4-10-2006

A Determination of Phylogeny and Hybridization History Within Clematis L. (Ranunculaceae) Using Actin and Nitrate Reductase Intron Sequences

Kimberly Fearn Do
University of South Florida

A Determination of Phylogeny and Hybridization History Within *Clematis* L. (Ranunculaceae) Using Actin and Nitrate Reductase Intron Sequences

by

Kimberly Fearn Do

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science
Department of Biology
College of Arts and Sciences
University of South Florida

Co-Major Professor: Frederick B. Essig, Ph.D.
Co-Major Professor: James R. Garey, Ph.D.
Richard P. Wunderlin, Ph.D.

Date of Approval:
April 10, 2006

Keywords: plant, molecular, DNA, clade, allele,

Table of Contents

List of Tables

List of Figures

A Determination of Phylogeny and Hybridization History Within *Clematis* L. Section *Viorna* (Ranunculaceae) Using Actin and Nitrate Reductase Intron Sequences

Kimberly Fearn Do

ABSTRACT

The phylogeny of *Clematis*, section *Viorna*, was characterized in this study using molecular data. Two nuclear introns were sequenced for a variety of taxa: actin and nitrate reductase. Actin intron sequence data yielded very little phylogenetic information. Some basal clades were resolved, but there were very few well supported relationships between species of the *Viorna* section in both the neighbor joining and maximum parsimony analyses. Nitrate reductase intron sequence data was slightly more variable. The number of well supported relationships in both the neighbor joining and maximum parsimony analyses for nitrate reductase was greater, but still not sufficient to yield an informative tree. Two possible explanations for the lack of variation are that these species have not evolved many differences in these intron sequences or that common alleles are flowing between the species. Hybrid analysis using the actin intron was inconclusive because the experimentally generated hybrid possessed an allele that neither parents

tested had. More sampling from multiple individuals from both parent and multiple hybrid individuals is necessary to answer this question. The hybrid specimen tested was homozygous for the nitrate reductase intron marker, and both parents also possessed the allele. This did not directly support or refute the use of these markers for tracking the hybridization history within *Clematis*.

Chapter One

INTRODUCTION

The Genus *Clematis*

Clematis is a major member of the family Ranunculaceae, known for its basal placement within the group of eudicots in the tree of life (Zomlefer, 1995). The Ranunculaceae still retain some of the characteristics of the ancestral eudicots such as an apocarpous gynoecium. Although it is known that the Ranunculaceae form a basal relationship within the eudicots, phylogenetic relationships have not been clearly resolved within the family. This study attempts to clarify some of the relationships within the genus *Clematis*.

Clematis, distributed world wide, is comprised of about 320 species in 19 sections (Johnson, 2001). The section of interest in the current study is *Viorna* Prantl, which consists of six subsections and 23 species. Nine of the species in this section have been included in this study representing all but one of the subsections (Table 1).

Table 1. List of described species included in the study. The author of the species name and the section of the species as classified by Tamura (1968) and modified by Johnson (2001) as well as voucher numbers included here. No vouchers were available for *C. ochroleuca*, *C. fremontii*, *C. fusca.* These species were purchased from nurseries and never flowered before dying. Identity of these can not be irrefutably verified. *C. socialis* samples were taken from two populations both of which are represented here. The identity of C. jackmanii, even with a voucher specimen, can also not be irrefutably verified.

Species Name	**Author**	**Tamura's Section**	**Voucher Number (USF)**
Clematis baldwinii	Torr. & A. Gray	*Viorna*	Essig 990406-1
Clematis glaucophylla	Small	*Viorna*	Essig 011001-6
Clematis reticulata	Walter	*Viorna*	Arias 71
Clematis crispa	L.	*Viorna*	Essig 981207-2
Clematis integrifolia	L.	*Viorna*	Essig 011001-2
Clematis socialis	Kral	*Viorna*	Gasden CS-58
Clematis socialis	Kral	*Viorna*	Laney CS-97
Clematis ochroleuca	Aiton	*Viorna*	
Clematis fremontii	S. Watson	*Viorna*	
Clematis fusca	Turcz.	*Viorna*	
Clematis campaniflora	Lodd. ex Steud	*Viticella*	Essig 040824-2
Clematis jackmanii x ?	T. Moore	*Viticella*	Essig 040824-1
Clematis lasiandra	Maxim.	*Connatae*	Essig 040824-3
Clematis drummondii	Torr. & A. Gray	*Clematis*	Gonzalez 7008
Clematis catesbyana	Pursh	*Clematis*	Hart 10-18-90
Clematis brachyura	Maxim.	*Pterocarpa*	Essig 020305-2
Clematis terniflora	DC.	*Flammula*	Essig 890904-1

Previous Classifications

There have been several proposed classification schemes for *Clematis*. The most widely accepted of these was proposed by Tamura (1968, 1989) and was based solely on morphological characters. Tamura's morphological classification was then revised in by Johnson (1997, 2001). The division of *Clematis* into sections and subsections tested in this study is based on this classification. Essig (1991) proposed a different classification based on seedling morphology. In this system, Essig proposed two divisions: Type I seedlings characterized by alternate phyllotaxy, absence of cataphylls, toothed eophyll margins, and an elongate hypocotyl, and Type II seedlings characterized by opposite phyllotaxy, presence of cataphylls, entire eophyll margins, and a compact hypocotyl. Essig's system is incongruent with that of Tamura.

In more recent years classification schemes have shifted from using morphological data to molecular data. Nucleotide sequence divergence has been a proven method for resolving phylogenetic questions resulting from morphological data analysis (Hoot and Palmer 1994; Cros et al. 1998; Soltis and Kuzoff 1995; Hardig et al. 2000). Both chloroplast and nuclear genomes have been utilized for these molecular analyses. These genomes are useful for different applications depending on the divergence level of the taxa in question. Chloroplast markers are useful for resolving deeper level (generic and familial) relationships. The chloroplast genome has a slower rate of nucleotide substitution than the nuclear genome of plants because of the nature of the genes carried on that genome (Hillis et al. 1996). For example, there are 37 tRNA coding genes carried on the chloroplast genome; most mutations to any of these genes would be deleterious to their function (Li 1997).

Miikeda et al. (1999) published a preliminary molecular classification of *Clematis* using a variety of coding and non-coding chloroplast genome markers. Although over 4000 base pairs were sequenced, the analysis only produced a poorly resolved molecular evidence tree. For species level phylogenetic questions, a nuclear marker is more appropriate. Specifically, a nuclear non-coding marker, such as an intron, would have a relatively fast substitution rate yielding sufficient variability to resolve some relationships (Li 1997). Recently, there have been many advances in the development of primers that amplify these highly variable nuclear introns. Strand et al. (1997) developed a methodological approach to finding informative nuclear characters. Using this procedure, Slomba et al. (2004) developed primers that amplify an intron of the nuclear gene actin.

Actin Intron Marker

The actin protein is an important element of the plant cytoskeleton. Actin is important in cellular expansion during growth, as well as movement of organelles and vesicles around the cell (Mathur and Hülskamp 2002; Mathur 2004). There is also some indication that actin is involved in the gravitropic response that plants display. Data suggests that statolith position is also affected by an actin myosin complex (Friedman et al. 2003).

Actin may be a useful marker because preliminary analysis shows that the non-synonymous rate of substitution is up to 19 times higher in angiosperm actin genes than mammalian or fungal actin genes (Moniz de Sá and Drouin 1996). This increased rate of

substitution may yield sufficient phylogenetically informative sites to resolve relationships between species hypothesized to be classified in the same section.

Using taxa chosen from across the genus, Slomba (2004), was able to develop a working hypothesis of the phylogeny of *Clematis* using the actin intron. He found two distinct clades that correspond to the two seedling morphology types of Essig (Slomba et al. 2004). It is hoped that expanded taxon sampling that encompasses a representation of the *Viorna* section may improve what is known about these relationships.

Nitrate Reductase Intron Marker

Although one marker may result in a phylogenetic hypothesis, multiple markers yielding the same phylogeny further support the hypothesis. Therefore, an additional nuclear intron was utilized in this study. Howarth and Baum (2002) developed primers for the second and third introns of nitrate reductase. They were successful in amplifying this intron in multiple species of *Scaevola* (Goodeniaceae). For Howarth and Baum, sequences of this intron were particularly useful because, in multiple genera studied, the size of the intron was large enough to yield much variability and many phylogenetically informative characters (from 85 to 1600 bp).

Nitrate reductase is a low copy number, nuclear encoded gene essential to nitrogen metabolism in plants. Nitrate reductase catalyzes the first step of nitrate assimilation where nitrate is reduced to nitrite (Sharma and Dubey 2005). Nitrate is the principal form of nitrogen available to plant roots in fertilizers and the action of nitrogen fixing bacteria. Nitrate must then be reduced into forms that can be used to make amino acids. Nitrate assimilation takes place in the leaves where light controls the activity of

nitrate reductase. This is because the reduction of nitrate to nitrite is closely coupled to both photosynthetic and respiratory reactions (Foyer et al. 2003).

Combined data from both actin and nitrate reductase intron sequences may yield more insights into the phylogeny of this subset of the genus *Clematis*. These intron sequences may be more informative than simply to resolve phylogenetic questions; it may be also possible to track the history of hybrid species using these markers. *Clematis* is a common garden ornamental, which gives this genus scientific and economic importance. One of the reasons that *Clematis* is such a popular garden plant is the relative ease of hybridization within the genus, especially between closely related species. Hybridization allows plant enthusiasts to manipulate morphological characters to their liking. For example, there are 14 different cultivars of *Clematis* that originated from hybridization between *C. texensis* and other large flowered species (Johnson, 2001). This ease of artificial hybridization leads to the hypothesis that many naturally occurring species may be the result of natural hybridization events.

Characterizing Hybrid History

According to Rieseberg (1997), hybrids are organisms formed by cross fertilization between two individuals of two populations distinguishable from one another by one or more heritable characters. Often, populations that interbreed to form a hybrid originate from two different species. The evolutionary influence of hybridization to speciation has likely been underestimated because of difficulties in detecting hybrid origins (Rossetto 2005). Difficulties arise in detecting hybrids, as well as in including hybrid species in phylogenetic analyses, because these species result from a reticulation

event rather than a bifurcation (Koontz et al. 2004). Before molecular techniques were available, intermediacy in morphological characters was the primary method for detecting a hybrid. Morphological evidence of hybrid origins can be very compelling, but it can also be misleading. Intraspecific variation may falsely indicate the presence of a hybrid or it may mask the existence of a true hybrid population or species. Recently, molecular tools have been utilized for determining the origins of putative hybrid species (Aguilar et al. 1999; Ferguson and Sang 2001; Franzke and Mummenhoff 1999; Koontz et al. 2004; Rossetto 2005; Zhou et al 2005).

Popular molecular markers for botanical studies have originated on the chloroplast genome. The paucity of variable sites on this genome makes it inadequate for differentiating haplotypes at very fine taxonomic levels like the species and population level (Rossetto 2005). Chloroplast markers are also not sufficient for hybrid studies because of the maternal inheritance of this genome. Koontz et al. (2004) used the cpDNA marker trnl-F to try to resolve the hybrid origins of *Delphinium gypsophilum* Ewan (Ranunculaceae) in conjunction with an ITS rDNA nuclear marker. These two markers were used because they represented genes with different modes of inheritance. The chloroplast marker is maternally inherited, while the nuclear marker is biparentally inherited. The hope of the study was that the chloroplast marker would identify the maternal parent of the putative hybrid and group in a clade with that maternal parent. However, even using the chloroplast marker in association with a nuclear marker could not completely resolve the hybrid relationships due to lack of variation. Since intron regions have a faster substitution rate than the chloroplast genome, using multiple nuclear

markers may make it possible to differentiate haplotypes of even recently diverged species (Li 1997).

Objectives of Current Study

The current study strives to resolve the phylogeny of the section *Viorna* of the genus *Clematis* using molecular intron markers. Another objective, also using molecular data, is to characterize the hybrid history of species within *Viorna* focusing on an experimentally generated hybrid.

Chapter Two

MATERIALS & METHODS

Specimen Collection and Preservation

Specimens were obtained from multiple sources. Fresh leaf material was collected from the yard of Dr. Frederick Essig (cultivated specimens, natural population unknown), the University of South Florida Botanical Gardens (cultivated specimens, natural population unknown), and Lettuce Lake Park in Tampa (natural population). Leaf material was examined under a dissecting scope to remove parasites and other foreign organic matter. Leaf material was then homogenized using liquid nitrogen and a mortar and pestle. Ground samples were stored in microcentrifuge tubes at -80°C.

Genomic DNA Extraction

Genomic DNA was extracted from ground leaf samples using the Plant DNA Isolation Kit (13000-50) from MO BIO (Solana Beach, CA). The kit bound DNA to a filter spin column to purify it from other plant cell contents. DNA was then eluted from the membrane in the spin column with an elution buffer and stored at -20ºC.

Polymerase Chain Reaction (PCR)

Polymerase chain reaction (PCR) was used to amplify desired genetic markers for all extracted DNA samples using a Biometra T3 thermal cycler (#1510303). Primers used for the various gene regions were manufactured by IDT (Integrated DNA Technologies Coalville, IA). Actin intron forward and reverse primers used were published by Slomba (2004). Nitrate reductase third intron forward and reverse primers used were published by Howarth and Baum (2002). PCR amplifications were set up in 200µl thin walled, hinge capped microcentrifuge tubes to a final volume of 50µl and consisted of 38.75µl molecular biology grade water, 5µl 10X Reaction Buffer (ID Lab Inc., London, ON, Canada), 1µl dXTP mix (1:1:1:1 of 12.5 mM dATP, dGTP, dCTP, and dTTP), 0.25µl high fidelity ID Proof DNA Polymerase (ID Lab Inc., London, ON, Canada), 1µl forward and reverse primer at 10µM, 2µl of 25mM $MgCl_2$, and 1µl template DNA (0.25µg/µl). Thermal cycling parameters for the actin intron were a 2 min. initial denature at 95°C followed by 35 cycles of 15 sec. at 95°C, then a 48°C annealing temp. for 30 sec. with a 72°C extension for 1 min. 30 sec. followed by a 10 min. final extension at 72°C. The nitrate dehydrogenase third intron used touch down thermal cycling parameters.

PCR Purification

All PCR prducts were purified to remove excess primers (primer-dimers) and unincorporated dNTP's before cloning. Purification was carried out using the QIAquick Gel Extraction Kit from Qiagen (Valencia, CA). The full volume of sample was electrophoresed onto a 0.9% agarose gel at 115V. DNA bands representing the desired

amplicon were then visualized with a longwave UV light source and excised from the gel into microcentrifuge tubes. The Qiagen kit dissolved the agarose gel and separated the DNA using a spin filter method. DNA was then eluted off the membrane with 30µl of molecular biology grade water (MBW).

Cloning and Plasmid Extraction

In order to test for the presence of multiple alleles of these introns, PCR product was cloned into One Shot Mach1-T1^{R} chemically competent *Escherichia coli* cells (Invitrogen; Carlsbad,CA) using a Topoisomerase TA Cloning Kit (Invitrogen; Carlsbad, CA). After 18 hours of incubation at 37ºC, colonies that grew on LB agar with ampicillin (50µg/ml) were those that contained an insert in the plasmid vector. These colonies were then picked off the agar plates and grown in LB broth on a microtitre plate and incubated overnight at 37°C (170 µl LB broth per well with ampicillin (50µg/ml)). One drop of glycerol was added to each well prior to being stored at -80°C. Once twelve specimens had viable colonies, a 96 well culture plate was inoculated with stored clones; each specimen represented by one row on the plate that was then incubated overnight at 37°C. Plasmid DNA containing the amplicon of interest was then extracted from the bacterial cells using the Perfectprep Plasmid 96 Vac, Direct Bind Kit (Eppendorf, Brinkmann Instruments, Inc; Westbury, NY). Final elution of DNA off of the filter plate was done with 50µl of molecular biology grade water rather than the specified elution buffer. It was also evident that ethanol was not removed sufficiently using the Perfectprep kit, therefore plates were also completely dried using a Savant SpeedVac Plus. DNA was then resuspending in 50µl of MBW.

Restriction Digestion of Plasmid DNA

To verify presence of proper insert in plasmid DNA preparations, restriction enzymes were used to cut the insert out of the plasmid. Reactions consisted of 14µl of MBW, 2µl plasmid DNA, 2µl EcoRI buffer (Promega; Madison, WI), 1µl EcoRI enzyme (Promega; Madison WI), and 1µl of BSA (bovine serum albumin)(Promega; Madison, WI). Reactions were incubated at 37°C for 2 hours; 10µl of each reaction was then eletrophoresed onto a 0.9% agarose gel run at 115V. The resulting image separated the false positive clones from those with an insert, and also removed those with an incorrectly sized insert. This also allowed for gel quantification of the amount of amplicon present in the sample.

Cycle Sequencing of Intron Markers

Sequencing was done on a Beckman Coulter CEQ8000 Genetic Analysis System using the GenomeLab Dye Terminator Cycle Sequencing with Quick Start Kit (Beckman Coulter; Fullerton, CA). Reactions were actually run at quarter strength of that recommended by Beckman with a final volume of 10µl. For most samples 4µl of template plasmid DNA had the appropriate concentration for sequencing. Template DNA was combined with 2µl of MBW and incubated at 96°C for 1 min. and then transferred to ice to uncoil the plasmid and increase primer binding efficiency. The preheated template DNA and MBW were combined with 2µl of plasmid specific primer at 3.2µM and 2µl of DTCS Quick Start Master Mix (Beckman Coulter; Fullerton, CA). Sequencing parameters for this reaction were 40 cycles of 96ºC for 20 sec. followed by 50ºC for 20 sec. and 60ºC for 4 min. The reaction could then be held indefinitely at 4ºC. Once the

thermal cycling program was finished, dye terminated fragments were then precipitated and purified. A solution was prepared fresh consisting of 2 parts 3M Sodium Acetate (pH 5.2), 2 parts 100mM Na_2-EDTA(pH 8.0), and 1 part 20 mg/ml glycogen; 5µl of this solution was added to each sequence in addition to 60µl of cold 100% ethanol. The mixture was then vortexed and centrifuged at 4°C for 15 min at 13,000 rpm. This should result in pelleted DNA; supernatant was removed and discarded. Pellet was rinsed twice with cold 70% ethanol and centrifuged for 2 min. at 4ºC. After each rinse supernatant was removed and discarded. Pellets were then dried for 4 minutes on low heat in the Savant SpeedVac Plus, and finally resuspended in 40µl of sample loading solution.

Analysis

Phylogenetic inferences were made from DNA sequence alignments generated by Clustal W (Thompson, et al. 1994; Higgins et al. 1996) and modified by multiedit (American Cybernetics, AZ). Nieghbor joining evolutionary distance trees were developed using MEGA 3.1 (Kumar et al. 2004).

Chapter Three

RESULTS

Actin Data Analysis

Actin intron sequence results were analyzed using a neighbor joining tree building algorithm. Each allele recognized for every species was included in the analysis. Numbers of alleles per species varied between two and one.

Actin intron sequences yielded a tree with very little resolution of relationships in the selected sampling of species. One of the only stable relationships was the grouping together of both alleles of *C. terniflora* with 99% bootstrap support (Letter A of Figure 1). There was also 99% bootstrap support for a relationship where the first allele of *C. catesbyana* was a sister group to a large clade comprised of all other taxa except *C. terniflora* (Letter B). A sister group relationship between the first allele of *C. reticulata* and a larger clade comprised of all taxa expect alleles of *C. terniflora* and the first allele of *C. catesbyana* is also supported with a bootstrap value of 63 (Letter C, Figure 1). The only other well supported clade is a grouping together of the first allele of *C. crispa* with the only allele of *C. fusca* for the actin intron sequences (bootstrap value 66, Letter D).

Beyond the relationships described above, there is no other resolution in the tree of Figure 1. In fact there is a large polytomy including all the members of the section *Viorna*; however, this polytomy does not only include member of *Viorna*. Surprisingly,

C. drummondii is a part of this polytomy. According to Essig (1991), *C. drummondii* is very different from members of the section Viorna because it is characterized by the type II seedling.

A maximum parsimony analysis was also performed on the actin intron sequence data set (Figure 2). Except for relationships involving *C. terniflora* and the first allele of *C. catesbyana,* there is no resolution in this tree. This is a direct result of the variation being concentrated in the sequences from *C. terniflora* and *C. catesbyana*. There is only one parsimony informative site in the actin data set when those taxa are removed, while there are 20 parsimony informative sites when all taxa are included. However, the relationships that are formed in the maximum parsimony analysis also appear as well supported relationships in the neighbor joining analysis.

As can be expected with very little support for branches, members of different sections based on the most recent Johnson (2001) classification are scattered throughout the tree. The section *Flammula*, represented by *C. terniflora*, is the only section supported by this analysis.

Although there was very little support for the neighbor joining tree resulting from the actin intron data, this data was still informative in diagnosing hybrid species. The experimentally generated hybrid was a cross between *C. crispa* and *C. reticulata*. Therefore, it was hypothesized that the hybrid sequences would group into two alleles; and that one allele would correspond with *C. crispa* and the other allele would correspond with *C. reticulata*.

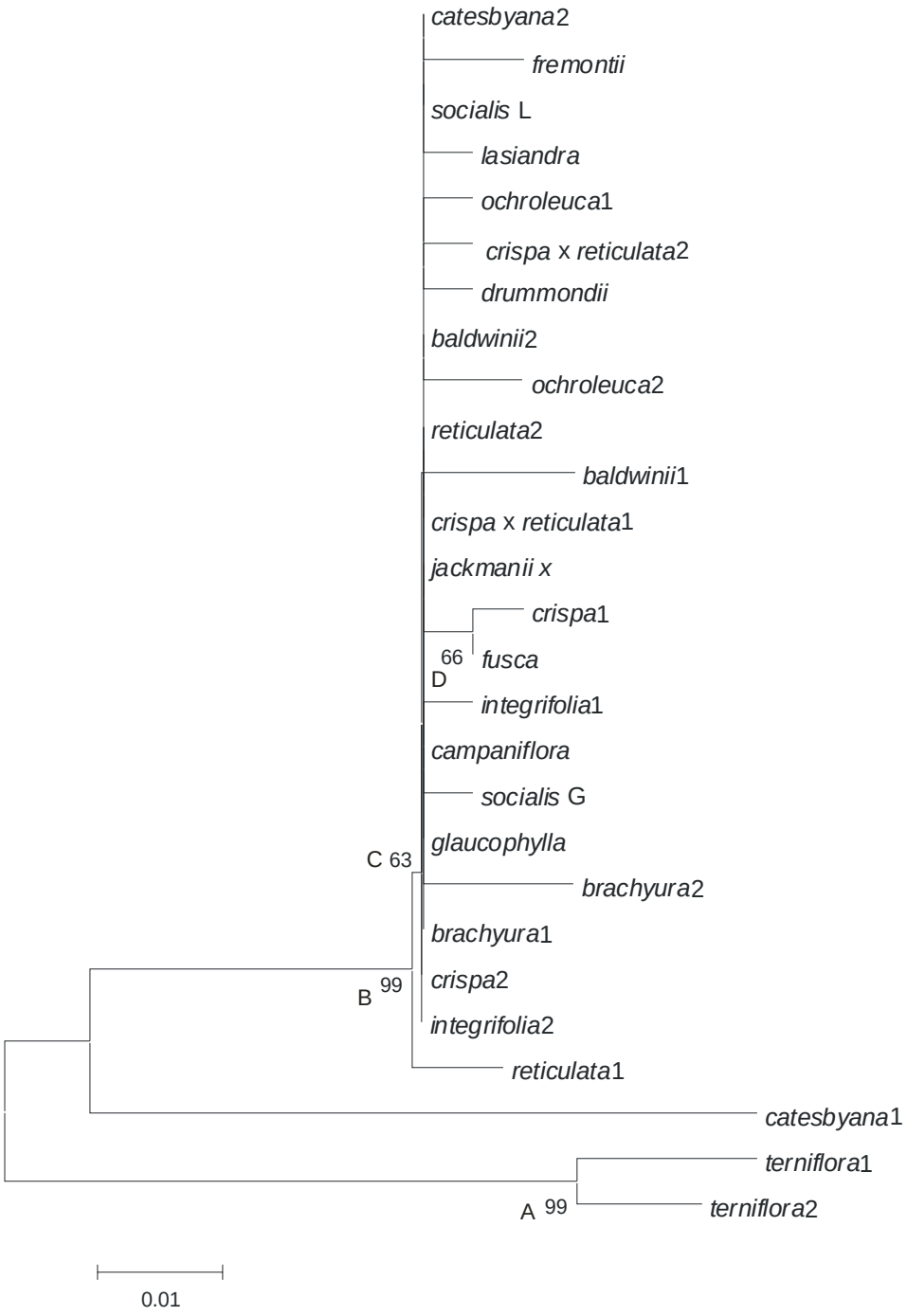

Figure 1. Neighbor joining, evolutionary distance tree for the actin intron sequence data. Bootstrap values over 50 (with 500 replicates) are associated with the branches, and these relationships are labeled with letters. Tree is drawn to scale.

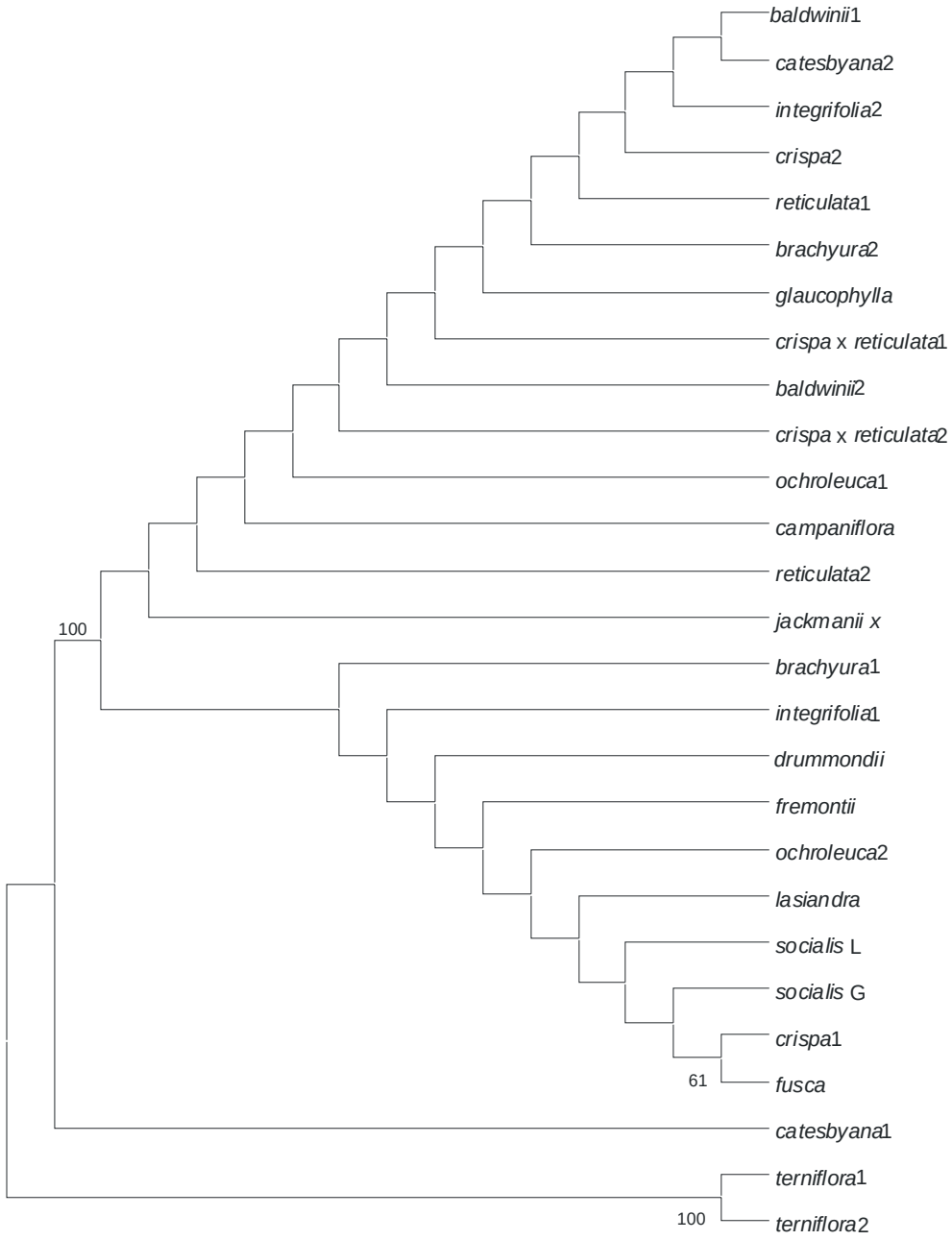

Figure 2. Maximum parsimony tree for the actin intron sequence data. Bootstrap values over 50 (with 100 replicates) are associated with the branches. Letters indicate relationships that appear and are well supported on the neighbor joining tree.

It was found that the sequence of the second allele of *C. crispa* and the second allele of *C. reticulata* were exactly the same as the first allele of the hybrid *C. crispa* x *C. reticulata* (Table 2). The sequence of the second allele of the hybrid matches the sequence of the first allele of *C. reticulata,* but these sequences are not exactly the same. The sequence of the hybrid has an insertion of three thymine bases and a transition from a thymine to a cytosine, while the sequence of the first allele of C. reticulata has a transition from a guanine to an adenine (Table 2). However, these sequences do share other characteristics; they both have an insertion of one thymine base at the same locus and another insertion of two thymine bases.

Table 2. List of substitutions and indels for the hybrid and parent alleles of actin intron sequences.

			Nucleotide Position									
Species/Allele	**Designation**	**Frequency**	37	42	68	83	84	85	86	87	149	176
C. crispa1	*A*	*2*	-	G	C	-	-	-	-	-	C	T
C. crispa2	*B*	*4*	-	G	T	-	-	-	-	-	T	T
*C. crispa x C. reticulata*1	*B*	*4*	-	G	T	-	-	-	-	-	T	T
*C. crispa x C. reticulata*2	*C*	*2*	T	G	T	T	T	T	T	T	T	C
*C. reticulata*1	*D*	*2*	T	A	T	-	-	-	T	T	T	T
*C. reticulata*2	*B*	*3*	-	G	T	-	-	-	-	-	T	T

Unsuccessful Marker Attempts

Before a successful second nuclear plant marker was found, some efforts were undertaken to sequence other markers. Introns of isocitrate dehydrogenase and glyceraldehyde 3-phosphate dehydrogenase were both attempted for collection of sequence data (Weese and Johnson 2005; Strand et al. 1997). In both cases, multiple bands resulted from PCR using published primers. Because size of introns between

genera is not conserved, only intron location is, the appropriate band (PCR amplicon) was not easily evident. There were also not consistent bands for different species. Determination of appropriate band would have required doing a test sequencing of all possible bands which is cost and labor intensive. However, if this was to be accomplished in the future, these two markers may yet be phylogenetically useful.

Nitrate Reductase Data Analysis

Sequences of the third intron of the nitrate reductase (NR) gene were analyzed using a neighbor joining tree building method. Associated with the branches are bootstrap values greater than 50 based on 500 replicates (Figure 3). *C. terniflora* was again used as an outgroup, and each allele present in the sequences for all the species was included in the analysis. Numbers of alleles per species varied again between one and two.

Like the actin intron data, the NR third intron data analysis yields a stable neighbor joining tree with some well supported relationships and other relationships with little to no support. *C. terniflora* was designated the outgroup, and the one allele of this species did group outside the rest of the species in the tree. The two alleles of *C. lasiandra* group together with 68% bootstrap support (Letter A), and the two alleles of *C. baldwinii* also group together with 82% bootstrap support (Letter F) (Figure 3).

According to this analysis, the only allele of *C. campaniflora* forms a sister group with *C. integrifolia* (86% bootstrap support, Letter H). Also, the first allele of *C. socialis* from the Laney population (population L) forms a sister group with both alleles of *C. baldwinii* (57% bootstrap support, Letter E). In addition, the second allele of *C. socialis*

population L forms a sister group with the only alleles of both *C. glaucophylla* and *C. socialis* the Gasden population (population G) (59% bootstrap support, Letter I).

Three more relationships in this analysis are supported with a 50 or more bootstrap value. Each of these is indicated by letters on Figure 3. Letter B represents a sister group relationship between the first allele of *C. brachyura* and a much larger clade consisting of all alleles except those of *C. fusca, C. lasiandra, C. terniflora* supported by a bootstrap value of 86%. Letter C represents a sister group relationship between the *C. jackmanii x* allele and a clade consisting of all alleles except those above and the first allele of *C. brachyura*. This relationship is supported with a 65% bootstrap value. Finally, letter D represents a sister group relationship between a clade consisting of the first allele of *C. socialis* population L and both alleles of *C. baldwinii* and a clade consisting of both alleles of *C. ochroleuca, C. crispa*, *C. reticulata,* and the hybrid and the second alleles of *C. brachyura* and *C. socialis* population L, and the only alleles of *C. fremontii*, *C. drummondii*, *C. campanifora*, *C. integrifolia*, and *C. socialis* population G. The relationship has bootstrap support of 64%.

The clade represented by the letter D is primarily composed of alleles of species in the *Viorna* section (Figure 3), however the relationships within the clade have very little support with the exceptions at letters E and G.

Figure 3. Neighbor joining, evolutionary distance tree for the nitrate reductase intron sequence data. Bootstrap values over 50 (with 500 replicates) are associated with the branches, and these relationships are labeled with letters. Tree is drawn to scale.

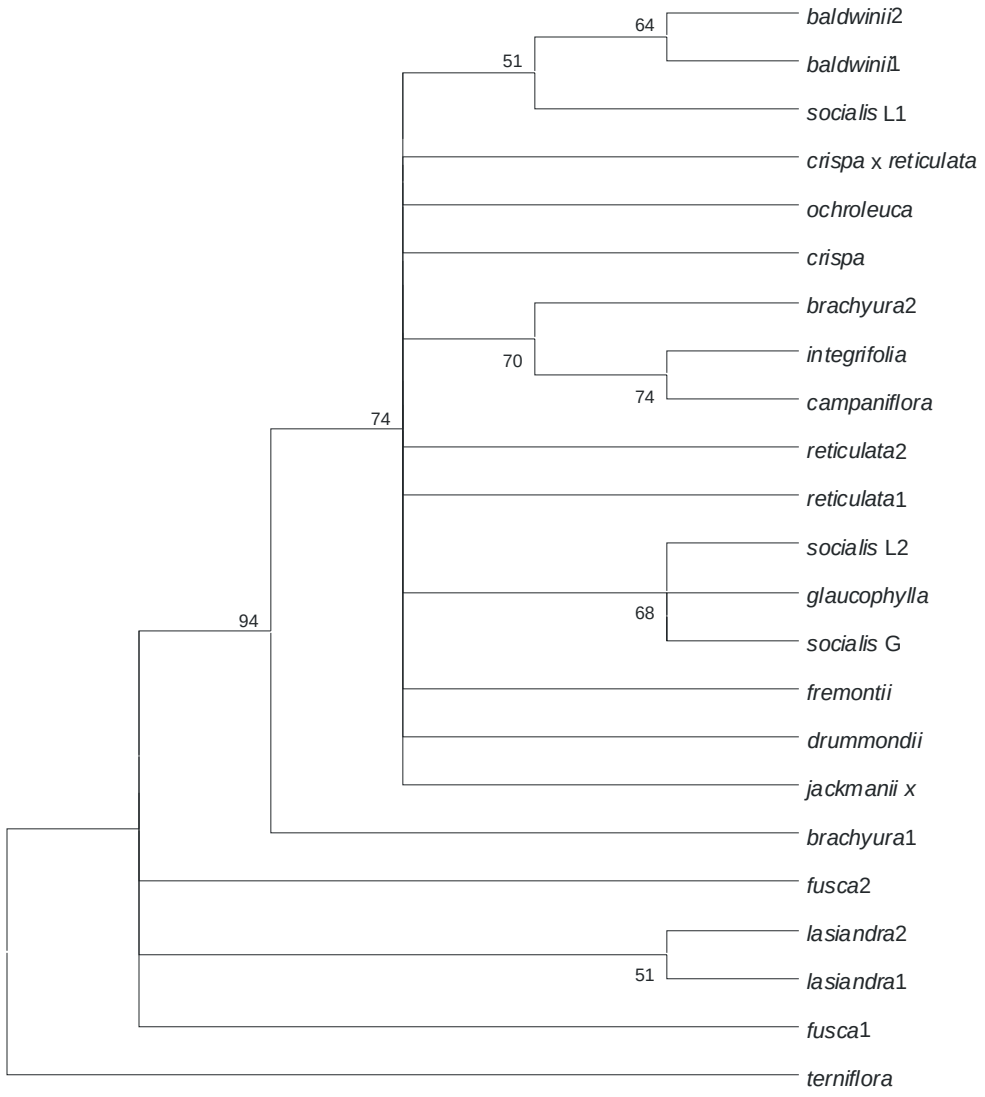

Figure 4. Maximum parsimony tree for the nitrate reductase intron sequence data. Bootstrap values over 50 (with 100 replicates) are associated with the branches, and these relationships are labeled with letters.

A maximum parsimony analysis was also conducted on the nitrate reductase data set (Figure 4). This data set has more variation than the actin data set. There were 25 parsimony informative sites (out of 194 overall sites) when all taxa were included. When the most different taxa were excluded, there were still 9 parsimony informative sites. Still, there are many more parsimony informative sites attributable to the few most distantly related taxa of the analysis than any of the closely related taxa. For this reason, there are well supported relationships between less closely related taxa than between taxa in the *Viorna* group. Many of these relationships remain unresolved.

The maximum parsimony tree did replicate all the relationships that were well supported in the neighbor joining tree except one (Letter D, Figure 3). The bootstrap values on these relationships were also greater than 50.

Similarly to the actin intron analysis, the neighbor joining tree of the third intron of NR is mostly uninformative. Unfortunately, in the case of the third intron of NR, this marker does not diagnose the parentage of the experimentally generated hybrid. Sequences from the *C. crispa* x *C. reticulata* specimen did not sort into two distinct alleles (Table 3). When alleles of both *C. crispa* and *C. reticulata* were determined, it was found that they shared an allele. This shared allele was passed to the hybrid making it homozygous for this marker. Although this does not diagnose the hybrid parentage, the nitrate reductase marker also does not refute the hybrid hypothesis.

Table 3. List of substitutions for the hybrid and parent alleles of nitrate reductase sequences.

			Nucleotide Position		
Species/Allele	**Designation**	**Frequency**	90	126	158
C. crispa	A	4	A	T	T
C. crispa x C. reticulata	A	6	A	T	T
C. reticulata 1	B	1	G	C	C
C. reticulata 2	A	3	A	T	T

Combined Actin and Nitrate Reductase Data Analysis

Sequences of both the actin and nitrate reductase introns were concatenated into one alignment for a final joint analysis. A neighbor joining tree was created with bootstrap values greater than 50 (based on 500 replicated) shown with the branches (Figure 5). Again, *C. terniflora* was used as an outgroup, and each allele of both markers was included in the analysis. Alleles for both actin and NR intron sequences were combined based on what taxa they originated from. In cases where one marker had two alleles, but the other only had one allele, the two unique alleles were concatenated with two sequences from one allele of the other marker within the same taxa.

Analysis of the concatenated data set yields a tree where some relationships are highly supported while others still have little to no support. Highly supported relationships are labeled A through L on Figure 5. Letter A represents the highly supported relationship (bootstrap value 99) that groups the two alleles of *C. terniflora* together. There is also a 99% bootstrap value (B) supporting grouping the two alleles of *C. lasiandra* together. At letter C there is a bootstrap value of 55 that joins the second allele of *C. fusca* into a sister group relationship with the two alleles of *C. lasiandra*.

The relationship at letter D (bootstrap value 86) unites the clade of *C. lasiandra* and the second allele of *C. fusca* and the first allele of *C. fusca* with a larger clade comprised of all other alleles except *C. terniflora*. There is also a highly supported relationship at letter E; the first allele of *C. brachyura* is a sister group to a much larger clade (bootstrap value 86). This clade encompasses all included alleles from the section *Viorna* (except *C. fusca*) in addition to the second allele of *C. brachyura*, *C. campaniflora*, and *C. drummondii*. Letter H also represents a well supported clade that unites both alleles of *C. baldwinii* into a one group (bootstrap value 85); letter G shows a relationship where the first allele of *C. socialis* population L is joined into a sister group relationship with both alleles of *C. baldwinii* (bootstrap value 57). Letter F joins the *C. jackmanii x* allele into a sister group relationship with a larger clade including both alleles of *C. crispa, C. ochroleuca, C. reticulata, C. integrifolia*, and the hybrid specimen, as well as the single alleles from *C. fremontii, C. glaucophylla, C. socialis* (population G), *C. campaniflora*, and the second alleles of both *C. socialis* (population L) and *C. brachyura*.

Two alleles of *C. integrifolia* also group together with a high level of support (95% bootstrap; letter K). These two *C. integrifolia* alleles form a sister group relationship with *C. campaniflora* (bootstrap value 73; letter J). This clade forms a sister group with the second allele of *C. brachyura*, at letter I, with 64% bootstrap value. Finally, a clade uniting *C. glaucophylla* with *C. socialis* (population G) and *C. socialis* (population L) is supported with a bootstrap value of 61% (letter L).

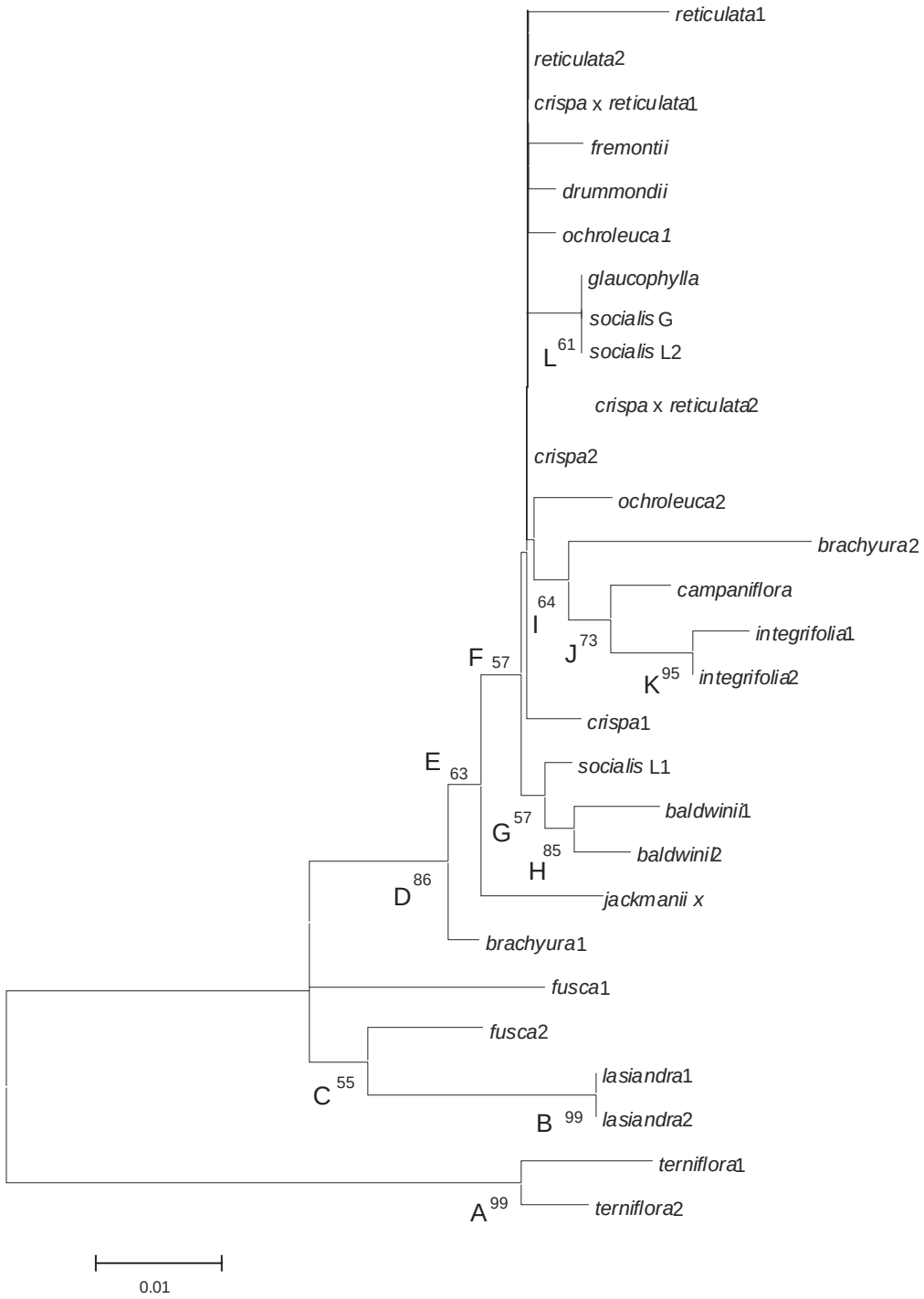

Figure 5. Neighbor joining, evolutionary distance tree for the concatenated sequence data set of both actin and nitrate reductase intron sequences. Bootstrap values over 50 (with 500 replicates) are associated with the branches. Tree is drawn to scale.

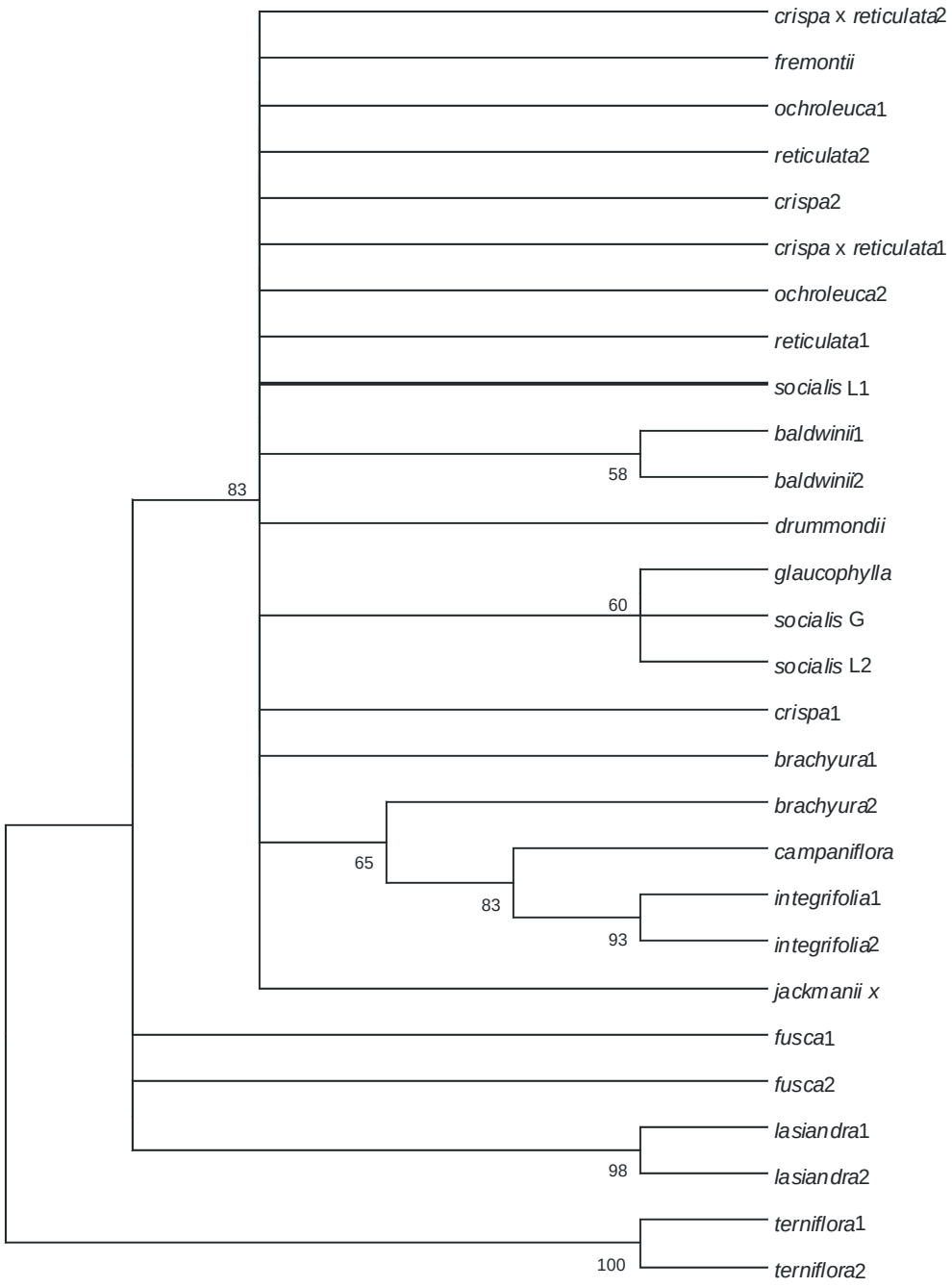

Figure 6. Maximum parsimony tree for the concatenated sequence data set of both actin and nitrate reductase intron sequences. Bootstrap values over 50 (with 100 replicates) are associated with the branches.

Many relationships in this tree, especially those among species within the section *Viorna* that are hypothesized to be more closely related, are still unresolved. Figure 5 illustrates the tree from the concatenated data set of both actin and NR intron sequences where all nodes without at least 50% bootstrap support are collapsed. Figure 5 shows how little information has actually been gleaned from the combined sequences of two intron markers in regards to determining very fine phylogenetic relationships.

A maximum parsimony analysis was also conducted on the concatenated data set (Figure 6). Not all the well supported clades in the neighbor joining tree were replicated in the maximum parsimony tree, but many were. The most interesting clade that appears on both trees is labeled F on Figure 5 and supported by a bootstrap value of 83 on the Figure 6. This clade creates a large polytomy of mostly species from the section Viorna, but not exclusively those species.

Combining the two markers does not give any more information regarding hybrid parentage. Because the hybrid is homozygous for the NR marker, the actin marker is the only one informative about this topic in this study. Because both putative parents also share the allele of the hybrid, the homozygous state of the hybrid for the NR marker does not refute the hybrid hypothesis. However, the NR marker also does not support the results of the actin marker.

Chapter Four

DISCUSSION

Actin Intron Topics

Phylogenetic relationships were not well resolved using actin intron sequences. Very few clades were well supported. One common relationship in all analyses was a clade formed with the two alleles of *C. terniflora*, and that this clade was always a sister group to rest of the tree (most basal). This relationship leads to the belief that this species is the least closely related to the others included in the study. This is reasonable considering that according to Johnson's classification, based on Tamura's original classification, *C. terniflora* is the only species included in the analysis that is a member of the section *Flammula*. *Flammula* is comprised predominantly of Asian species, with a few European, but no species from the New World. Unlike *Viorna*, the flowers have four tepals but are not urn-shaped or campanulate, and they do not have the leathery texture of flowers of *Viorna* (Johnson 2001).

One major reason that there are so few well supported clades in the actin intron sequence alignment, is that there is not enough variation in the data to tease apart the relationships. When considering all taxa, 21% of sites were variable, but only 7.5% of sites were parsimony informative leaving 13.5% of overall sites or 64% of variable sites singletons. Overall, this amount of variation seems sufficient. However, when variable

sites are counted again with out *C. terniflora* or *C. catesbyana* (first allele) in the data set, there are only 7.9% variable sites, 0.3% parsimony informative sites, and 7.5% singletons. The sharp decline in variable sites and, especially, parsimony informative sites after removing the most basal taxa shows that those taxa contain most of the variation. Therefore, it will never be possible to resolve relationships among the other taxa without more data. The data would have to either be another marker with more variability or just keep adding more markers to the current data set.

In the actin data set the most variable region is a thymine rich zone where the number of T nucleotides varies widely among the taxa. The T-rich region is the location of much of the variation observed for this marker in the taxa sampled. However, this variation is not very suitable for phylogenetic reconstructions because it can change too easily from one individual to another regardless of which species that individual belongs to. This region is also riddled with gaps making it less useful for phylogenetic analysis.

Many factors contribute to the paucity of informative variation within the actin data set. Using an intron sequence increases the amount of potential variation because the sequence mutations are not constrained by function. This gives introns a higher mutation rate than either coding regions of nuclear DNA or chloroplast DNA. Unfortunately, of the approximately 266 base pairs that make up the actin marker, only about 105 bp make up the intron. The rest of the sequence is made up of flanking exon regions that have slower mutations rates due to functional constraints.

Another factor contributing to the small amount of observed variation, is the occurrence of one identical allele in ten different taxa analyzed. Not surprising, *C.*

glaucophylla, *C. crispa*, *C. reticulata*, *C. baldwinii*, and *C. integrifolia* share this allele and occur throughout Florida. These species are also all members of the section *Viorna* as proposed by Tamura and Johnson (1968, 2001). Somewhat surprising, *C. brachyura*, and *C. campaniflora* also have this allele. These species are members of the sections *Viticella* and *Pterocarpa*, respectively, and do not grow in close vicinities either. Very suprising, *C. catesbyana* also possesses this allele. The other allele of *C. catesbyana* was one of the most variable taxa in the analysis. This was growing in a botanical garden in the vicinity of other *Clematis,* one of which was *C. baldwinii* which also possessed this allele. However, Essig's seedling morphology classification identified *C. catesbyana* as a having a type II seedling while *C. baldwinii* has a type I seedling. The ability of a type I seedling plant to share genetic material with a type II seedling plant is very unlikely, although technically it is hypothesized that the species are interfertile. The experimentally generated hybrid and a hybrid from Gig Harbor, Washington, *C. jackmanii x,* both possess this allele as well. Possession of this allele by the experimentally generated hybrid is not surprising, since both parent species also possess the allele. Possession of this allele by another hybrid (parentage unknown) from the opposite coast is more puzzling. Obviously, other species in the *Viorna* group included in this study did not share this particular allele, so this sequence does not occur in all *Viorna* or only in *Viorna*. This allele being so common, but inconsistent, may be the result of multiple different processes. The allele may be selected for in some way that is not understood, considering it is hypothesized that typical introns are not selected for or against. This also may have been the original allele of the genus *Clematis,* and not all species that have radiated since the advent of the genus have had mutations to change this allele. One more

possible explanation lies in the fact that, since mutations are random, the occurrence of the same allele in many species is the result of convergent evolution, but is not due to common ancestry. This last explanation is more unlikely due to the fact the allele occurs in ten out of the 17 taxa included in the study. The possibility of the same allele occurring at random two different times seems credible. However, the probability of the same allele occurring at random ten different times out of 17 is so small it is negligible. One last possible explanation is the fact that the different species of *Clematis* sampled are interfertile enough that there is gene flow between them. This gene flow may cause a homogenization of the alleles present in the different species. This gene flow also may have happened in the past, especially at the conclusion of the last ice age when the range of these plants was much smaller than it is now, and more of the species may have overlapped in geographic distribution.

Future studies may sample many individuals of the same taxa to determine how many alleles each one has. Testing many individuals of one taxa for variation may also yield more information about whether all the members of *Viorna* actually do have the common allele in their gene pool. This analysis was conducted with just one individual for most taxa; if more individuals were tested, the common allele may have been a part of a different individual's genome, even though it was not in the first individual's genome.

One possible conclusion that can be made from the lack of variation, even in an intron sequence, is that the time since divergence for this group of taxa is very small. In other words, there has not been sufficient time for all the different taxa to accumulate mutations for all unique alleles.

Another possible conclusion that can be made from the lack of variation is that, especially for closely related species, reticulate evolution (hybridization) has clouded the signal of the molecular data. If different species are not completely reproductively isolated, as many plants are not, gene flow is possible between species to varying degrees. In this way common alleles can be passed between species almost as if closely related species are one large population.

Using the actin intron marker to diagnose hybrid species may not be as clear cut as once hypothesized. When the experimentally generated hybrid (*C. crispa* x *C. reticulata*) actin intron sequences were first analyzed, there were clearly two alleles present. Two different alleles first support the use of this marker to diagnose hybrids and determine their origin. However, one of the two alleles is the very common allele. Unfortunately, *C. crispa* and *C. reticulata* both possess the common allele as well. Therefore, it is not immediately clear which parent species contributed that allele to the hybrid. When the other hybrid allele is considered, it is not found to exactly match any of the other parent alleles.

There are many possible explanations as two why the second hybrid allele does not match any other parent allele. Either the second hybrid allele was contributed by *C. crispa* and one or both underwent further mutations so that the two sequences no longer match, or the second hybrid allele was contributed by *C. reticulata* and one or both mutated. Another possibility is that an individual of *C. crispa* or *C. reticulata* does have an allele that exactly matches that of the hybrid, but this was not sequenced in this study. The experimentally generated hybrid was developed more than ten years ago, which does provide for enough time for mutations to occur in the hybrid or the putative parents

which would cause the sequences of this marker to no longer match. However, the chance that three separate mutations all accumulated in such a short time is very small.

Another explanation is that, since the hybrid does not exist in a large population of similar hybrid taxa, but exists in a small area where many other species of Clematis grow, there may have been further hybridization with another taxa at a later unknown time. This unknown taxa may have contributed the second allele sequenced in this study.

The most likely explanation is that because the exact same plants that created the hybrid were not sampled, only plants of the same species, the exact same allele that was contributed to the hybrid was not sequenced.

One more complicating factor is the occurrence of the exact same allele, except for one transition, in *C. ochroleuca*. This coincidence must be the result of convergent evolution, not common ancestry, because at no point was *C. ochroleuca* growing in the vicinity of the experimentally generated hybrid. Therefore, *C. ochroleuca* could not have contributed its allele to the hybrid. However, if this hybrid was truly from unknown origin, sequence data would certainly suggest that *C. ochroleuca* was a parent species.

It is impossible at this time to determine the history of the hybrid using the available data. Because one of the alleles present is shared by both parent species, determining which parent contributed that marker is impossible. Because the other allele is not matched exactly by either parent, this also is impossible to trace. Even though, the first allele of *C. reticulata* only differs from the second hybrid allele in three ways, it is highly unlikely that there were three mutations from the recent hybridization event to present. The *C. reticulata* allele has one transition from a guanine to an adenine which no other allele shares. This substitution may have occurred after *C. reticulata* contributed the

allele to the hybrid. The second hybrid allele also has one transition substitution that is shared by no other allele. This transition from a thymine to a cytosine may also have occurred after the inception of the hybrid. The last difference is in the number of thymine nucleotides in a very T-rich area of the intron. In the C. reticulata allele there are nine thymine bases in a row, while in the hybrid there are 12. The most likely mechanism for this disparity in length of T-rich region is replication slippage. Replication slippage is common in areas of microsatellite DNA and involved DNA polymerase pausing and momentarily dissociating from the template during replication. When DNA polymerase reforms the bond with the template DNA, it is not in the same place as where it left off (Viguera et al 2001). If the DNA polymerase reattaches further downstream from the site of dissociation, the replicated DNA will have a deletion of the skipped nucleotides. If the DNA polymerase reattaches further upstream from the site of dissociation, there will be an insertion into the new replicated DNA strand. A preliminary survey of all the alleles reveals a great disparity in the number of thymine nucleotides in this region across all taxa. The most common number being seven thymine bases, but there may be as many as twelve in other alleles.

This process of insertion is what is hypothesized to have happened in the hybrid second allele to give it three more thymine nucleotides than the parent *C. reticulata* allele. Any one of the above mutations alone may be conceivable, but having all three in such a short time does not seem feasible.

To improve upon this in the future a new cross should be made where the hybrid origin is known and the actual parent plants are sampled. Multiple individuals of the hybrid from multiple seeds of one hybrid cross should also be sampled. By just including

one hybrid individual not all possible allelic combinations were represented. One hybridization event would have produced many seeds, providing multiple pollen grains came into contact with multiple stigmas. Some percentage of those seeds would be viable depending on compatibility between the two parent species' genomes (Judd et al. 1999). Planting all those seeds would yield hybrid plants that may differ in which alleles came from which parent. By analyzing sequences from multiple hybrid plants derived from seeds from the same flower, it would be possible to find out more about what the actin intron sequences can tell us about hybridization.

The taxa from gig harbor, *C. jackmanii x* was a putative hybrid. Unfortunately, the actin intron sequences for this taxa did not sort into two distinct alleles. All sequences for this taxa were the same allele, specifically the very common allele. Therefore, the actin intron marker can not support or refute the hybrid hypothesis for *C. jackmanii x*.

Nitrate Reductase Intron Topics

In general, many of the themes of the actin intron sequence analysis hold true for the NR intron sequence analysis as well. The neighbor joining tree that resulted from the NR intron sequences also had few well supported relationships. At the base of the tree are the alleles of *C. terniflora* (grouping together with a bootstrap value of 99), *C. lasiandra* (grouping together with a bootstrap value of 66), and *C. fusca*. There is not much resolution in relationships between these species, however, these three species all have Asian geographic distributions. *C. terniflora, C lasiandra,* and *C. fusca* are all classified into different sections by Tamura and later by Johnson as well (1987, 2001). These Asian taxa must be the most different from the other (mostly North American) taxa which is

why they are grouping outside of the major clade of the tree at letter C (Figure 3). Not all the Asian taxa fall unresolved at the base of the tree. The two alleles of *C. brachyura* fall within the clade marked at letter C, but they do not fall together. The first allele of *C. brachyura* falls at the base of the clade, forming a sister group with the rest of the North American and European species which form a clade at letter D.

One other well supported clade, at letter F of Figure 3, also groups together taxa of similar geographic distribution. *C. campaniflora* and *C. integrifolia* are the two European species included in the study. These two species are also not in the same section according to Tamura (1987). However, the clade at letter F is better evidence for possibly modifying the existing classification than the unresolved relationship of the Asian species.

Among North American species there is very little resolution. Like the actin intron analysis, the NR intron analysis was also lacking in sufficient variation to tease apart the relationships of the closest related species. In the nitrate reductase data set 31% of sites were variable, but 58% of variable sites were singletons. Singletons do not contribute much information to a neighbor joining tree. In the data set, 12.9% of sites were parsimony informative. As with the actin data set, the variation is concentrated in the sequences of the few least closely related taxa. When the sequences of *C. terniflora*, *C. lasiandra*, and *C. fusca* were removed from the data set, the number of variable sites was reduced to 14.9%, 69% of these variable sites being singletons. Removing those taxa also reduced the parsimony informative sites to 4.6%. This again explains why the most basal groups are fairly well supported, but why there is very little resolution in the rest of the tree in both the neighbor joining and maximum parsimony analyses.

Again there were many taxa that shared one most common allele. In this case *C. crispa, C. fremontii, C. reticulata, C. ochroleuca* and *C. crispa* x *C. reticulata* all shared one allele and are all members of the proposed *Viorna* section. However, *C. drummondii* also shared this allele, but is considered a member of the *Viorna* section. Obviously, there was no resolution obtained in the relationships of the alleles with the same sequence.

In two separate analyses, using two different intron markers, *C. drummondii* grouped in a clade with other members of the *Viorna* section. Judging by morphological data, *C. drummondii* has quite disparate characters than those of the *Viorna* section. For example, *C. drummondii* (as well as the other members of its subsection) are dioecious, and have white flowers, while members of section *Viorna* are all monoecious with more showy flowers typically urn shaped. Admittedly, the clade that united *C. drummondii* with members of the *Viorna* section has almost no further resolution, and *C. drummondii* may be grouping with the other taxa based on their North American connection.

Another interesting relationship from the NR intron sequence analysis is the clade, letter H, uniting the first allele of *C. socialis* population L into a sister group relationship with the two alleles of *C. baldwinii*. This is interesting because the other allele of *C. socialis* population L falls into a different polytomic clade with *C. socialis* population G and *C. glaucophylla*. This may support a hypothesis that *C. socialis* population L is a hybrid between *C. socialis* population G and *C. baldwinii*. This may also just be a coincidence that the two alleles of *C. socialis* population L are very different from one another and one allele happens to have a sequence very similar to *C. baldwinii*. The coincidence hypothesis is supported by the fact that *C. socialis* is an endangered species with all four known populations occurring in Alabama (Kral 1982;

Boyd and Hilton 1994). *C. baldwinii*, however, has only been documented to occur in Florida (Johnson 2001). Another possibility is that *C. socialis* population L is a hybrid, but not between *C. socialis* population G and *C. baldwinii*, but between *C. socialis* population G and another species not included in this study, but with an NR intron sequence related to that of *C. baldwinii*.

Unfortunately, the experimentally generated hybrid was homozygous for the NR intron marker. This makes sense because both *C. crispa* and *C. reticulata* possess the most common allele for this marker. Since one allele is so prevalent for this marker, it is not a good hybrid diagnostic tool. The chance of the same allele coming together from two different taxa is too great, and when that process occurs detecting the hybrid state is impossible.

In general, there are fewer alleles in this marker than were found for the actin intron marker. Only six of the 17 taxa analyzed had two alleles rather than only one. However, the differences that did exist were more pronounced than those for the actin intron marker, in general. Disregarding the six taxa that shared the most common allele, the other alleles present in the analysis were very different from one another, with many substitutions. Even though the marker was only about 200 base pairs long, the whole length of it was made up of intron. Therefore, there were many substitutions throughout the sequences for different alleles. Unlike the actin intron sequences, there was no region dominated by one base or any short repeats so replication slippage was not an issue. Ultimately, this means that the phylogenetic utility of this marker is greater than the actin intron marker, even though overall the NR intron is shorter.

This analysis was specifically conducted using the third intron of the nitrate reductase gene. Primers were also available for the second intron of the nitrate reductase gene. Preliminary analysis suggests that this intron may be longer than the third, and therefore have more phylogenetic informative characters. Future analyses should include exploring the phylogenetic utility of this intron in conjunction with the third nitrate reductase intron and the actin intron. PCR needs to be optimized to amplify the second intron preferentially, before sequencing efforts can be efficient.

Combined Actin and Nitrate Reductase Intron Topics

When data sets are combined, the goal is to get more information from the sum of the parts than is available by the parts individually. Molecular data sets are combined by concatenating the sequences. Very little information was added to the analyses from the concatenated data set. The major relationships delineated in the nitrate reductase analysis were reiterated in the analysis of the concatenated data set, but with lower bootstrap values. The neighbor joining tree resulting from the concatenated data set further supports the conclusion that the actin data set does not yield informative phylogenetic relationships because the tree had no further resolution with the added actin sequence data.

Previous work on this genus using the actin intron resulted in a tree with better resolution and higher bootstrap values, but the taxa included were from many different sections within *Clematis*. When the taxa sampling is limited to mostly one section, *Viorna*, there is little to no resolution within the section. In the previous study, *C. brachyra* showed anomalous relationships within the trees developed (Slomba 2002);

anomalous results were consistent in this study. *C. brachyura* had two alleles for both markers tested; when the markers where concatenated based on allele the two sequences grouped into very different places on the tree (Figure 3). The first allele of *C. brachyura* groups at the base of a larger clade composed mostly of members of *Viorna* (which *C. brachyura* is not). Placement here would be hypothesized. However, the second allele of *C. brachyura* groups well within the clade of *Viorna* species. In fact, the second allele groups in a sister relationship to a clade of *C. campaniflora* and *C. integrifolia*. It may be that *C. brachyura* has a faster rate of evolution than other taxa in *Clematis*. Some alleles are becoming less closely related to each other than they are to other alleles from different taxa. This would make *C. brachyura* appear more closely related to other taxa, but this relationship would actually be due to random mutation and not close common ancestry. One way to test the relationships of *C. brachyura* would be to sample from many individuals to see what alleles are present in the population, rather than in just one individual. This may refute the results shown here and in previous analyses, or it may verify an interesting evolutionary phenomenon in *C. brachyura*.

Conclusions

Overall conclusions that come from this study include the fact that intron sequence data will not provide a well resolved tree for closely related species of *Clematis* in the *Viorna* section using typical phylogenetic methods. The probability of interfertility allowing gene flow between species (hybridization) is so great that reticulate evolution is confounding the image of a tree that assumes bifurcation. Because of this problem, adding more intron sequences may not provide any more resolution for these

phylogenies. Instead, a population level approach should be taken in the future that would allow for frequencies of alleles to be calculated and different methods of phylogeny can be employed like population analyses that make minimum spanning networks. In order to accomplish this, more sampling should be done to multiple individuals within the current species included. Sampling more individuals would also help overcome the problem encountered when diagnosing hybrid history. Many, if not all, possible allelic combinations must be sampled to determine the hybrid parentage molecularly.

LITERATURE CITED

Aguilar, J.F., J.A. Rosselló, G.N. Feliner. 1999. Nuclear ribosomal DNA (nrDNA) concerted evolution in natural and artificial hybrids of *Armeria* (Plumbaginaceae). Molecular Ecology. 8:1341-1346.

Boyd, R.S. and C.D. Hilton. 1994. Ecological studies of the endangered species *Clematis socialis* Kral. Castanea. 59(1):31-40.

Cros, J., M.C. Combes, P. Trouslot, F. Anthony, S. Hamon, A. Charrier, and P. Lashermes. 1998. Phylogenetic analysis of chloroplast DNA variation in *Coffea* L. Molecular Phylogenetics and Evolution. 9(1): 109-117.

Essig, F.B. 1991. Seedling morphology in *Clematis* (Ranunculaceae) and its taxonomic implications. Sida. 14(3): 377-390.

Ferguson, D. and T. Sang. 2001. Speciation through homoploid hybridization between allotetraploids in peonies (*Paeonia*). PNAS. 98(7):3915-3919.

Franzke, A. and K. Mummenhoff. 1999. Recent hybrid speciation in *Cardemine* (Brassicaceae)—conversion of nuclear ribosomal ITS sequences in situ nascendi. Theor. Appl. Genet. 98:831-834.

Friedman, H., J.W. Vos, P.K. Hepler, S. Meir, A.H. Halevy, and S. Philosoph-Hadas. 2003. The role of actin filaments in the gravitropic response of snapdragon flowering shoots. Planta. 216:1034-1042.

Foyer, C.H., M. Parry and G. Noctor. 2003. Markers and signals associated with nitrogen assimilation in higher plants. Journal of Experimental Botany. 54(382): 585-593.

Hardig, T.M., P.S. Soltis, and D.E. Soltis. 2000. Diversification of the North American shrub genus Ceanothus (Rhamnaceae): conflicting phylogenies from nuclear ribosomal DNA and chloroplast DNA. American Journal of Botany. 87(1): 108-123.

Higgins, D.G., J.D. Thompson, and T.J. Gibson. 1996. Using CLUSTAL for multiple sequence alignments. Meth. Enzymol. 266:383-402.

Hillis, D.M., C. Moritz, and B.K. Mable. 1996. Molecular Systematics 2nd ed. Sinauer Assoc. Inc. Sunderland, MA.

Hoot, S.B. and J.D. Palmer. 1994. Structural rearrangements, including parallel inversions, within the chloroplast genome of *Anemone* and related genera. Journal of Molecular Evolution. 38:274-281.

Howarth, D.G. and D.A. Baum. 2002. Phylogenetic utility of a nuclear intron from nitrate reductase for the study of closely related plant species. Molecular Phylogeny and Evolution. 23:525-528.

Johnson, Magnus. 2001. The Genus Clematis. Magnus Johnson Plantskola AB. Södertälje, Sweden.

Judd, W.S., C.S. Campbell, E.A. Kellogg, and P.F. Stevens. 1999. Plant Systematics: A Phylogenetic Approach. Sinauer Associates Inc. Sunderland, Massachusetts.

Kral, R. 1982. A new *Clematis* from northeastern Alabama. Rhodora. 84(838):285-291.

Koontz, J.A., P.S. Soltis, and D.E. Soltis. 2004. Using phylogeny reconstruction to test hypotheses of hybrid origin in *Delphinium* section *Diedropetata* (Ranunculaceae). Systematic Botany. 29(2): 345-357.

Kumar, S., K. Tamura, and M. Nei. 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings in Bioinformatics. 5:150-163.

Li, W.H. 1997. Molecular Evolution. Sinauer Assoc. Inc. Sunderland, MA.

Mathur, J. and M. Hülskamp. 2002. Microtubules and microfilaments in cell morphogenesis in higher plants. Current Biology. 12: R669-R676.

Mathur, J. 2004. Cell shape developments in plants. Trends in Plant Science. 9(12):583-590.

Miikeda, O., S. Koga, T. Handa, and T. Yukawa. 1999. Subgeneric Relationships in *Clematis* (Ranunculaceae) by DNA sequences. In: S. Andrews, A.C. Leslie and G. Alexander (Editors). Taxonomy of Cultivated Plants: Third International Symposium, pp. 355-358. Royal Botanic Garderns, Kew.

Moniz de Sa, M. and G. Drouin. 1996. Phylogeny and substitution rates of angiosperm actin genes. Molecular Biology and Evolution. 13(9): 1198-1212.

Rieseberg, L.H. 1997. Hybrid Origins of Plant Species. Annu. Rev. Ecol. Syst. 28:359-89.

Rossetto, M. 2005. A simple molecular approach for identifying a rare *Achronychia* (Rutaceae) provides new insights on its multiple hybrid origins. Biological Conservation. 121:35-43.

Sharma, P., R.S. Dubey. 2005. Modulation of nitrate reductase activity in rice seedlings under aluminium toxicity and water stress: role of osmolytes as enzyme protectant. Journal of Plant Physiology. 162: 854-864.

Slomba, J.M., J.R. Garey, and F.B. Essig. 2004. The Actin I intron – A phylogenetically informative DNA region in *Clematis* (Ranunculaceae). Sida. 21(2):879-886.

Soltis. D.E. and R.K. Kuzoff. 1995. Discordance between nuclear chloroplast phylogenies in the *Heuchera* group (Saxifragaceae). Evolution. 49(4):727-742.

Strand, A.E., J. Leebens-Mack, and B.G. Milligan. 1997. Nuclear DNA-based markers for plant evolutionary biology. Molecular Ecology. 6:113-118.

Tamura, M. 1968. Morphology, ecology and phylogeny of the Ranunculaceae. Sci. Rep. S. Coll. N. Coll. Osaka Univ. 16:21-43.

Tamura, M. 1987. A classification of the genus *Clematis*. Acta Phytotax. Geobot. 38:33-44.

Thompson, J.D., D.G. Higgins, and T.J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-80.

Viguera, E., D. Canceill, and S. Dusko Ehrlich. 2001. Replication slippage involves DNA polymerase pausing and dissociation. The EMBO Journal. 20(10):2587-2595.

Weese, T.L., L.A. Johnson. 2005. Utility of NADP-dependent isocitrate dehydrogenase for species-level evolutionary inference in angiosperm phylogeny: A case study in *Saltugilia*. Molecular Phylogenetics and Evolution. 36:24-41.

Zhou, R., S. Shi, and C. Wu. 2005. Molecular criteria for determining new hybrid species—An application to the *Sonneratia* hybrids. Molecular Phylogenetics and Evolution. 35:595-601.

Zomlefer, W.B. 1994. Guide to Flowering Plant Families. University of North Carolina Press. Chapel Hill, North Carolina. Pg. 42-45.