C; Hurvich and Tsai, 1989). TIM1+ G was found to be the best model for the present molecular data. An Analysis of Molecular Variance (AMOVA, Excoffier et al., 1992) was conducted to determine the degree of genetic differentiation among and within the three clusters identified by morphology. AMOVA was conducted in Arlequin

v 3.5 (Excoffier and Lischer, 2010) considering mutational differences between haplotypes according to Tamura and Nei's distance (Tamura and Nei, 1993) method. This model is the closest available model in the software that matches with the substitution model suggested for the current DNA data by jModeltest. Statistical significance was determined by conducting 3,000 permutations and Φ-statistics were considered significant at $p = 0.05$. Because there were few shared haplotypes between clusters, a fixation index (F_{ST}, Wright, 1965) was also estimated to infer the amount of haplotypic variation among morphological clusters without considering mutational differences. This analysis was done based on haplotype frequency information in Arlequin v 3.5 (Excoffier and Lischer, 2010) and the significance ($p = 0.01$) was tested with 3,000 permutations.

Maximum Parsimony (MP) and Bayesian approaches were used to infer phylogenetic relationships among the haplotypes of *P. dilatata*. *Platanthera aquilonis* was used as an outgroup in both analyses. MP analysis was done using PAUP* v 4.0b10 (Swofford, 2002). The gaps in the sequences were coded using modified complex indel coding (MCIC) method (Muller, 2005) which is a simplified version of CIC but requires minimum number of assumptions on gap evolution and is practical to implement (Muller, 2005). The heuristic search was done with 100 replicates of random sequence addition while the branch swapping was done by tree bisection reconnection method. A maximum of 100 trees with a score of 1 or above were saved in each replicate. The branch support values for the nodes were assessed using 1000 replicates of bootstrap analyses using the same parameters mentioned above except that 10 replicates of random

sequence addition were conducted per bootstrap replicate. The tree was visualized in FigTree v1.3.1 (Rambaut, 2012).

Bayesian analysis was done using MrBayes v3.2.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) with general time reversible (GTR) and an invariable gamma substitution model. This is the closest available model that matches the best fitting model suggested by AIC_C and that can be implemented in MrBayes. Alignment gaps were coded as simple indels (Simmons and Ochoterena, 2000), and the sequences and gaps were treated as unlinked loci. The option 'variable' substitution model was used for gaps. Markov Chain Monte Carlo (MCMC) algorithm was run for 1.5 million generations with two incrementally heated chains until the branches converged as indicated by an average standard deviation ≤ 0.01. One of every 100 generations of trees was randomly sampled. The first 25% of trees were discarded as burn-in while remaining trees were used to create a strict consensus tree. Posterior probabilities were used to assess support for clades in the resulting phylogeny. The tree was visualized in FigTree v1.3.1 (Rambaut, 2012).

The genealogical sorting index, *gsi* (Cummings et al., 2008) was used to estimate the amount of monophyly of each of the morphological clusters of *P. dilatata*. *Gsi* is a statistical way of quantifying the amount of exclusive ancestry of a group of individuals (leaves) in a rooted phylogenetic tree. The *gsi* ranges from 0 to 1, where 0 indicates polyphyly (no shared ancestry) of the group members and 1 indicates reciprocally monophyly (Cummings et al., 2008). The Bayesian strict consensus tree (Fig. 2.3) was used to compute the *gsi* for each of the three morphological clusters in a web-based

application (Cummings et al., 2010). Statistical significance of *gsi* for each cluster was tested using 10,000 permutations at a level of $p = 0.05$.

Given that the three clusters did not exhibit reciprocal monophyly in MP or Bayesian phylogenies and this is expected to be due to ancestral polymorphism, a species tree approach, which fits gene trees into a single species tree using a coalescent model (Kingman 1982), was used to infer phylogenetic relationships among the morphological clusters. Species tree approaches account for gene tree-species tree discrepancies observed in cases of reticulate evolution and recent divergences (Liu et al., 2008) and is, thus, expected to better illuminate phylogenetic structure not recognized by traditional gene trees topology (Liu and Pearl, 2007; Liu et al., 2008). The approach implemented through *BEAST (Heled and Drummond, 2010) was used to estimate a species tree for the three clusters. This method is does not require a specified outgroup, and other assumptions like a constant population size and a uniform species tree prior. Importantly, it estimates all the gene trees and the species tree simultaneously in a single MCMC analysis (Heled and Drummond, 2010). The software package BEAST v. 1.7.0 (Drummond et al., 2012) was used to infer the species tree based on the following run parameters: a GTR model, empirical base frequencies, a Yule species tree prior, and a random starting tree. The analysis was conducted for 50 million generations with sampling every 1000^{th} iteration. Three independent runs of this analysis were conducted to ensure that convergence occurred and the MCMC chains were mixed. LogCombiner (Drummond et al, 2012) was used to combine multiple log files from multiple runs. Tracer v. 1.5 (Rambaut and Drummond, 2007) was used to evaluate convergence of the run. All parameters had reached an effective sample size (ESS) greater than 250 within

50 million generations suggesting mixing and convergence of MCMC. Tree Annotator v. 1.7 (Drummond et al, 2012) was used to estimate the species tree the pool of gene trees after discarding 10% of the trees. The tree was visualized in FigTree v1.3.1 (Rambaut, 2012). Attempts were also made to include indel characters (as SIC) obtained from alignment gaps into the analysis; however convergence was not obtained despite multiple parallel runs. For this analysis, data were divided into separate sequences and indel (binary characters) files. The indels were obtained from sequence gaps coded by SIC method as described in the Bayesian analysis. Both the data types were combined to generate a single XML file. All the model parameters and the prior were used as specified in the analyses above except that a Stochastic Dollo model was used for indel characters. Three parallel runs were conducted with each run replicated for 100 million generations. Each of these runs yielded a very low ESS (ESS < 50.0) suggesting that there was no convergence of MCMC chains.

Correlation analyses

Mantel tests (Mantel, 1967) were conducted to determine if genetic distance among the populations of *P. dilatata* was correlated with i) morphological distance and/or ii) geographic distance, and iii) if morphological distance was correlated with geographic distance. Pairwise population genetic distances were calculated based on the concatenated cpDNA sequences using Arlequin v 3.5 (Excoffier and Lischer, 2010). Tamura-Nei (TrN) genetic distance (Tamura and Nei, 1993) and a gamma a value of 0.054 (obtained from jModelTest) were used. Pairwise population morphological distances was estimated using the software package PASSaGE v 2 (Rosenberg and Anderson, 2011). For this, a mean of each morphological character was computed for

each population. These measurements were used to compute an overall pairwise population morphological distance matrix based on Euclidean distance. Pairwise population geographic distance (i.e., Euclidean distance) was computed using geographic coordinates of the populations using software program Geographic Distance Matrix Generator v 1.2.3 (Ersts, 2012). Separate Mantel tests were conducted between i) genetic and morphological distances, ii) genetic and geographic distances, and iii) morphological and geographic distances. PASSaGE v 2 was used for all the Mantel tests. The statistical significance of each correlation was estimated based on 10,000 permutations and significance of the relationships was assessed at $p = 0.05$. Since eastern populations were geographically distantly separated from the western populations, these Mantel tests were also conducted with only western populations to see if samples from the east could have biased the comparisons.

Results

Morphological variation

A significant correlation was found between 33 out of a total of 36 pairwise tests among each of the floral characters used in this study (values not shown). Lateral petal width was not significantly correlated with dorsal sepal width, lip length and spur length. The correlation coefficients that were significant ($p \leq 0.05$) ranged from very high ($r = 0.96$, lengths of lateral sepal vs. lateral petal) to very low ($r = 0.27$, length vs. width of lateral petal). Widespread correlation among the floral traits indicates that they could be evolving in a concerted manner in *P. dilatata* thus justifying the use of combined morphological data for computing overall floral morphological divergence. Using as many characters as possible is generally recommended in morphometric analyses (Sneath

and Sokal, 1973; Stuessy, 2009). The UPGMA dendrogram based on total morphological distance (Fig. 2.2) showed three primary clusters within *P. dilatata*.

Kruskal-Wallis tests revealed that all nine morphological characters were significantly variable among the morphological clusters at the significance level of $p \leq 0.05$. Dunn's multiple comparison showed that the three clusters were significantly different from each other in the five length characters (i.e., dorsal sepal, lateral sepal, lateral petal, lip and spur), but differences in width characters were mixed (Table 2.3). For example, cluster I varied significantly from the other clusters in dorsal sepal width and lateral sepal width, but varied only from cluster III in lip width. Similarly, cluster III varied significantly from the other clusters in lip width, but varied only from cluster I in dorsal sepal width, lateral sepal width and lip width. Finally, cluster II varied from cluster I in dorsal sepal width and lateral sepal width while it varied from cluster III in lateral petal width and lip width. The three clusters exhibit significantly different spur lengths. Cluster I (Fig. 2.2) was dominated by short spurred individuals (2 - 4.1 mm, mean = 3.1) while cluster II and cluster III included individuals with medium (3.6 - 5.8 mm, mean = 4.8) and long (5.4 - 12.9 mm; mean = 8.0) spurs, respectively (Table 2.3).

Molecular variation

The six chloroplast markers produced a total of 2,511 bases of unambiguously aligned sequence. The lengths and the number of informative sites of each of the six cpDNA regions and the concatenated sequence are shown in Table 2.4. The lengths of individual cpDNA regions ranged from 59 (*psbC_trnS*) to 704 (*psaI-accD*) bases. All regions contained more than one gap position, thus indels are a major part of genetic variation within this data set. Gap positions gave a total of 74 indel characters from

simple indel coding and a total of 22 characters from modified complex indel coding method.

A total of 57 haplotypes were identified within *P. dilatata,* haplotype diversity was 0.9897 (SD = 0.0038) and nucleotide diversity was 0.00137 (SD = 0.00024) (Table 2.5). When comparing among the groups, cluster III had the highest number of haplotypes (30) and haplotype diversity (0.9784; SD = 0.0105) while cluster I had the highest nucleotide diversity (0.00211; SD = 0.0002). From the individual chloroplast region analyses (Appendix A), it was found that the number of haplotypes for the species was highest in *psaI-accD* (30 haplotypes) and *trnV-ndhC* (16) followed by *rpL16* and *psbC-trnS* (13 each). Among other regions, *psbA-trnK* and *atpF-atpH* had 12 and 11 haplotypes respectively. The haplotype diversity for the species ranged from as high as 0.9337 (SD = 0.0179) in *psaI-accD* to as low as 0.7619 (SD = 0.0393) in atpF-atpH. Similarly, the nucleotide diversity for the species ranged from 0.00188 (SD = 0.00038) in *psaI-accD* to 0.00016 (SD = 0.00015) in *atpF-atpH*. Few haplotypes were shared among the clusters. While cluster I did not share any haplotypes with the other clusters, one population of cluster II (ME-1) and one of cluster III (CAN-1), both from eastern North America, shared a haplotype. The AMOVA based on Φ statistics (Excoffier et al., 1992) revealed substantial (20.52%) and highly significant ($p \leq 0.01$) molecular variation among three morphological clusters while within-cluster variation was 79.48%. The fixation index (Wright, 1965) based on haplotype frequency among morphological clusters was also significant ($p \leq 0.01$), although the index value was lower ($F_{ST} = 0.025$) while there was high amount of haplotypic variation within the clusters (97.51%).

A Bayesian strict consensus tree (Fig. 2.3) showed haplotypes of *P. dilatata* to form a distinct monophyletic group relative to *P. aquilonis* with 100% support. However, none of the three morphological clusters formed a monophyletic clade in the tree. Even the haplotypes of eastern populations of *P. dilatata*, which are geographically distant from western populations, did not form a clade. Some structure was still observed among the haplotypes of *P. dilatata* as several small clades with moderate to high support were identifiable (e. g. clade A, B). However, none of these clades formed exclusive geographic or morphological groups within *P. dilatata.* Maximum Parsimony analysis resulted in a topologically similar consensus tree to that of Bayesian tree and is not presented here. The genealogical sorting index (*gsi*) for all three morphological clusters was highly significant. Cluster III had a higher *gsi* (0.455, $p \leq 0.01$) when compared to cluster I (*gsi* = 0.168, $p \leq 0.01$) or cluster II (*gsi* = 0.256, $p \leq 0.01$). The species tree (Fig. 2.4) also showed *P. dilatata* to be a distinct monophyletic group relative to *P. aquilonis* with strong support (pp = 100%). This analysis resolved relationships among the morphological clusters with high support values. In this species tree, cluster I, with short spur length, was not only divergent from the other two clusters, with longer spurs, but also appeared to be basal in the tree (pp = 96%). Clusters II and III are united in a clade (pp = 92%).

Correlations

Although morphologically divergent populations also exhibited genetic divergence, pairwise population morphological distances were not significantly correlated with pairwise population genetic distances as indicated by the Mantel test (r = 0.05181; p = 0.694). Similarly, population genetic distances were not significantly

correlated with geographic distances (r = -0.17054; p = 0.24). Finally, population morphological distances were not significantly correlated with geographic distances (r = -0.14064; p = 0.24447). The scatter plots of each pair of these distances are presented in Fig. 2.5. Finally, none of the comparisons involving only western populations were significant at p = 0.05 (values not reported).

Discussion

Taxonomic evaluation

The current taxonomy of *P. dilatata* recognizes three varieties that are diagnosable by differences in morphological traits and possibly geographic distributions. Results from the present study support Wallace's (2003) view that there are divergent lineages within this complex. However, it is suggested that these taxa should be elevated to the rank of species as first recognized by Rydberg (1901). The clusters that were identified in UPGMA and supported as distinct in genetic analyses are concordant with the Genotypic Cluster Criterion. It is important to note that the three groups are able to maintain their morphological and genetic identity despite shared areas of distribution (Fig 2.1). The three groups also fulfill the criteria of the Phenetic Species Concept (PSC, Sokal and Crovello, 1970) by forming distinct diagnosable groups based on overall morphological similarity. The PSC assumes that morphology reflects genetics, which is true except in cases of phenotypic plasticity and evolutionary convergence.

The three clusters identified by UPGMA are diagnosable by numerous individual morphological characters. For example, cluster I (Fig. 2.2) had small flowers with short spurs (2 - 4.1 mm, mean = 3.1 mm), lips (3.3 - 5.2 mm, mean = 4.3 mm) and lateral petals (2.1 - 4.3 mm; mean = 3.24 mm); cluster II had medium length spurs (3.6 - 5.8

mm, mean = 4.8 mm), lips (4.0 - 6.17 mm, mean = 4.97 mm), and lateral petals (3.0 - 5.0 mm; mean = 3.9 mm); and cluster III had long spurs (5.4 - 12.9 mm; mean = 8.0 mm), lips (4.9 - 8.17 mm; mean = 6.3 mm) and lateral petals (3.8 - 7.17 mm; mean = 5.2 mm). The three groups were also found to be significantly distinct genetically, as indicated by very rare sharing of haplotypes, highly significant differences in haplotype frequencies in AMOVA (Φ_{ST} = 20.52%; $p \leq 0.001$), and highly significant *gsi* values (cluster I, *gsi* = 0.168, $p \leq 0.01$; cluster II, *gsi* = 0.256, $p \leq 0.01$; cluster III, *gsi* = 0.455, $p \leq 0.01$). An absence of widespread shared haplotypes between groups suggests that hybridization is rare between them.

The three identified groups correspond to the three varieties (*albiflora, dilatata* and *leucostachys*) previously proposed in this system (Luer, 1975; Sheviak, 2002), but they may not perfectly correspond to the geographic distributions described by Luer (1975) or (Sheviak (2002). Cluster III is comprised of populations from eastern and western North America (Fig. 2.2) while corresponding var. *leucostachys* was not earlier reported from eastern North America (Luer, 1975; Sheviak, 2002). These long-spurred individuals may be the 'robust' forms noted by Luer (1975) in this region. While cluster I only contained populations from Montana and Wyoming in the present study, corresponding var. *albiflora* was previously described from a broader area in western North America (Luer, 1975; Sheviak, 2002) and in areas that were sampled in this study, such as the southern Rockies.

Among other examples that used the GCC to test species hypotheses (e. g. Noble et al., 2010; Pettengill and Neel, 2011; Reeves and Richards, 2011; Verbruggen et al., 2005), some are analogous to the present study. For example, in a study (e. g. Reeves

and Richards, 2011) of Wild North American Hops (*Humulus lupulus*, Cannabaceae), the authors raised the status of three morphological varieties to three species based on the GCC. They also found that their data met the criteria of monophyly (Donoghue, 1985) and diagnosability (Nixon and Wheeler, 1990). In a different case (Pettengill and Neel, 2011), the previously recognized species, *Agalinis acuta* (Orobanchaceae), was merged with *A. decemloba* because the former species did not form a distinct morphological and genetic cluster as defined by the GCC. There are additional studies that compared morphological and genetic variations to evaluate species or lower level taxonomic hypotheses in plants (e. g. Barrett and Freudestein, 2009; Hansen et al., 2000; Harastova-Sobotkova, 2005). However, these studies did not explicitly mention what species concept/criteria they intended to use. Also, the studies using GCC showed that there is no consistency in the analytical methods used to estimate/compare distinctiveness of potential species. The GCC, in fact, is not restricted to a particular method of analysis (except that all are quantitative) as are some other methods of "operational species criteria" (see Sites and Marshall, 2004).

Although recognizable as three species in the *P. dilatata* complex, these data suggest that they are in the early stages of divergence. The present results, in fact, suggested that *P. dilatata* could represent a good example of incomplete lineage sorting. First, haplotype variation within identified species was extensive (within group variation 97.51%; $p \leq 0.01$) meaning that the haplotypes are not yet fixed in these groups. Furthermore, the rarity of shared haplotypes among identified species indicates that gene flow is strongly restricted among them. Finally, phylogenetic structure was observed

among clusters only in those analyses that considered incomplete lineage sorting (i. e. *gsi* and species tree), but not in the gene trees that do not account ancestral polymorphism.

Patterns of evolutionary divergence

In addition to taxonomic evaluation, current data also suggested possible causes and direction of divergence within *P. dilatata*. First, since the morphological partitioning (i. e. three suggested species) was also concordant with genetic divergence, plasticity as an explanation of morphological variation in this complex is refuted. On the other hand, while a positive correlation between morphological and neutral genetic divergence would suggest genetic drift as a possible cause of morphological divergence (Hodges and Arnold, 1995; Rieseberg et al., 2003), a significant correlation was not obtained in the present study (Fig. 2.5a). This, instead, may suggest that morphology is under divergent selection, such that morphological divergence was not captured by the neutral genetic divergence. Finally, there was also a lack of correlation between genetic and geographic distances (Fig. 2.5b) which suggests that isolation by distance may not be the explanation of genetic divergence (Good and Wake, 1992) in *P. dilatata*. Thus, there, more likely, is some other cause(s) of genetic divergence in this system. The floral morphological variation within this complex may suggest that divergent selection and subsequent reproductive isolation (thus, gene flow) could be one possible explanation of genetic divergence. Finally, a lack of correlation between morphological and geographic distances also indicates that morphological divergence may be associated with factors other than spatial separation alone, and this pattern may reflect localized selection by pollinators.

Floral radiation by means of pollinator-mediated selection has long been a leading hypothesis for orchid floral diversification (e. g. Darwin, 1862, 1877; Dodd et al., 1999; Fenster et al., 2004; Kolreuter, 1761), and this has also been empirically demonstrated (e. g. Xu et al., 2011). There are several lines of evidence within the genus *Platanthera* where species have developed unique floral traits to attract specific groups of pollinators (Catling and Catling, 1991; Hapeman, 1997; Hapeman and Inoue, 1997; Inoue, 1983; Nilsson, 1978, 1983; Robertson and Wyatt, 1990). For example, the length of the nectar spur is suggested to determine whether and where on the body of a pollinator the pollinia are attached and whether a successful transfer occurs to the stigma of the next flower. Evidence of spur length evolution and associated variation in pollinators has been observed in *P. ciliaris* (Robertson and Wyatt, 1990), *P. bifolia* (Boberg and Agren, 2009; Maad, 2000; Maad and Alexandersson, 2004), *P. chlorantha* (Darwin, 1877; Nilsson, 1988), and *P. mandarionorum* (Inoue, 1986).

Interestingly, it has been documented that, within *P. dilatata*, medium-spurred forms (var. *dilatata;* cluster II) and long-spurred forms (var. *leucostachys;* cluster III) exhibit variation in pollinators. The former variety is pollinated by Noctuid moths and Skippers both during the day and night (Boland, 1993) while the latter by large nocturnal Noctuid moths (Kipping, 1971). There are no pollinator studies in the short spurred forms (var. *albiflora*, clusterI); however, it could be pollinated by short-tongued insects as in *P. stricta* (Patt et al., 1989) which also has shorter nectar spurs. Thus, pollinator-mediated selection could significantly limit gene flow among the three morphological forms within *P. dilatata* by means of pre-zygotic isolation.

Pollinator-mediated evolution in orchids is usually considered to be associated with highly specialized plant-pollinator interactions (e. g. Xu, et al., 2011; Schiestl and Schluter, 2009). However, whether evolutionary divergence can also occur in orchids like *P. dilatata* which are pollinated by a wider assemblage of pollinators (Kipping, 1971; Boland, 1993) has not been established. The evidence obtained in the present study may support the hypothesis that evolutionary divergence in flowering plants may occur even in the absence of strictly specific plant-pollinator interactions if variation occurs in the most active groups of pollinators that effect most pollinations (e. g. Olsen, 1997; Schemske and Horvitz, 1984). This phenomenon is described as the most effective pollinator principle (Stebbins et al., 1970; Mayfield et al., 2001).

If the flower morphology is under selection, then it is of interest to understand the direction of evolution of the three morphological forms as a test of directional pollinator-mediated evolution of spur length found in other species of *Platanthera*. The most common trend of spur length evolution in *Platanthera* is from shorter to longer spurs, although reversals from longer to shorter spurs are known (Hapeman and Inoue, 1997). The species tree obtained in the present study (Fig. 2.4) also suggests that short-spurred forms (cluster I) that occupied the basal part of the tree are more likely to be the ancestral while longer spurred forms (cluster II and III) are derived. It is important to note that other floral traits were also found to be variable among clusters and are correlated with spur length. This suggests that individual flower organs may evolve in a concerted way. Thus, it is likely that multiple floral traits are under selection, but this needs to be empirically tested in the future.

Conclusion

The data presented in this study indicate that there are three morphologically and genetically distinct clusters within *P. dilatata*. I suggest that these clusters be elevated to the ranks of species based on sufficient variation consistent with the Genotypic Cluster Criterion (Mallet, 1995) and Phenetic Species Concept (Sokal and Crovello, 1970). These clusters correspond to vars. *albiflora*, *dilatata* and *leucostachys* as defined by morphology but not by geography (Luer, 1975; Sheviak, 2002). Rare haplotype sharing among these groups indicates that there is restricted gene flow among them, possibly due to differences in pollinators that lead to prezygotic reproductive isolation. High haplotype diversity within each of these groups and their failure to exhibit reciprocal monophyly in gene trees indicates extensive ancestral polymorphism. This may suggest their recent divergence. Finally, high haplotypic variation within morphological clusters suggests a high degree of genetic diversity within *P. dilatata*. Also, morphological variants with intermediate spur lengths that Sheviak (2002) noted in western North America and the 'robust' forms of plants that Luer (1975) observed in eastern North America could indicate that additional cryptic species are present within this complex. Conservation of the *P. dilatata* complex throughout its range is important to explore additional cryptic species that may exist. The phylogenetic relationship revealed by the species tree indicates that short spurred forms could be ancestral form in this complex while longer spurred forms are derived. Ultimately, pollinator-mediated selection in this *P. dilatata* complex needs to be confirmed by quantifying pollinators across populations with different floral morphologies and by conducting artificial cross-pollinations to test for genetic compatibility. More inclusive genomic sampling, particularly from nuclear

regions, and more inclusive geographic sampling of populations could improve resolution of the specie trees.

Table 2.1 Ranges of spur and lip lengths for the three varieties of *Platanthera dilatata* documented in previous studies.

Morphological character	**var. *albiflora***	**var. *dilatata***	**var. *leucostachys***	**Reference**
Spur length (mm)	Up to 10	5.0 - 10.0	10.0 - 20.0	Luer (1975)
	2.0 - 7.0	4.0 - 12.0	8.0 - 20.0	Sheviak (2002)
	2.5 - 3.7	3.5 - 5.8	7.0 - 10.6	Wallace (2003)
Lip length (mm)	Longer than lip	5.0 - 10.0	5.0 - 13.3	Luer (1975)
	6.0 - 10.0	5.0 - 10.0	4.0 - 11.0	Sheviak (2000)
	3.9 - 5.0	3.3 - 6.1	4.8 - 7.1	Wallace (2003)

Table 2.2 Locations of the populations of *Platanthera dilatata* used in this study.

Population #	Population code	Latutude	Longitude	Voucher	Herbarium acronym
1	MT-1	45.416667	-112.85	L.E. Wallace 241	OS
2	MT-2	46.25	-113.783333	L.E. Wallace 238	OS
3	MT-3	46.25	-113.816667	L.E. Wallace 237	OS
4	MT-4+	48.06681	-115.92027	L.E. Wallace 298	MISSA
5	MT-5+	46.73741	-114.54655	L.E. Wallace 335	MISSA
6	MT-6	48.095308	-113.977532	L.E. Wallace 214	OS
7	MT-7	46.05	-114.283333	L.E. Wallace 236	OS
8	WY-1	44.433333	-110.7	L.E. Wallace 252	OS
9	WY-2	45.018518	-109.94585	L.E. Wallace 246	OS
10	WY-3	42.566667	-108.833333	L.E. Wallace 257	OS
11	WY-4	43.103803	-109.940529	L.E. Wallace 259	OS
12	CO-1+	40.41404	-105.81868	L.E. Wallace 275	MISSA
13	CO-2+	39.87367	-107.59344	L.E. Wallace 276	MISSA
14	CO-3+	40.23147	-105.89046	L.E. Wallace 300	MISSA
15	ME-1	45.933333	-68.366667	L.E. Wallace 228	OS
16	NY-1	43.316667	-75.033333	L.E. Wallace 220	OS
17	VT-1	44.45	-72.216667	L.E. Wallace 224	OS
18	AK-1	60.566667	-149.566667	J. V. Freudenstein2635a	OS
19	CAN-1	44.733333	-81.266667	L.E. Wallace 233	OS
20	WA-1+	47.78242	-120.83926	L.E. Wallace 320	MISSA

Populations for which only molecular data were collected are indicated by an asterisk (*) and those for which morphological data were newly collected during this study are indicated by a plus (+). The first letters of the population code indicates the state of origin. CAN = Canada. NA = information not available.

Table 2.2 (continued)

21	WA-2+	48.51815	-118.80063	L.E. Wallace 332	MISSA
22	OR-1+	43.82256	-121.79754	L.E. Wallace 325	MISSA
23	OR-2	45.310586	-117.306304	L.E. Wallace 201	OS
24	OR-3*	44.483333	-121.833333	NA	NA
25	CA-1*	39.419167	-120.246111	S. Datwyler 148	SACT
26	UT-1+	37.5911	-112.90249	L.E. Wallace 331	MISSA

Table 2.3 Summary of morphological variation among three clusters of *Platanthera dilatata*.

	Cluster I (n = 28)	**Cluster II (n = 28)**	**Cluster III (n = 34)**
DSpLn	3.7 [a]	4.07 [b]	5.04 [c]
SD	0.386	0.478	0.716
DSpWd	1.38 [a]	2.39 [b]	2.35 [b]
SD	0.218	0.461	0.732
LSpLn	4.16 [a]	4.64 [b]	6.05 [c]
SD	0.508	0.512	0.901
LSpWd	1.89 [a]	2.29 [b]	2.46 [b]
SD	0.355	0.451	0.589
LPtLn	3.24 [a]	3.91 [b]	5.21 [c]
SD	0.515	0.62	0.843
LPtWd	2.46 [ab]	2.22 [b]	2.65 [ac]
SD	0.292	0.605	0.64
LipLn	4.29 [a]	4.97 [b]	6.3 [c]
SD	0.463	0.58	0.889
LipWd	2.06 [a]	2.08 [a]	2.52 [b]
SD	0.206	0.473	0.477
SprLn	3.14 [a]	4.84 [b]	8.16 [c]
SD	0.575	0.612	1.991

The mean and standard deviation (SD) is reported for each trait. Unlike superscript letters after mean values indicate a significant difference ($p \leq 0.05$) between clusters. n = individual cluster sample size, DSpLn = Dorsal sepal length, DSpWd = Dorsal sepal width, LSpLn = Lateral sepal length, LSpWd = Lateral sepal width, LPtLn = Lateral petal length, LPtWd = Lateral petal width, LipLn = Lip length, LipWd = Lip width, SprLn = Spur length.

Table 2.4 Loci and primers used in PCR and sequencing of *Platanthera dilatata* samples and the variability obtained from each region.

Locus	Aligned length (bp)	# variable sites (PIC)	No. of indel positions	Primer sequence (5'-3')	Reference
rpL16-F71	648	9 (5)	3	GCTATGCTTAGTGTGTGACTCGTTG	Jordan et al., 1996
rpL16-R622				CCAACCCAATGAATCATTAGGATT	Les et al., 2002
trnV(uac)x2	653	0 (0)	6	GTCTACGGTTCGARTCCGTA	Shaw et al., 2007
ndhC				TATTATTAGAAATGYCCARAAAATATCATATTC	Shaw et al., 2007
psaI-75R	704	10 (9)	6	AGAAGCCATTGCAATTGCCGGAAA	Shaw et al., 2007
accD				AATYGTACCACGTAATCYTTTAAA	Shaw et al., 2007
atpF	168	1 (0)	2	CCACGGAAACAAAAGAATCG	This study
atpH				GGACTGGTCGTGGCATTAGC	This study
psbC	59	0 (0)	2	AGGAATCGATCGTGATTTGG	This study
trnS(uga				GTTCGAATCCCCCTCTCTCC	This study
psbA	279	0 (0)	3	TTCGCGTCTCTCTAAAATTGC	This study
trnK				AAACTGCAAGCACGATTTGG	This study
Total	**2511**	**20 (14)**	**22**		

The length of the aligned sequence, number of variable sites excluding gaps, parsimony-informative characters (PIC), and number of indels are included for each cpDNA regions separately and for the concatenated data.

Table 2.5 Chloroplast DNA diversity metrics for *Platanthera dilatata* and each of three morphological clusters.

Groups	**Cluster I**	**Cluster II**	**Cluster III**	***P. dilatata***
# haplotypes	13	16	30	57
Haplotype diversity	0.9608	0.9714	0.9784	0.9897
SD	0.0301	0.0234	0.0105	0.0038
Nucleotide diversity	0.00211	0.00093	0.00082	0.00137
SD	0.0002	0.00023	0.00017	0.00024

These diversity metrics are based on combined cpDNA sequence with the alignment gaps included (SD = standard deviation).

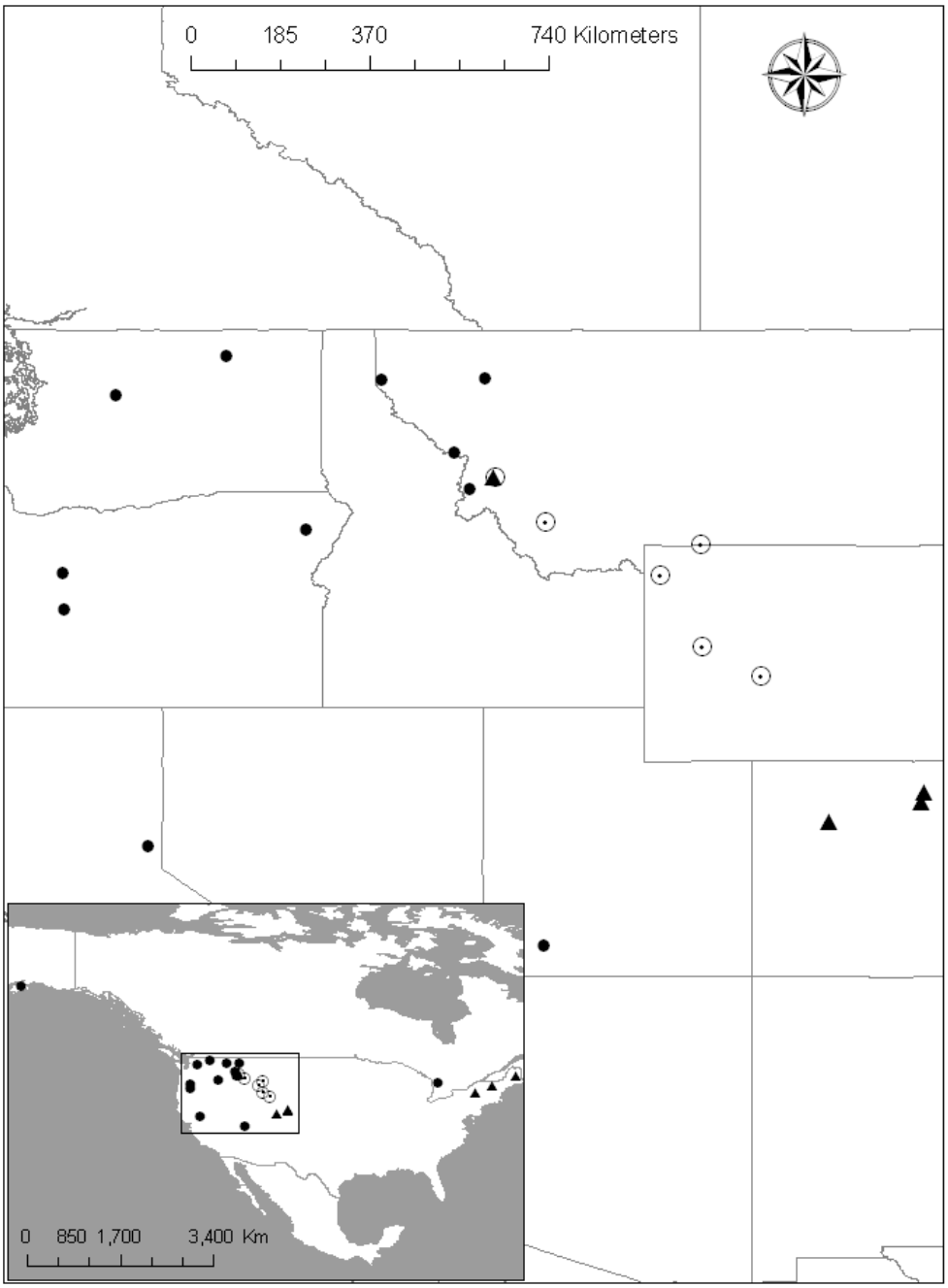


Figure 2.1 Locations of populations of *Platanthera dilatata* sampled for the present study.

Cluster I, open circle with dot; cluster II, solid triangle; cluster III, solid circle. Highlighted area in western North America is enlarged to display densely located populations. Information on population codes, geographic locations and voucher information are provided in Table 2.2.

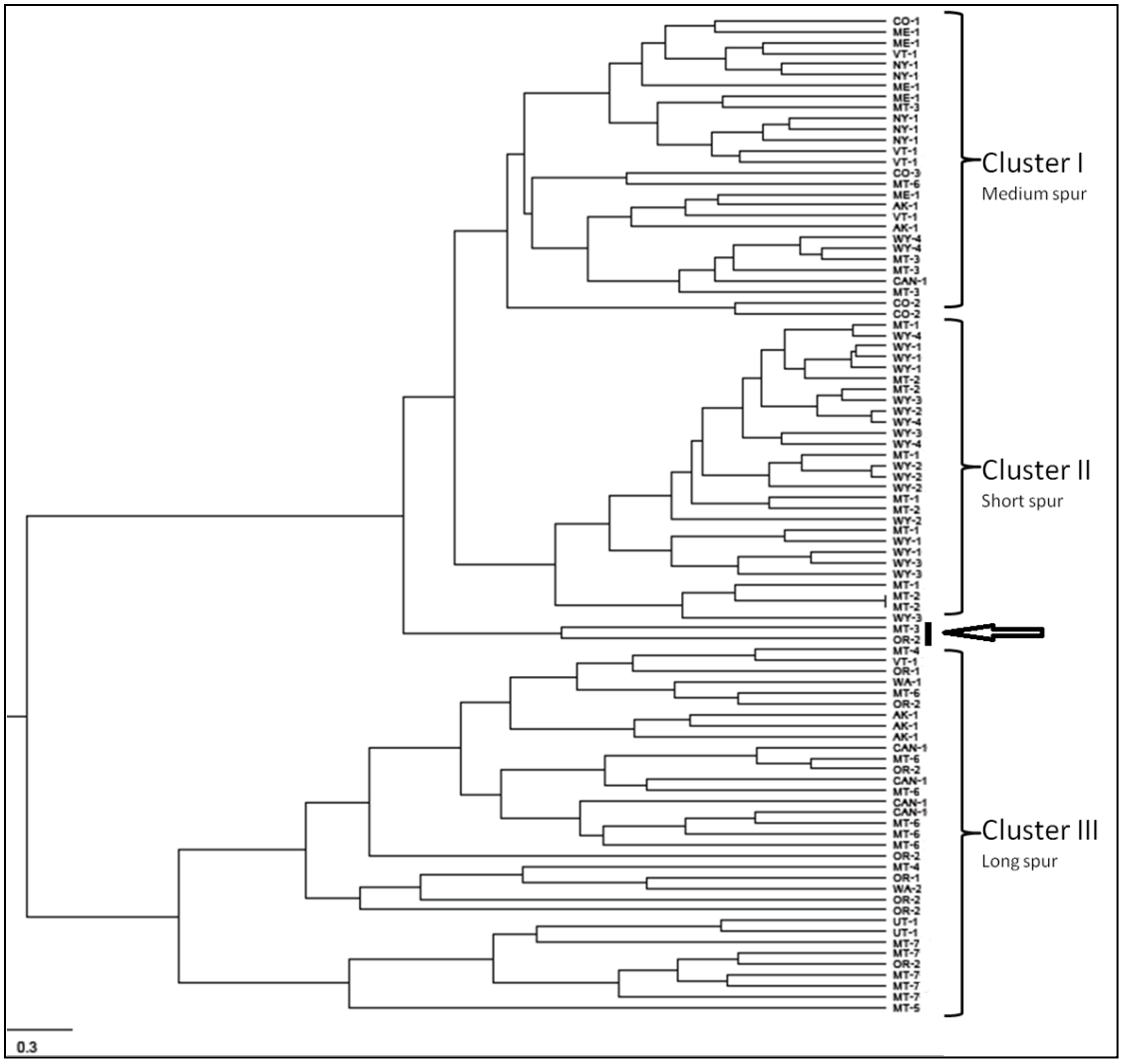


Figure 2.2 UPGMA dendrogram of the individuals of *Platanthera dilatata* based on Euclidean morphological distances.

The Euclidean distance was based on nine floral morphological characters among individuals (N = 92). The individuals are labeled by respective population codes. The three clusters referred to in the text are also labeled. One individual each of MT-3 and OR-2 (pointed by arrow) were not assigned to either cluster for morphological analyses.

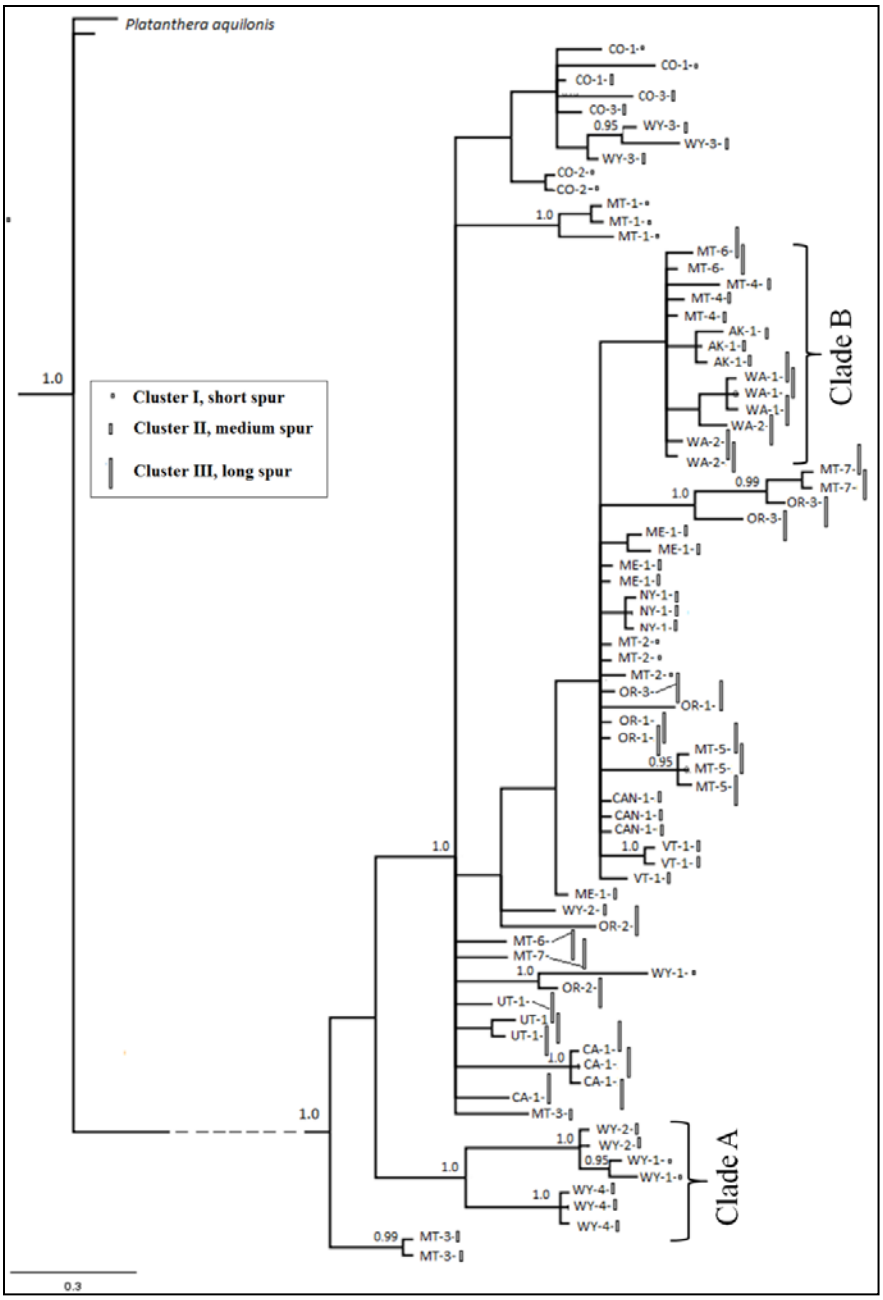


Figure 2.3 Bayesian strict consensus phylogram of the haplotypes of *Platanthera dilatata.*

Concatenated cpDNA sequences and gaps (indels) were used in the analysis (N = 78). Haplotypes are labeled by population codes followed by short, medium or long vertical bars to represent clusters. The clades described in the text are labeled. Posterior probability ≥ 0.70 are displayed on the branches. The branch connecting the outgroup was broken to accommodate space and is represented by a dashed line.

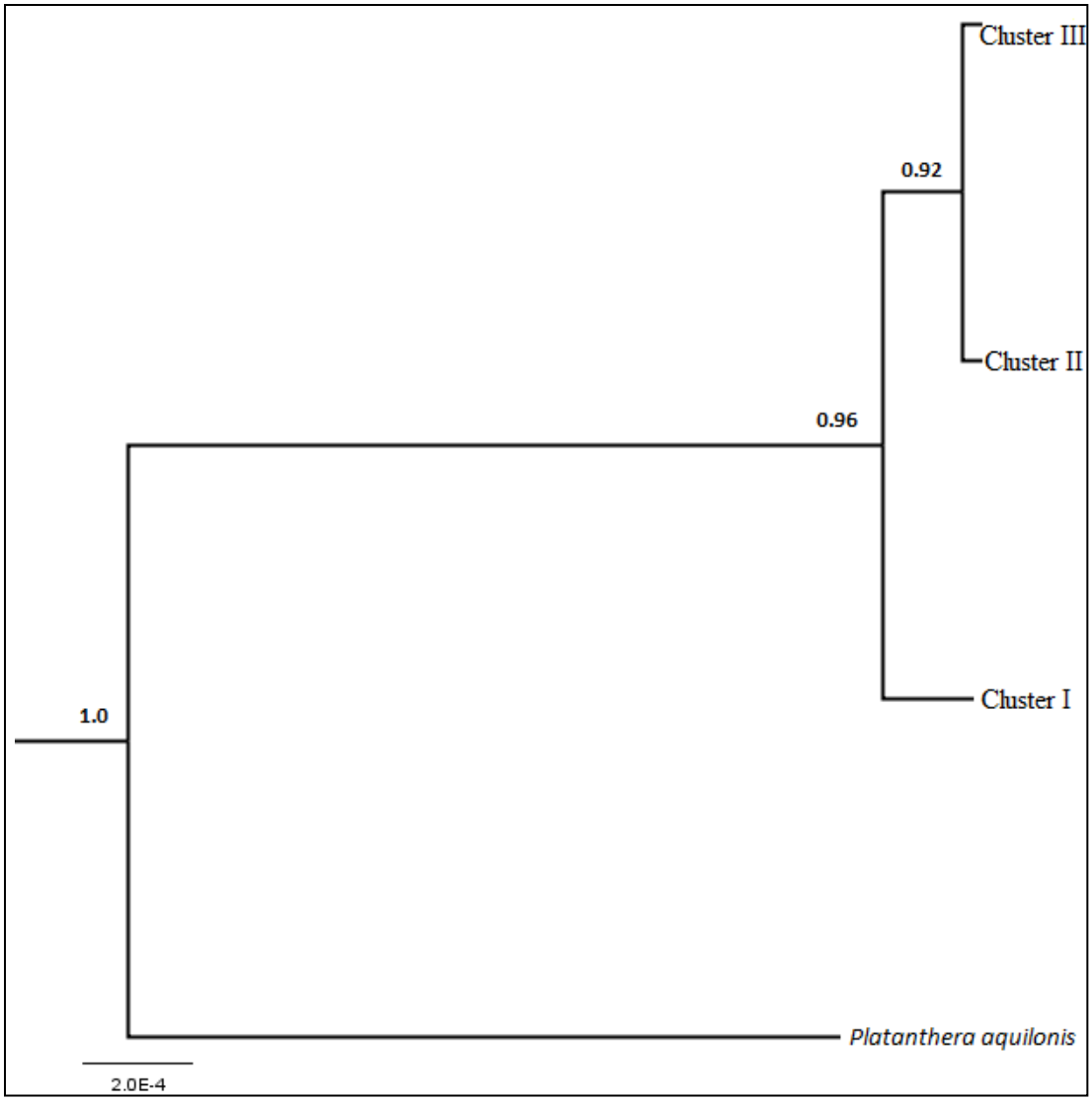
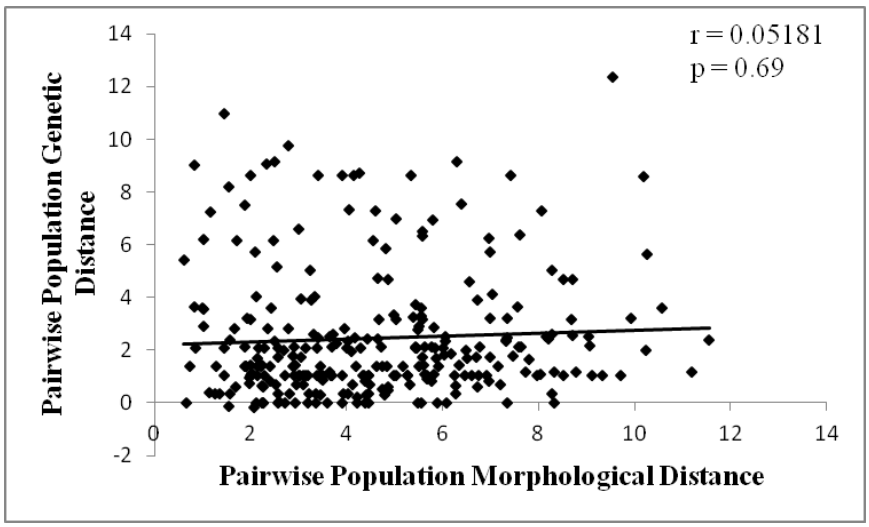
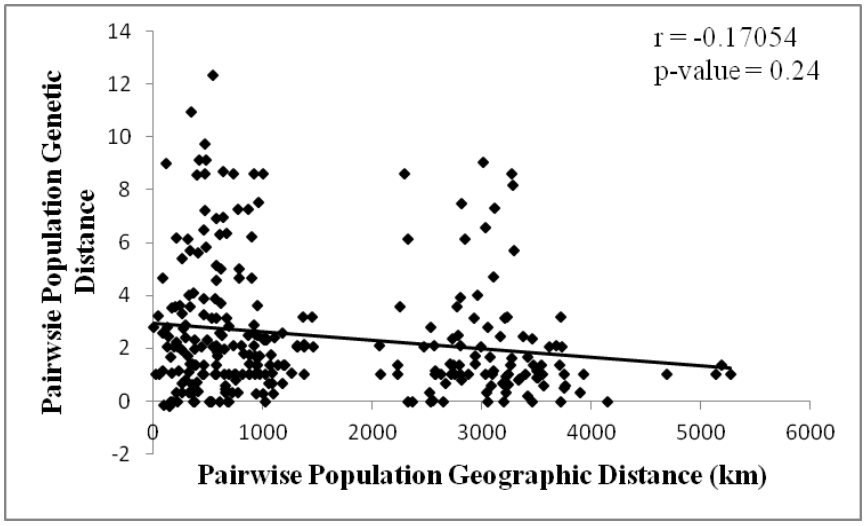


Figure 2.4 A species tree showing the relationship among three clusters of *Platanthera dilatata.*

These relationships are based on combined sequences of six cpDNA regions of *P. dilatata* (N = 78). Posterior probability support values are displayed on the branches. *Platanthera aquilonis* served as an outgroup species.



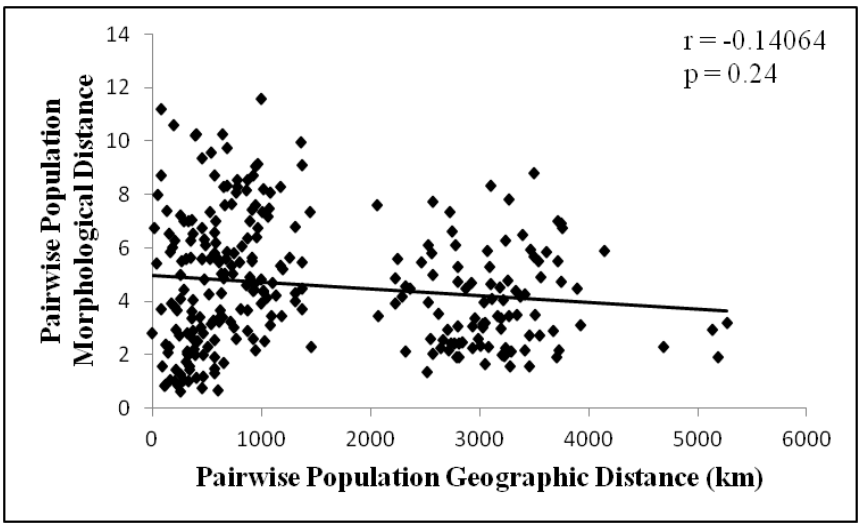
a



b

Figure 2.5 Scatter plots of correlations of morphological, genetic and geographic distances among populations of *Platanthera dilatata.*

These results are based on Mantel tests of population genetic distance with: a) morphological and b) geographic distances; and c) between morphological and geographic distances. The significance of the correlations were tested under 10000 permutations (p = 0.05).



c

Figure 2.5 (continued)

References

Ames, O. 1910. Orchidaceae: Illustrations and studies of the family Orchidaceae, Vol. IV: The Genus *Habenaria* in North America. Merrymount Press, Boston, MA.

Barrett, C. F., and J. V. Freudenstein. 2011. An integrative approach to delimiting species in a rare but widespread mycoheterotrophic orchid. *Molecular Ecology* 20:2771-2786.

Bikandi, J. 2010. BioPHP. Microsatellite Repeats Finder [Online]. Website:http:/www.biophp.org/minitools/microsatellite_repeats_finder/demo.php (Accessed on 20 August 2012).

Blaxter, M. L. 2004. The promise of a DNA taxonomy. *Philosophical transactions of the Royal Society of London B* 359:669-679.

Boberg, E., and J. Agren. 2009. Despite their apparent integration, spur length but not perianth size affects reproductive success in the moth-pollinated orchid *Platanthera bifolia.* *Functional Ecology* 23:1022-1028.

Boland, J. T. 1993. The floral biology of *Platanthera dilatata* (Pursh.) Lindl. (Orchidaceae). M.S. Thesis, Memorial University of Newfoundland, St. Johns, Newfoundland.

Cardoso, A., A. Serrano, and A. P. Vogler. 2009. Morphological and molecular variation in tiger beetles of the *Cicindela hybrida* complex: is an integrative taxonomy possible? *Molecular Ecology* 18:648-664.

Carstens, B. C., and L. L. Knowles. 2010. Navigating the unknown: model selection in phylogeography. *Molecular Ecology* 19:4581-4582.

Catling, P. M., and V. R. Catling. 1997. Morphological discrimination of *Platanthera huronensis* in the Canadian Rocky Mountains. *Lindleyana* 12:72-78.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Inc., Sunderland, MA.

Cronquist, A. 1978. Once again, what is a species? *In* L. Kuntson [ed.], Biosystematics in agriculture, 3-20. Alleheld Osmun, Montclair, NJ.

Cummings, M. P., M. C. Neel, K. L. Shaw, and S. Otto. 2008. A genealogical approach to quantifying lineage divergence. *Evolution* 62:2411-2422.

Cummings, M., M. Neel, K. Shaw, and A. Bazinet. 2010. Genealogical Sorting Index (Web-based software program). http://www.genealogicalsorting.org/ (accessed 3.05.2012).

Darwin, C. R. 1859. On the origin of species by means of natural selection, or the preservation of favoured reces in the struggle for life. London: John Murray.

Darwin, C. R. 1862. On the Various Contrivances by which British and Foreign Orchids are Fertilised by Insects. John Murray, London, England.

Darwin C R 1877. The different forms of flowers on plants of the same species. John Murray, London, England.

Dayrat B., 2005. Towards integrative taxonomy. *Biological Journal of the Linnean Society* 85:407-415.

de Queiroz, K. 1998. The general lineage concept of species, species criteria, and the process of speciation: A conceptual unification and terminological recommendations. *In* D. J. Howard, and S. H. Berlocher [eds], Endless Forms: Species and Speciation, 57-75. Oxford University Press, Oxford, England.

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

Dodd, M., E. J. Silvertown, and M. W. Chase. 1999. Phylogenetic analysis of trait evolution and species diversity variation among angiosperm families. *Evolution* 53:732-744.

Donoghue, M. J. 1985. A critique of the biological species concept and recommendations for a phylogenetic alternative. *Bryologist*. 88:172-181.

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

Drummond A. J., M. A. Suchard, D. Xie, and A. Rambaut 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*. doi:10.1093/molbev/mss075.

Duminil, J.,and M. Di Michele. 2009. Plant species delimitation: a comparison of morphological and molecular markers. *Plant Biosystem* 143:528-542.

Ersts, P. J. 2012. Geographic Distance Matrix Generator (version 1.2.3). American Museum of Natural History, Center for Biodiversity and Conservation. Available from http://biodiversityinformatics.amnh.org/open_source/gdmg. Accessed on 2012-5-20.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotype: applications to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.

Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10:564-567.

Fenster, C. B., W. S. Armbruster, P. Wilson, M. R. Dudash, and J. D. Thomson. 2004. Pollination syndromes and floral specialization. *Annual Review of Ecology, Evolution, and Systematics* 35:375-403.

García, P. E., P. Schönswetter, J. F. Aguilar, G. N. Feliner, and G. M. Schneeweiss. 2009. Five molecular markers reveal extensive morphological homoplasy and reticulate evolution in the *Malva* alliance (Malvaceae). *Molecular Phylogenetics and Evolution* 50:226-239.

Gaston, K. J., and M. A. O'Neil. 2004. Automated species identification: why not? *Philosophical Transactions: Biological Sciences* 359(1444):655-667.

Givnish, T. J., and K. J. Sytsma. 1997. Homoplasy in molecular and morphological data: The likelihood of correct phylogenetic inference. *In* T. J. Givnish, and K. J. Sytsma [eds], Molecular evolution and adaptive radiation, 55–101. Cambridge University Press, New York.

Good, D. A. and D. B. Wake. 1992. Geographic variation and speciation in the torrent salamanders of the genus *Rhyacotriton* (Caudata: Rhyacotritonidae). *University of California Publications in Zoology* 126:1-91.

Grant, V. 1963. The origin of adaptations. Columbia University Press, New York and London.

Haig, S. M., E. A. Beever, S. M. Chambers, H. M. Draheim, B. D. Dugger, and et al. 2006. Taxonomic considerations in listing subspecies under the US Endangered Species Act. *Conservation Biology* 20:1584-1594.

Hansen, K.T., R. Elven, and C. Brochmann. 2000. Molecules and morphology in concert: tests of some hypotheses in arctic *Potentilla* (Rosaceae). *American Journal of Botany* 87:1466-1479.

Hapeman, J. R. 1997. Pollination and floral biology of *Platanthera permnoena. Lindleyana* 12:19-25.

Hapeman, J. R., and K. Inoue. 1997. Plant-pollinator interactions and floral radiation in *Platanthera* (Orchidaceae). *In* T. J. Givnish and K. J. Sytsma [eds.], Molecular Evolution and Adaptive Radiation, 433-454. Cambridge University Press, Cambridge, England.

Harastova-Sobotkova, M., J. Jersakova, P. Kindlmann, and L. Curn. 2005. Morphological and genetic divergence among populations of *Neotinea ustulata* (Orchidaceae) with differing flowering phonologies. *Folia Geobotanica* 40:385-405.

Hebert, P. D. N., A. Cywinska, S. L. Ball, and J. R. deWaard. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences* 270:313–321.

Hedges, S. B., and L. R. Maxson. 1997. Complementary uses of molecules and morphology. *Molecular Phylogenetics and Evolution* 8:445.

Helbig, A. J., A. G. Knox, D. T. Parkin, G. Sangster, and M. Collinson. 2002. Guidelines for assigning species rank. *Ibis* 144: 518-525.

Heled, J., and A. J. Drummond. 2010. Bayesian Inference of Species Trees from Multilocus Data. *Molecular Biology and Evolution* 27(3):570-580.

Hennig, W. 1966. Phylogenetic systematics. Urbana: University of Illinois Press.

Hey, J., R. S. Waples, M. L. Arnold, R. K. Butlin, and R. G. Harrison. 2003. Understanding and confronting species uncertainty in biology and conservation. *Trends in Ecology and Evolution* 18:597-603.

Hodges, S. A., and M. L. Arnold. 1995. Spurring plant diversification: are floral nectar spurs a key innovation? *Proceedings of the Royal Society of London, Series B* 262:343-348.

Hurvich, C. M., and C.-L. Tsai. 1989. Regression and time series model selection in small samples. *Biometrika* 76:297-307.

Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17:754-755.

Ingvarsson, P. K., S. Ribstein, and D. R. Taylor. 2003. Molecular evolution of insertions and deletion in the chloroplast genome of Silene. *Molecular Biology and Evolution* 20:1737-1740.

Inoue, K. 1983. Systematics of the genus *Platanthera* (Orchidaceae) in Japan and adjacent regions with special reference to pollination. *Journal of the Faculty of Science of the University of Tokyo, Section III, Botany* 13:285-374.

Inoue, K. 1986. Experimental studies on male and female reproductive success: effects of variation in sur length and pollinator activity on *Platanthera mandarinorum* spp. *Hachijoensis* (Orchidaceae). *Plant Species Biology* 1:207-215.

Isaac, N. J. B., J. Mallet, and G. M. Mace. 2004. Taxonomic inflation: its influence on macroecology and conservation. *Trends in Ecology and Evolution* 19:464-469.

Jordan, E. C., M. W. Courtney, and J. T. Neigel. 1996. Low levels of intraspecific genetic variation at a rapidly evolving chloroplast DNA locus in North American duckweeds (Lemnaceae). *American Journal of Botany* 83: 430-439.

Kelchner, S. A. 2000. The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of the Missouri Botanical Garden* 87:482-498.

Kingman, J. F. C. 1982. The coalescent. *Stochastic Process and their Application* 13:235-248.

Kipping, J. L. 1971. Pollination studies of native orchids. M.S. Thesis, San Francisco, CA.

Kjár, A., A. S. Barfod, C. B. Asmussen, and O. Seberg. 2004. Investigation of Genetic and Morphological Variation in the Sago Palm (*Metroxylon sagu*; Arecaceae) in Papua New Guinea. *Annals of Botany* 94:109-117.

Kölreuter, J. G. 1761. Vorläufige Nachrichten von einigen das Geschlect der Pflanzen betreffenden Versuchen und Beobachtungen. Gleditschischen Handlung, Leipzig.

Les, D. H., M. L. Moody, S. W. L. Jacobs, and R. J. Bayer. 2002. Systematics of seagrasses (Zosteraceae) in Australia and New Zealand. *Systematic Botany* 27:468-484.

Librado, P., and J. Rozas. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.

Liu, L., and D. K. Pearl. 2007. Species tree from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Systematic Biology* 56: 504-514.

Liu, L., D. K. Pearl, R. T. Brumfield, and S. V. Edwards. 2008. Estimating species tree using multiple-allele DNA sequence data. *Evolution* 62: 2080-2091.

Luer, C. A. 1975. The Native Orchids of the United States and Canada, Excluding Florida. New York Botanical Garden, New York, NY.

Maad, J. 2000. Phenotypic selection in hawkmoth-pollinated *Platanthera bifolia*: targets and fitness surfaces. *Evolution* 54:112-123.

Maad, J., and R. Alexandersson. 2004. Variable selection in *Platanthera bifolia* (Orchidaceae): phenotypic selection differed between sex functions in a drought year. *Journal of Evolutionary Biology* 17:642-650.

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

Mallet, J. A. 1995. Species definition for the modern synthesis. *Trends in Ecology and Evolution*. 10:294-299.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27:209-220.Mayden, R. L. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem. *In* M. F. Claridge, H. A. Dawah, and M. R. Wilson [eds], Species: the units of biodiversity, 381-324. Chapman & Hall, London.

Mayden, R. L. 1997. A hierarchy of species concepts: the document of the species problem. *In* Claridge, M. F., H. A. Dawah, and M. R. Wilson [eds], The units of biodiversity - species in practice, Special vol. 54:381-424. London: Systematics Assocication, Chapman and Hall Ltd.

Mayfield, J. A., G. E. Reiber, C. Maynard, J. M. Czerniecki, M. T. Caps, and B. J. Sangeorzan. 2001. Survival following lower-limb amputation in a veteran population. *Journal of Rehabilitation Research and Development* 38(3):341-345.

Mayr, E. 2000. The biological species concept. *In* Q. D. Wheeler, R. Meier [eds], Species concepts and phylogenetic theory, 17-29. Columbia University Press, New York.

McDade, L. A. 1995. Species concepts and problems in practice: insights from botanical monographs. *Systematic Botany* 20:606-622.

Moritz, C. 1994. Applications of mitochondrial DNA analysis in conservation: a critical review. *Molecular Ecology* 3:401-411.

Muller, K. 2005. SeqState – primer design and sequence statistics for phylogenetic DNA data sets. *Applied Bioinformatics* 4:65-69.

Naomi, S. I. 2011. On the integrated frameworks of species concepts: Mayden's hierarchy of species concepts and de Queiroz's unified concept of species. *Journal of Zoological Systematics and Evolutionary Research* 49:177-184.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Nei, M., and F. Tajima. 1987. Problems arising in phylogenetic inference from restriction-site data. *Molecular Biology and Evolution* 4:320-323.

Nilsson, L. A. 1978. Pollination ecology and adaptation in *Plafanfhem chforanfha* (Orchidaceae). *Botaniska Notiser* 131:35-51.

Nilsson, L. A. 1983. Processes of isolation and introgressive interplay between *Platanthera bifolia* (L.) Rich.and *P. chlorantha* (Custer) Reichb. (Orchidaceae). *Botanical Journal of the Linnaean Society* 87:325-350.

Nilsson, L. A. 1988. The evolution of flowers with deep corolla tubes. *Nature* 334:147-149.

Nixon, K. C., and Q. D. Wheeler. 1990. An amplification of the phylogenetic species concept. *Cladistics* 6(3):211-223.

Noble, D. W. A., Y. Qi, and J. Fu. 2010. Species delineation using Bayesian model-based assignment tests: a case study using Chinese toad-headed agamas (genus *Phrynocephalus*). *BMC Evolutionary Biology* 10:197.

Olmstead, R. G., and J. D. Palmer. 1994. Chloroplast DNA systematics: A review of methods and data analysis. *American Journal of Botany* 81(9):1205-1224.

Olsen, K. M. 1997. Pollination effectiveness and pollinator importance in a population of *Heterotheca subaxillaris* (Asteraceae). *Oecologia* 109:114-21

Padial, J. A. Miralles, I. De la Riva, and M. Vences. 2010. The integrative future of taxonomy. *Frontiers in Zoology* 7:16.

Patt, J. M., M. W. Merchant, D. R. E Williams, and B. J. D. Meeuse. 1989. Pollination biology of *Platanthera stricta* (Orchidaceae) in Olympic National Park, Washington. *American Journal of Botany* 76:1097-1106.

Pettengill, J. B., and M. C. Neel. 2011. A sequential approach using genetic and morphological analyses to test species status: the case of United States federally endangered *Agalinis acuta* (Orobanchaceae). *American Journal of Botany* 98(5):859-871.

Posada, D. 2008. jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution* 25: 1253-1256.

Puorto, G., S. M. Da Graca, R. D. G. Theakston, R. S. Thorpe, D. A. Warrell, and et al. 2001. Combining mitochondrial DNA sequences and morphological data to infer boundaries: phylogeography of lancehead pitvipers in the Brazilian Atlantic forest, and the status of *Bothrops pradoi* (Squamata: Serpentens: Viperidae). *Journal of Evolutionary Biology* 14:527-538.

Rambaut, A. 2010. Se-AL: Sequence Alignment Editor. Available online at http://tree.bio.ed.ac.uk/software/seal/.

Rambaut A. 2012. FigTree v1.1.1: Tree figure drawing tool. Available: http://tree.bio.ed.ac.uk/software/figtree/. Accessed 10 June 2012.

Rambaut, A., and A. J. Drummond. 2007. LogCombiner v1.4, Available from http://beast.bio.ed.ac.uk/LOgCombiner.

Rambaut, A., and A. J. Drummond. 2007. Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer.

Rambaut, A., and A. J. Drummond. 2007. TreeAnnotator v1.4, Available from http://beast.bio.ed.ac.uk/TreeAnnotator.

Reeves, P. A., and C. M. Richards 2011. Species delimitation under the general lineage concept: an empirical example using wild North American Hops (Cannabaceae: *Humulus lupulus*). *Systematic Biology* 60(1):45-59.

Rieseberg, L. H., S. A. Church, and C. L. Morjan. 2003. Integration of populations and differentiation of species. *New Phytologist* 161:59-69.

Robertson, J. L., and R. Wyatt. 1990. Evidence for pollination ecotypes in the yellow fringed orchid, *Platanthera ciliaris*. *Evolution* 44:121-133.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-1574.

Rosenberg, M. S., and C. D. Anderson. 2011. PASSaGE: Pattern Analysis, Spatial Statistics, and Geographic Exegesis. Version 2. *Methods in Ecology and Evolution* 2(3):229-232.

Rozen, S., and H. J. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. *In* S. Krawetz, and S. Misener [eds], Bioinformatics Methods and Protocols: Methods in Molecular Biology, 365-386. Humana Press, Totowa, NJ.

Rundle, H. D., and P. Nosil. 2005. Ecological speciation. *Ecology Letters* 8:336-352.

Rydberg, P. A. 1901. The American species of *Limnorchis* and *Piperia*, north of Mexico. *Bulletin of the Torrey Botanical Club* 28:605-643.

Ryder, O. A. 1986. Conservation and systematics: The delimitation of sub-species. *Trends in Ecology and Evolution*. 1:9-10.

Sattler, R., and R. Rutishauser. 1997. The fundamental relevance of morphology and morphogenesis to plant research. *Annals of Botany* 80:571-582.

Schemske, D.W., and C. C. Horvitz. 1984. Variation among floral visitors in pollination ability a precondition for mutualism specialization. *Science* 225:519-521.

Schiestl, F. P., and P. M. Schluter. 2009. Floral isolation, specialized pollination, and pollinator behavior in orchids. *Annual Review of Entomology* 54:425-46.

Schrenk, W. J. 1978. North American platantheras: evolution in the making. *American Orchid Society Bulletin* 47: 429-437.

Shaffer, H. B., and R. C. Thomson. 2007. Delimiting species in recent radiations. *Systematic Biology* 56(6):896-906.

Sheviak, C. J. 2002. *Platanthera*. *In* Flora of North America Editorial Committee, (eds.). Flora of North America North of Mexico Vol. 26, 551-571. Oxford University Press, Oxford, England.

Sheviak, C. J., and M. Bracht. 1998. New Chromosome Number Determinations in *Platanthera*. *North American Native Orchid Journal* 4:168-172.

Shaw, J., E. B. Lickey, E. E. Schilling, and R. L. Small. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany* 94(3): 275-288.

Simmons, M. P., and H. Ochoterena. 2000. Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* 49:369-381.

Simmons, M. P., H. Ochoterena, and T. G. Carr. 2001. Incorporation, relative homoplasy, and effect of gap characters in sequence-based phylogenetic analyses. *Systematic Biology* 50:454-462.

Sites, J. W. Jr., and J. C. Marshall. 2004. Operational criteria for delimiting species. *Annual Review of Ecology, Evolutionary and Systematics* 35:199-227.

Sokal, R. R., and P. H. A. Sneath. 1963. Principles of Numerical Taxonomy. San Francisco: W. H. Freeman.

Sneath, P. H. A., and R. R. Sokal. 1973. Numerical Taxonomy: The Principles and Practice of Numerical Classifications, First edition. San Franscisco, W H Freeman & Co (Sd).

Sokal, R. R., and J. Crovello 1970. The biological species concept: a critical evaluation. *American Naturalist* 104:127-153.

Stebbins Jr., G. L. 1970. Adaptive radiation of reproductive characteristics in Angiosperms, I: Pollination mechanisms. *Annual Review of Ecology and Systematics* 1:307-326.

Stuessy, T. F. 2009. Plant taxonomy: The systematic evaluation of comparative data, 2nd ed. Columbia University Press, New York.

Swofford, D. L. 2002. PAUP* Phylogenetic Analysis Using Parsimony and Other Methods, Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tamura, K., and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology Evolution* 10(3):512-26.

Thorpe, R. S., H. Black, and A. Malhotra. 1996. Matrix correspondence tests on the DNA phylogeny of the Tenerife Lacertid elucidate both historical causes and morphological adaptation. *Systematic Biology* 45:335-343.

Valdecasas, A. G., D. Williams, and Q. D. Wheeler. 2005. Integrative taxonomy then and now: a response to Dayrat (2005). *Biological Journal of the Linnaean Society* 93:211-216.

van Oppen, M. J., C. Rico, G. F. Turner, and G. M. Hewitt. 2000. Extensive homoplasy, nonstepwise mutations, and shared ancestral polymorphism at a complex microsatellite locus in Lake Malawi cichlids. *Molecular Biology and Evolution* 17:489-498.

Verbruggen, H., O. De Clerck, T. Schils, W. H. C. F. Kooistra, and E. Coppejans. 2005. Evolution and phylogeography of *Halimeda* section *Halimeda* (Bryopsidales, Chlorophyta). *Molecular Phylogenetics and Evolution* 37:789-803.

Wallace, L. E. 2003. An evaluation of taxonomic boundaries in *Platanthera dilatata* (Orchidaceae) based on morphological and molecular variation. *Rhodora* 105:322-336.

Waples, R. S. 1991. Pacific salmon, *Oncorynchus* spp., and the definition of a "species" under the Endangered Speceis Act. *Marine Fisheries Review* 53:11-22.

Wiens, J. J. 2004. What is speciation and how should we study it? *American Naturalist* 163:914-923.

Wiens, J. J., and T. L. Penkrot. 2002. Delimiting species based on DNA and morphological variation and discordant species limits in spiny lizards (*Sceloporus*). *Systematic Biology* 51:69-91.

Wiens, J. J., P. T. Chippindale, and D. M. Hillis. 2003. When are phylogenetic analyses misled by convergence? A case study in Texas cave salamanders. *Systematic Biology* 52:501-514.

Wiley, E. O. 1981. Phylogenetics: The theory and practice of phylogenetic systematics. New York: Wiley.

Wolfe, K. H., W. H. Li, and P. M. Sharp 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences USA* 84:9054-9058.

Wright, S. 1965. The Interpretation of Population Structure by F-Statistics with Special Regard to Systems of Mating. *Evolution* 19(3):395-420.

Xu, S., P. M. Schluter, G. Scopece, H. Breitkopf, K. Gross, S. Cozzolino, and F. P. Schiestl. 2011. Floral isolation is the main reproductive barrier among closely related sexually deceptive orchids. *Evolution* 65(9): 2606-2620.

Yue, J-X, J. Li, D. Wang, H. Araki, D. Tian, and S. Yang. 2010. Genome-wide investigation reveals high evolutionary rates in annual model plants. *BMC Plant Biology* 10:242.

Zar, J. H. 1996. Biostatistical Analysis. Prentice Hall, Upper Saddle River, NJ.

APPENDIX A

INDIVIDUAL cpDNA REGION DIVERSITY METRICS FOR *PLATANTHERA DILATATA*

cpDNA regions	Groups	# haplotypes	Haplotype diversity (SD)	Nucleotide diversity (SD)
rpL16 (648 bp)	Cluster I	4	0.6797 (0.0795)	0.000151 (0.00021)
	Cluster II	5	0.6619 (0.0825)	0.00059 (0.00024)
	Cluster III	9	0.7814 (0.0499)	0.00093 (0.00025)
	P. dilatata	**13**	**0.7749 (0.0351)**	**0.00112 (0.0002)**
trnV-ndhC (653 bp)	Cluster I	6	0.8366 (0.0533)	0.00162 (0.00034)
	Cluster II	7	0.8048 (0.0627)	0.0016 (0.00026)
	Cluster III	8	0.7688 (0.0370)	0
	P. dilatata	**16**	**0.8365 (0.0272)**	**0**
psaI-accD (704 bp)	Cluster I	9	0.9150 (0.0348)	0.00382 (0.00053)
	Cluster II	8	0.7238 (0.1005)	0.00108 (0.00047)
	Cluster III	16	0.9325 (0.0191)	0.00085 (0.00024)
	P. dilatata	**30**	**0.9337 (0.0179)**	**0.00188 (0.00038)**
psbC-trnS (59 bp)	Cluster I	5	0.8039 (0.0491)	0.0000073 (0.00271)
	Cluster II	9	0.8632 (0.0486)	0
	Cluster III	9	0.6444 (0.0800)	0
	P. dilatata	**13**	**0.8428 (0.0276)**	**0**
psbA-trnK (279 bp)	Cluster I	7	0.8431 (0.0521)	0.00081 (0.00045)
	Cluster II	7	0.7048 (0.0948)	0.00036 (0.00032)
	Cluster III	8	0.7681 (0.0547)	0
	P. dilatata	**12**	**0.8567 (0.0177)**	**0**
atpF-atpH (168 bp)	Cluster I	4	0.5948 (0.1086)	0
	Cluster II	6	0.6952 (0.0908)	0
	Cluster III	7	0.8205 (0.0300)	0.00031 (0.00029)
	P. dilatata	**11**	**0.7619 (0.0393)**	**0.00016 (0.00015)**

Chloroplast DNA diversity metrics for *Platanthera dilatata* and the three morphological groups at each of the cpDNA regions used in this study (SD = standard deviation). Sequence alignment gaps were included when estimating these matrices.