INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

BIOSYSTEMATICS OF THE GENUS *PHYSALIS* (SOLANACEAE)

by

Mary Kathryn ("Maggie") Whitson

Biology Department
Duke University

Date: Dec. 3, 2001

Approved:

Paul S. Manos, Supervisor

D. E. Stone

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Biology in the Graduate School
of Duke University

2001

UMI Number: 3041321

UMI Microform 3041321

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT
(Botany - Systematics and Evolution)

BIOSYSTEMATICS OF THE GENUS *PHYSALIS* (SOLANACEAE)

by

Mary Kathryn ("Maggie") Whitson

Biology Department
Duke University

Date: Dec. 3, 2001

Approved:

Paul S. Manos, Supervisor

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biology in the Graduate School of Duke University

2001

ABSTRACT

Physalis (Solanaceae) includes 75+ species of yellow-flowered herbs, commonly called tomatillos. Mexico is the center of diversity, and *Physalis* is almost entirely limited to the New World, with a single species native to the Old World. *Physalis* is characterized by a papery, inflated calyx which surrounds the berry. However, similar structures are found in the physaloid genera. To define the generic limits of *Physalis*, sequence data from the ITS region and the *waxy* gene were used to generate a phylogeny of subtribe Physalinae. Morphologically typical New World species of *Physalis* formed a monophyletic group, which did not include *P. alkekengi*, the type species and only Old World native taxon. The monotypic genus *Margaranthus* nested within *Physalis*, and should be recognized as *P. solanaceous* (Schlecht.) Axelius. All other physaloid genera, including *Quincula*, *Chamaesaracha,* and *Leucophysalis*, were distinct from *Physalis*.

A monophyletic species complex, the U.S. rhizomatous perennial *Physalis,* was used as a study group to examine interspecific patterns of ITS variation, and the interplay between secondary structure and levels of nucleotide variation. A *waxy* gene tree provided the phylogenetic hypothesis, and the results of phylogenetic analysis of ITS sequences were compared to this. This comparison indicated that putative hybridization and incomplete lineage sorting were acting to obscure phylogenetic patterns. Secondary structure was linked to the placement of large indels, which all occurred at the tips of single-stranded loop structures.

A variety of techniques, including phylogenetic analysis of DNA sequences, elliptic Fourier analysis of leaf shapes, and crossing studies were used to study the

biosystematics of an endemic *Physalis* from northern Florida and related species. The DNA data suggested that *P. heterophylla* was sister to the Floridian taxon. The two species can be crossed, but are geographically isolated, and morphological and molecular characters generally distinguish them. Thus, the Floridian taxon was described as a new species, *P. diminuta*.

Finally, to solve the problem of *P. alkekengi*, a type species not closely related to the rest of its genus, a proposal was made to conserve *Physalis* with a conserved type (*P. pubescens* L.), and provide a new genus for *P. alkekengi*.

ACKNOWLEDGEMENTS

This work would not have been possible without the help and encouragement which so many people generously provided. I would especially like to thank Mona and James Whitson, my parents, for always being there when I needed them, and for enthusiastically supporting my desire to become a botanist. My mother has been a willing field assistant and intrepid collector of plants. She creatively used a business trip to Kansas as an opportunity to enrich the Duke University herbarium, selflessly smuggling pressed *Physalis* specimens and other secondary-compound-enriched dicots through the airport. My father, number one plant spotter and field assistant extraordinaire, braved ticks, banana spiders, heat, humidity, and mosquitoes the size of pelicans, all for the pleasure of driving me through the back roads of Florida, Texas and Mississippi in search of *Physalis*. Dad has been a good sport and great company, even when pressing spiny *Solanum*s, digging plants with rhizomes to China, and dealing with suspicious locals.

I would also like to thank my advisor, Paul Manos, for his positive attitude and lots of good advice. He has patiently read, re-read and then read again various grant proposals, chapter drafts, re-drafts and other works in progress. He has also carefully fostered my caffeine addiction, introduced me to XTC music, and has not complained about the presence of slugs in the lab.

My committee members, Vicki Funk, John Mercer, Jon Shaw, and Don Stone, are also deserving of many thanks. They have always taken the time to answer my (many) questions, and have provided lots of good suggestions and reading material. They

did a great job reading my thesis, had lots of constructive comments, and the resulting papers will be much improved because of their input.

I couldn't have done any of this research without funding. The majority of my funding came from the Mellon Foundation, by way of the Duke University systematics program. Sigma Xi and the Aleane Webb Dissertation Research Fund also generously provided monetary aid.

I have gotten lots of great Solanaceae information from both Lynn Bohs and Dick Olmstead. They have also very generously provided me with some rather crucial DNA samples, for which I am extremely grateful. At Duke, Dr. Wilbur has provided lots of encouragement, good gossip, nomenclatural advice, herbarium space, and help in unearthing various obscure taxonomic references, and Dr. W. Culberson wrote the majority of the botanical Latin which was necessary for my new taxon descriptions.

For helping to make graduate school a more entertaining and enriching experience, I would like to thank my friends and fellow graduate students at Duke, including but not limited to: Janneke HilleRisLambers, Anne Pringle, Leonie Moyle, Jason McLachlan, Stuart McDaniel, Verena Lu, Mac Alford, and Bryan Smith. The Manos lab crew, both past and present, including Jenny Arrington, Chuck Cannon, Navina Hamilton, Jay Horn, Alejandra Jaramillo, and Kyle Williams, has also been a great source of advice, inspiration, and people to go get coffee with.

I would also like to thank the botany and biology department staff members who have so patiently helped me with various necessary tasks. Without Beverly Calhoun, Marcia Kirinus, and the other great people in the greenhouse, my plants would have been

much less happy. Sherri Herndon and Jody Soeder have done a great job mounting my herbarium specimens and helping me keep up with my loans. Susan Gerbeth Jones and Susan Gillispie have happily helped solve many of my computer problems. Anne Lacey, Caroline Usher and Blythe Boquist have listened to me whine about everything from photocopiers to TA assignments, and have always been willing to help. Renee Puryear and Jim Tunney have kept my lab funds flowing, and Lisa Bukovnik has helped me spend them well.

Finally, I would like to thank all of the members of the Duke botany and biology departments for making my time in graduate school enriching, entertaining, and educational in a number of ways. It has been a wonderful six and a half years.

TABLE OF CONTENTS

Chapter 1.

Untangling *Physalis* from the physaloids: a two-gene phylogeny of the Physalinae

INTRODUCTION:

The most arresting feature of the genus *Physalis* L. is the calyx, which becomes greatly expanded in fruit, inflating until it completely envelops the berry. *Physalis* is one of the largest genera in the Solanaceae, with 75-90 species, most of which occur in Mexico. With one notable exception, all species are native to the New World, though cultivated species and weedy annuals have been introduced to warm areas worldwide. Of the species grown for their edible fruits, one of the best known is *P. philadelphica*, the tomatillo, which is used in salsa verde. The Chinese lantern plant, *P. alkekengi*, is also cultivated for its fruits, but its use is more often ornamental than culinary – the fruiting branches, with large, inflated red calyces embellishing their berries, are often used in flower arrangements. Several species have been used medicinally, and recent research has focused on potential antibacterial and antitumor properties of their secondary compounds (Chiang et al., 1992; Kennelly et al., 1997; Pietro et al., 2000).

The typical *Physalis* is an herb with solitary, bee-pollinated, axillary yellow flowers (Sullivan, 1984b). The nodding, bell-shaped corollas are unlobed, with darkly-spotted throats. Once pollination has occurred, the calyx expands until the developing berry is entirely hidden, often touching the fruit only at the base. In some species the

mature calyces turn yellow or orange. The many-seeded berries range from greenish to yellow to tangerine, and are sometimes flushed red or purple.

Historically, *Physalis* has been divided into several large species groups on the basis of gross morphological characters (e.g. habit, hair type, calyx angles) (Martinez, 1993). The most recent infrageneric revision used both traditional morphological characters, as well as some new ones, such as trichome surface sculpturing, to explicitly define subgenera and sections (Martinez, 1999. See Table 1.1). However, the relationships both within and between the sections remain unclear.

Accressant calyces appear throughout the Solanaceae, but the highly inflated calyx found in *Physalis* is unusual. Though this feature makes *Physalis* one of the easiest Solanaceous genera to recognize, the circumscription of the genus has been confounded by several small genera called physaloids. As in *Physalis*, these taxa have axillary flowers and anthers which open by longitudinal slits. In fruit, the physaloids share some amount of calyx expansion, though the calyces do not necessarily become inflated. Previous taxonomic treatments disagree on which of the physaloid genera should go into *Physalis*, as well as on which genera are physaloids. *Archiphysalis*, *Athenaea*, *Deprea*, *Exedeconus*, *Jaltomata*, *Physalisatrum*, *Nicandra*, *Saracha* and a host of others have all been associated with *Physalis*. In North America, the "debate" has centered on five small physaloid genera, *Chamaesaracha*, *Leucophysalis*, *Margaranthus*, *Oryctes*, and *Quincula*, whose affinities to *Physalis* remain uncertain (Rydberg, 1896; Averett, 1970; Barboza, 2000). With the exception of *Oryctes*, all of these genera have at one time been

placed in *Physalis* (Axelius, 1995). Basically, the calyx is key. If it has a big one, maybe it's related.

Chloroplast DNA data have greatly clarified the situation by helping refine the circumscription of the tribe Physaleae, which contains many of the physaloid genera (D'Arcy and Averett, 1996; Olmstead et al., 1999). Four subtribes have also been proposed on the basis of cpDNA phylogenies, including the Physalinae, which contains *Physalis* and seven other genera (Table 1.1). Included among these are the five North American physaloid genera, *Witheringia*, and its segregate *Brachistus* (D'Arcy et al., 1981). Because most species of *Witheringia* lack any calyx expansion in fruit, it has not traditionally been considered physaloid. However, cpDNA data placed it firmly at the base of the Physalinae clade (Figure 1.1), and *Witheringia* does share the nodal inflorescences and longitudinally dehiscent anthers characteristic of the physaloids. Finally, the recent segregation of *Tzeltalia* from *Physalis* has given rise to a new physaloid genus which should be considered a member of subtribe Physalinae (Estrada and Martinez, 1998).

Physaloid genera aside, there are other taxonomic issues within *Physalis*. There are several morphologically unusual species of *Physalis* which have been noted by many a taxonomist, but have not been removed from the genus. Oddly enough, the type of the genus, *P. alkekengi* L., may be numbered among the morphologically atypical species. *Physalis alkekengi* is the only Eurasian-native species in the genus, as well as in the Physalinae. Because of its attractive red fruiting calyces, this species has long been grown as an ornamental throughout China, Japan and Europe. It is unclear where *P.*

alkekengi originated, though China has been suggested (Hendrych, 1989; Olmstead et al., 1999). In his 1896 monograph of the genus, after establishing *Leucophysalis*, Rydberg stated, "If...*P. alkekengi* could be also removed, the genus would be a very natural one." Though the morphological differences may seem slight, cpDNA phylogenies do place the North American genera *Chamaesaracha* and *Margaranthus* between *P. alkekengi* and other species of *Physalis* (Mione et al., 1994; Olmstead et al., 1999).

Since the 1950s, a wide range of cytological, biochemical and morphological data have been collected for various sets of *Physalis* species and physaloid genera, but few studies have included both the majority of these genera and a broad sampling of species from within *Physalis*. The purpose of this study is to examine the relationships between the genera in subtribe Physalinae and the species relationships within *Physalis*. Phylogenetic analysis of DNA sequences from the ITS region of nrDNA and from the nuclear gene *waxy* was used to address the following questions:

1) What are the relationships between the physaloid genera within subtribe Physalinae?
2) How divergent is the Old World *P. alkekengi* from the New World species of *Physalis*?
3) Are the sections of the genus *Physalis*, established primarily on the basis of morphology, generally congruent with DNA data?

METHODS:

Taxon Sampling: *ITS* and *waxy* sequences were obtained from representatives of each of the eight genera in the Physalinae, including the monotypic genera *Oryctes*, *Quincula* and *Margaranthus*, as well as two species of *Chamaesaracha* and two species of *Leucophysalis*. One species of *Tzeltalia* was included as well. Several samples were extracted from herbarium specimens, and DNAs for *Margaranthus*, *Quincula*, *Chamaesaracha* and *Leucophysalis* were kindly provided by R. Olmstead. See Table 1.2 for voucher information.

Thirty-five species of *Physalis* were sequenced, representing all four subgenera, three of which contain only 1-2 species. Seven of the nine sections of subgenus *Rydbergis*, which encompasses the bulk of *Physalis*, were sampled. Material of the monotypic section *Tehuacanae* was not available, and only an ITS sequence was obtained for *P. minimaculata*, one of the two species in section *Rydbergae*. Though most of the Mexican and southwestern United States samples were extracted from herbarium specimens, several species were grown from seed supplied by the Solanaceae Germplasm Collection at Nijmegen, the Netherlands. Samples of southeastern United States native species were extracted from fresh material.

DNA Extraction: Fresh leaves were extracted using a miniprep modification of Doyle and Doyle's CTAB procedure (1987), or using DNeasy Plant Mini kits (Quiagen Inc., Valencia, California, USA). Herbarium material was extracted using the CTAB procedure, and was cleaned using the Elu-quik DNA purification kit (Schleicher and Schuell, Keene, New Hampshire, USA).

General PCR Protocols: PCR protocols were similar for both ITS and *waxy*. PCR reactions were carried out in 25 uL volumes, using Perkin Elmer (Norwalk, Connecticut, USA) AmpliTaq, Mg+ buffer, and dNTPs. The cocktail included: 0.75 uL H2O, 1.25 uL DMSO (ITS only; Buckler et al., 1997), 2.5 uL dNTPs, 4.15 uL Mg+ buffer, 1.25 uL forward primer, 1.25 uL reverse primer, 1.25 uL glycerol and 0.1 uL Taq. This was added to 12.5 uL of diluted DNA sample (1/50-1/100 for high quality DNAs, 2/25 for herbarium DNAs and some *waxy* reactions). The thermocycler program used was: 96° 2min; 30 cycles x 96° 1 min, 50° 1 min, 72° 45 sec; 72° 7 min; 4° hold. PCR products were cleaned using Quiaquick PCR purification kits (Quiagen Inc., Valencia, California, USA), or, in cases where the product was smeary or had multiple bands, using Quiaquick gel cleanup kits (Quiagen Inc., Valencia, California, USA).

Herbarium PCR Protocols: DNA from herbarium material was often fragmented and very limited in quantity. To improve amplification, the general PCR file was modified to: 96° 2 min; 10 cycles x 96° 1 min, 50° 1.5 min, 72° 1 min; 25 cycles x 96° 1 min, 50° 1 min, 72° 45 sec; 72° 7 min; 4° hold. When amplification was weak or undetectable, 1 uL of the 'failed' PCR product was added to new cocktail and re-PCRed, which typically resulted in better amplification. The product from multiple reactions was often pooled to obtain enough for sequencing. Sequences from different individuals of the same species, with one or both obtained from re-PCRing herbarium extractions, clustered together in the analysis, which suggests that any errors resulting from re-PCRing are not enough to significantly affect the phylogenetic signal.

ITS primers: Initially, ITS was amplified and sequenced using ITS 2, 3, 4 and 5 (White et al., 1990). As the project progressed, other primers were included. Samples from fresh material were amplified using Leu1 (L. E. Urbatsch), ITS-5A (K. Wurdak) or ITS-5 and ITS-4 or 4A (external to ITS-4: 5' GGAATCCTTGTAAGTTTC 3'). For DNA extracted from herbarium material, ITS was amplified in two halves, using Leu1 x ITS-2 or 2C and ITS-3 x ITS-4 or 4A . Re-PCRs were done using more internal primers, when possible. Samples were sequenced using mostly internal primers, including those mentioned above and ITS-3i (internal to ITS-3: 5' AATGCGATACTTGGTGTGAA 3'). Two to four sequencing reactions were done for each sample, such that most of the resulting sequence was double stranded. Some taxa could not be directly sequenced, and were cloned using either the Invitrogen TA cloning kit, or the Topo TA kit (Invitrogen, San Diego, California, USA).

Waxy primers: In samples from fresh material, approximately 620 bp of *waxy*, between exons 8 and 10, was amplified using *waxy* 3F by 2R, primers originally designed for the Convolvulaceae (Miller et al., 1999). Amplification of *waxy* was limited to this region to avoid length variation within the introns, and direct sequencing of PCR products was successful for the majority of taxa. For herbarium samples *waxy* was amplified in two halves, using 3F by SR (5' AAAGGTTCAGAYATTCTTGT 3') and 2R by SF (5' AGACTTGARGAGCAGAAAGG 3'). Primers SR and SF were designed for this study, and face each other, so that after sequencing, some overlap joins the two amplified segments.

DNA Sequencing and Alignment: dRhodamine dyes (Applied Biosystems Inc., Foster City, California, USA) was used for cycle sequencing reactions, following the manufacturer's protocols. The resulting products were sequenced on an ABI 377 or an ABI 3700 Automated Sequencer. Sequences were initially corrected and aligned using Sequencher 3.1.1 (Gene Codes Corp., Ann Arbor, Michigan, USA). Further alignment was done by eye.

Missing and composite taxa: The ITS region was successfully amplified in 75 taxa, but *waxy* proved to be more difficult, and was amplified for only 50 of these. About 45 species of *Physalis* were included in the ITS data set, whereas the *waxy* data set had approximately 35.

Leucophysalis nana was the one taxon represented by sequences from two different individuals in the two-gene data set. Neither of the two accessions used would produce both an ITS sequence and an entire *waxy* sequence. The ITS sequence from one individual was combined with the *waxy* sequence from the other. Using the partial sequences from both accessions, *L. nana* was shown to be monophyletic in a preliminary parsimony analysis of the ITS data.

Leucophysalis viscosa was the only taxon included in the two gene data set which had a significant amount of missing data. While represented by an entire ITS sequence, 42% of *waxy* (3F to SR) would not amplify.

Outgroup Selection: Several taxa were examined as potential outgroups, including species of *Capsicum*, *Deprea*, *Iochroma*, *Jaltomata*, *Lycianthes*, *Vassobia*, *Withania* and *Witheringia*. Taxa were initially selected on the basis of cpDNA

phylogenetic analyses, or because their calyx morphology suggested physaloid affinities (D'Arcy and Averett, 1996; Olmstead et al., 1999). DNA from *Capsicum*, *Iochroma*, *Lycianthes*, *Vassobia* and *Witheringia* was kindly provided by L. Bohs and R. Olmstead.

Phylogenetic Analyses: The ITS and *waxy* data sets were analyzed both separately and in combination using PAUP* (Swofford, 1998). Heuristic searches were run with maximum parsimony as the optimization criterion, the amb- function off, and gaps treated as missing data.

The 75-taxon ITS data set was analyzed using a simple addition sequence to produce the starting tree, followed by TBR branch swapping, with no max trees limit. Due to the high number of taxa and relatively low numbers of parsimony informative characters, the bootstrap analysis of this data set was done using the fast heuristic search method, with 10,000 replicates and no branch swapping.

The analysis of the 50-taxon ITS data set was similar to that of the larger data set, but instead of simple addition, the heuristic search used 100 replicates of random addition. To provide bootstrap values comparable to all of the data sets analyzed, two bootstrap analyses were performed, one full heuristic search of 100 replicates with TBR branch swapping, and one fast heuristic search with 10,000 replicates.

Because the *waxy* data set had few parsimony informative characters, full heuristic parsimony searches and bootstrap runs were computationally impossible. Searches were terminated after several thousand most parsimonious trees had been generated, and bootstrapping was done with 10,000 fast bootstrap replicates.

Two methods were used to determine if the ITS and *waxy* data sets should be combined. A 100 replicate partition homogeneity test, with one tree saved per replicate, was performed to help determine if the two data sets were significantly different from one another. The bootstrap consensus trees of the individual data sets were also compared as a means to evaluate topological congruence.

Finally, the combined data set was analyzed using a heuristic search with 100 random addition replicates. Several bootstrap runs were done, two with 100 and 1000 full heuristic replicates, and two with 10,000 and 100,000 fast bootstrap replicates.

RESULTS:

Outgroup Selection: Preliminary analyses suggested that most of the non-Physalinae taxa were too divergent from *Physalis* to be appropriate outgroups. *Witheringia*, the basal-most clade in the cpDNA phylogeny of the Physalinae (Olmstead et al. 1999. See Figure 1.1), proved to be the most suitable choice, and three species were used.

75-Taxon ITS data set: This data set consisted of ITS sequences from 75 taxa. The first 68 base pairs (the end of the 18S gene) were dropped from the analysis, since this was the region of overlap between the three forward primers (Leu 1, ITS 5A, ITS 5) used. The final data set included 744 base pairs of aligned ITS sequences. ITS1 started at base pair 11, and was 269 bp long. The 5.8 S gene was 164 bp long, and ITS2 was 243 bp long, ending at base pair 686. For this data set, 511 characters were constant, 61 were

variable but not parsimony informative, and 172 were parsimony informative. Analysis of these data resulted in 6691 most parsimonious trees, each of 784 steps.

50-Taxon ITS data set: The 50-taxon data set included only those taxa which also had waxy sequences. Of the 744 characters used, 213 were variable, and 146 were parsimony informative. The 220 most parsimonious trees were all 675 steps long.

ITS gene trees: A comparison of one of the most parsimonious gene trees from the 75-taxon data set (Figure 1.2) to one from the 50-taxon data set (Figure 1.3), revealed similar structures. Both trees have a strongly supported branch separating the outgroup, *Witheringia*, and the closely-related *Brachistus*, from the remainder of the Physalinae. *Tzeltalia* and *Leucophysalis viscosa* hold basal positions in both trees, followed by a grade of morphologically atypical *Physalis* species intermixed with physaloid genera. A major feature of both ITS gene trees is the strongly supported, monophyletic clade made up of all of the morphologically typical species of *Physalis*, and the notable exclusion of *P. alkekengi*, the type of the genus, from this clade. In both trees, *P. alkekengi* (China) is strongly supported as being sister to *P. carpenteri* (southeastern United States).

The morphologically typical *Physalis* clade is generally congruent with subgenus *Rydbergis*. The three other subgenera of *Physalis* (*Physalis*, *Physalodendron*, and *Quincula*) did not group with the *Rydbergis* clade, but were resolved as a grade of physaloid taxa and morphologically unusual *Physalis* species at its base. For both analyses, support for species groups within the *Rydbergis* clade was weak, but it is apparent that most of the sections of subgenus *Rydbergis* are probably not monophyletic.

With the exception of *Margaranthus*, all of the physaloid genera were distinct both from the main *Physalis* clade and from each other. *Margaranthus* nested within the *Rydbergis* clade. *Physalis* subgenus *Physalodendron* (*P. arborescens* and *P. melanocystis*) and *Tzeltalia* were strongly supported as monophyletic in the 75-taxon tree (Figure 1.2). In addition, both trees showed that *Chamaesaracha* was monophyletic, and *Leucophysalis* was paraphyletic.

Waxy gene trees: The *waxy* data set had 612 characters, but only 66 were parsimony informative. An initial heuristic search was terminated early after generating thousands of 249 step trees. A set of 112 most parsimonious trees was generated by conducting 1000 replicates of random addition and saving one tree per replicate if it was 249 steps or shorter. Branch swapping on these trees was stopped after 42,647 trees had been saved. A second search was done using simple addition, and was terminated after 85,791 trees had been found. The same consensus tree resulted from either set of trees.

The strict consensus of the *waxy* gene trees was similar to, but less resolved than, that of the 50-taxon ITS gene tree. In one of the most parsimonious *waxy* trees (Figure 1.4), the main, monophyletic clade of *Physalis* species was resolved, but with less support (63%) then in either ITS analysis. Though the species relationships within the main *Physalis* clade are not well-resolved, there is strong support for one monophyletic group of species, this being the stellate-haired perennial taxa, most of which belong to *Physalis* section *Viscosae*. As in the ITS trees, *Chamaesaracha* was monophyletic and *Leucophysalis* was paraphyletic. Unlike the ITS tree, *Oryctes* is strongly supported as being sister to *Leucophysalis*, and the sister-species relationship between *P. carpenteri*

and *P. alkekengi* was not detected. Though the *L. viscosa waxy* sequence was incomplete, analyses of the individual ITS and *waxy* data sets both placed *L. viscosa* at the base of the tree, suggesting that the lack of 30 parsimony informative characters (from a total of 66 for *waxy*) did not have a profound phylogenetic effect.

Clade Support: Fast bootstrapping was the method most suitable for use with the 75-taxon ITS data set and the *waxy* data set because the small number of parsimony informative characters for a relatively large number of taxa made it difficult to complete a full heuristic search. Prior to using this method, its properties were examined more closely. First, several 10,000 replicate fast bootstrap runs for the data set were compared to see how much variation in values was seen between runs. Generally, there was 1-2% of variation per node, though one case of 6% was seen. Finally, a 10,000 replicate fast bootstrap run was compared to a 100,000 replicate run to see how much variation resulted from increased replication. The results were similar to what was seen with the earlier comparisons. Thus it was decided that 10,000 fast bootstrap replicates were adequate for providing reliable bootstrap values.

Both the 50-taxon ITS data set and the combined data set had enough characters to complete full heuristic searches, raising the question of how bootstrap values from fast and full heuristic analyses compare. Bootstrap values from a 100 replicate full heuristic search were very similar to those from a 1000 replicate search, suggesting that 100 replicates is adequate for providing reliable bootstrap values. However, when these values were compared to those from 10,000 and 100,000 replicate fast bootstrap searches, fast bootstrap searches generally showed much less support for nodes that appeared to be

well-supported in the full heuristic searches. Thus the fast bootstrapping method seems to be more conservative that the full heuristic method. This agrees with the findings of DeBry and Olmstead (2000), who examined the effects of using less rigorous tree search methods to save time during bootstrap analysis.

Comparison of the different bootstrapping methods between the 50-taxon ITS data set and the combined 50-taxon data set showed that 100 replicate full heuristic bootstrap runs resulted in more resolved consensus trees than did the fast bootstrap runs. For the heuristic analyses, the 50-taxon ITS tree had 24 nodes with greater than 50% bootstrap support, and the tree from the combined data set had 28, as opposed to 16 for ITS and 17 for the combined data set in the fast bootstrap consensus trees. The most noticeable difference between heuristic and fast bootstrap analyses was that in the heuristic bootstrap analysis of the combined data set, *Oryctes* was weakly supported as being sister to *Leucophysalis* (Figure 1.5), a result not seen in the fast bootstrap analysis.

Partition Homogeneity Test: An ILD (incongruence length difference) test produced a set of trees which were all longer than the sum of tree lengths from the original partition, indicating that the two data sets were significantly different (p=0.01). However, this test is known to be conservative (Yoder et al., 2001), and the bootstrap consensus trees from the two data sets had no strongly supported areas of incongruence. Thus, the data sets were combined for the final analysis.

Combined analysis: The final, combined data set consisted of 212 parsimony informative characters for 50 taxa. This data set produced 170 most parsimonious trees of 966 steps.

Examining one of the most parsimonious trees from the two-gene analysis (Figure 1.5) indicated that it is more similar in structure to the gene tree from the 50-taxon ITS data set than it is to the *waxy* gene tree, though bootstrap support, especially along the backbone of the tree, is slightly higher than in either individual analysis. *Brachistus* and *Witheringia* were strongly supported as being monophyletic (100%), and together with *L. viscosa*, formed the basal-most members of the Physalinae. *Tzeltalia* was the next taxon to branch in this grade, and was well supported as being both separated from the basal taxa (86%) and from the tip clades (93%). *Physalis alkekengi* and *P. carpenteri* were still strongly supported as sister taxa (93%), branching after *Tzeltalia*, along with *P. microphysa*, and a weakly supported (53%) clade containing *Oryctes* and the North American species of *Leucophysalis*. There was no bootstrap support for the backbone of the tree between these three clades. There was weak support (59%) for *P. arborescens* being the next taxon to diverge, and for *Quincula* (54%) or *Chamaesaracha* to diverge after that. One of the most strongly supported (100%) branches in the tree separated the physaloid taxa from the morphologically typical species of *Physalis*. There was weak support (64%) for *P. crassifolia* and *P. acutifolia* being the basal-most members of the *Physalis* clade, as well as for *Margaranthus*, *P. philadelphica* and *P. microcarpa* being the next branches to diverge. There was little resolution among the majority of *Physalis*

species, except for two well supported clades, one including 11 U.S. perennial species (86%), the other a group of stellate-haired species (81%).

DISCUSSION:

Overview of *Physalis*: As currently circumscribed, *Physalis* is a paraphyletic genus. Though this had been suggested by earlier cpDNA work (Mione et al., 1994; Olmstead et al., 1999), the more extensive sampling used in this study clarified this point: the highly inflated fruiting calyx, considered so definitive of *Physalis*, has arisen multiple times throughout the Physalinae. The morphologically typical, New World members of *Physalis* form a strongly supported monophyletic group, but *P. alkekengi*, the type of the genus, is not a member of this clade. *Physalis alkekengi* is highly divergent from the bulk of *Physalis* species, and is separated from them by a grade of physaloid genera and morphologically unusual species of *Physalis*. While support for the order of taxa within this grade is weak, there is strong support for separating these taxa from the four, basal-most taxa in the Physalinae (*Brachistus*, *Leucophysalis viscosa*, *Tzeltalia* and *Witheringia*).

Thirty of the 35 *Physalis* species sampled formed a clade at the tip of the Physalinae clade. This clade generally corresponds to *Physalis* subgenus *Rydbergis*, and is morphologically homogenous, consisting of mostly herbaceous plants, with solitary flowers, unlobed yellow corollas and highly inflated fruiting calyces. Included in this group is *Margaranthus solanaceous*, supporting its current inclusion in *Physalis* (Martinez, 1993; Axelius, 1995). A wide variety of morphological data also supports the

placement of *Margaranthus* within *Physalis*. The urceolate flowers of *Margaranthus*, though unique within *Physalis*, resemble partially open *Physalis* flowers, with their yellow base color largely obscured by the large, dark basal spots which extend most of the length of the corollas. Indeed, the corolla vasculature is quite similar to that of *Physalis* (Averett, 1979). In fruit, *Margaranthus* resembles a typical annual *Physalis*, and its chromosome number (x=12) and morphology are similar to those of the annual species surveyed by Menzel (1950).

The *Physalis* species which did not fall within the *Rydbergis* clade are all species which are morphologically atypical, either because they have clustered flowers, lobed corollas, or unusual fruiting calyx morphology. Included among these 'atypical' species are the three remaining subgenera of *Physalis* which were recognized by Martinez (1999): *Physalis* (*P. alkekengi*), *Physalodendron* (*P. arborescens* and *P. melanocystis*) and *Quincula* (*P. lobata*).

Relationships among the basal taxa of the Physalinae: *Brachistus* is a small genus of three species, and has been considered a section of *Witheringia* (D'Arcy et al., 1981). Morphologically, *B. stramonifolium* is very similar to *Witheringia*, the main difference being the slight expansion of the fruiting calyces, which causes them to gently clasp the sides of the maturing berries. The 2-gene phylogeny supports *Brachistus* and *Witheringia* as sister taxa, though the 75-taxon ITS phylogeny places the two accessions of *B. stramonifolium* in different positions. However, the genus *Brachistus* is poorly known, and it is quite possible that there are more than the three currently recognized species. *Brachistus* was segregated from *Witheringia* partly on the basis of its comparatively

primitive calyx morphology (D'Arcy, 1986), and its recognition as a distinct genus warrants further study.

Tzeltalia is a recent segregate of *Physalis*, and consists of the two species from Guatemala and southern Mexico, which were formerly placed in *Physalis* section *Calidariae* (*P. calidara* and *P. amphitricha*) (Estrada and Martinez, 1998). ITS data support both the monophyly of this small genus and its recognition as distinct from *Physalis*.

The taxon currently called *Leucophysalis viscosa* has been placed in no less than six genera: *Athenaea*, *Chamaesaracha*, *Jaltomata*, *Physalis*, *Saracha*, and *Witheringia*. Its placement at the base of the Physalinae clade, as opposed to sister to the two North American species of *Leucophysalis*, indicates that this species should be removed from *Leucophysalis*. Where it should be placed next remains uncertain. Preliminary analyses using more distant outgroups than *Witheringia* (not shown) also placed *L. viscosa* at the base of the Physalinae, suggesting that *L. viscosa* does belong in the subtribe. An earlier cpDNA study showed that it did not belong in *Saracha* or *Jaltomata* (Mione et al., 1994). Though further sampling of the basal Physalinae may clarify the affinities of *L. viscosa*, providing a new genus for this distinctive taxon might be the best way to deal with the nomenclatural dilemma.

The "physaloid grade": Four physaloid genera and four morphologically atypical species of *Physalis* form a grade at the base of the *Rydbergis* clade. These taxa encompass most of the morphological variation within the Physalinae. The trait they all share is some amount of calyx expansion in fruit, and most of them have clustered

flowers. Support for the backbone of this grade is weak, but morphology provides some additional information about the relationships between these taxa.

Physalis alkekengi and *P. carpenteri* form one of the most strongly supported pairs of sister taxa in the physaloid grade, as well as one of the most interesting. They are morphologically dissimilar: *P. alkekengi* is a rhizomatous, perennial herb with 5-lobed, white corollas and red-orange fruit and fruiting calyces, while *P. carpenteri* is a woody, taprooted perennial, with unlobed, yellow corollas and brown fruiting calyces around yellow berries. An even more striking difference is the disjunction in their ranges. *Physalis alkekengi* is native to China, while *P. carpenteri* is a rare species from the southeastern United States. Sequence data from the chloroplast gene ndhf also support this relationship (Bohs, unpublished). The floristic affinities between eastern North America and eastern Asia are firmly established (Graham, 1999), and these taxa may exemplify this pattern. *Physalis alkekengi* is one of the most temperate species within the Physalinae, and it is not implausible to hypothesize that one of its ancestors ranged broadly throughout the Arcto-Tertiary geoflora (Graham, 1999). Fossil evidence does demonstrate that members of the subfamily Solanoideae once ranged much further north than they presently do. Fossil seeds resembling those of *Solanum* or *Physalis* have been identified from the Mary Sachs Gravel, an 18 million year old deposit from Banks Island, in northern Canada (Matthews and Ovenden, 1990).

The placement of *Leucophysalis* in this analysis is congruent with the placement of this taxon in earlier cpDNA studies (Olmstead et al., 1999; Mione et al., 1994), though the sequence of relationships is not entirely clear. What is clear is that *L. nana* and *L.*

grandiflora are sister taxa, which is also supported by the very similar morphology and flavonoid chemistry of the two species (Averett, 1979). Like *P. alkekengi* and *P. carpenteri*, though certainly less extreme, these taxa have a broad range disjunction. *L. nana* is a plant of the Sierra Nevada Range (Averett, 1979), while *L. grandiflora* is one of the few north temperate members of the Physalinae, occurring in the Great Lakes region of the United States and Canada. This is additional evidence for widespread ancestral physaloids at higher latitudes.

Oryctes nevadensis has been recognized as a physaloid genus since 1896 when Rydberg treated it in his monograph of *Physalis* and related genera. The molecular data strongly support the placement of this taxon within the Physalinae, a result which had been suggested by cpDNA data (Olmstead et al., 1999), but its exact affinities remain unclear. The derived physaloids, including *P. arborescens*, *Chamaesaracha*, *Quincula*, and members of the *Rydbergis* clade, all share a 19 bp deletion in ITS1, which is also found in *Oryctes*, arguing for its placement among these taxa. *Oryctes* is very rare and poorly known, and our current knowledge of its morphology offers little information about its affinities. Its tubular, purple corollas are unique within the subtribe, but its flavonoids are typical of much of the Solaneae, as well as *Leucophysalis* and *Chamaesaracha* (Averett and D'Arcy, 1983). Its unique seed testa pattern argues against a close affinity to *Physalis* or *Chamaesaracha*, but determining which, if any, of these taxa share this character awaits further information about seed testa characters in other physaloids (Axelius, 1992).

P. microphysa is one of the more unusual of the 'anomalous' species of *Physalis*. While most of the morphologically odd species are notable for their clustered flowers and/or lobed corollas, *P. microphysa* has typical, *Physalis*-like, solitary yellow flowers, but very unusual fruiting calyces. The calyces only enlarge to about 1/2" long, are rather deeply lobed, and never close at the apices. Rydberg (1896) suggested removal of this species from *Physalis*, and Martinez (1999) agreed, declining to treat this species in her revision of the genus, and suggesting that it and its putative sister species, *P. parvianthera*, be placed in a new genus, *Cascada*. The placement of this species in the physaloid grade, and its lack of association with any of the other taxa in this grade, supports its placement in a new genus.

The most 'strongly' supported node (56% bootstrap) along the backbone of the physaloid grade separates the basal members of the grade from *P. arborescens*, *Quincula*, and *Chamaesaracha*. Though ITS data supports *P. arborescens* and *P. melanocystis* as sister species, and thus supports the monophyly of *Physalis* subgenus *Physalodendron*, too little is known about the morphology of these species to discuss their affinities within the physaloid grade.

The node separating *P. arborescens* from *Quincula* and *Chamaesaracha* is weakly supported, but several morphological characters affirm the close relationship of these two taxa. Pinnatifid leaves occur in both *Chamaesaracha* and *Quincula*, a trait unique within the Physalinae. The two taxa also share fruit with basal placentation and similar corolla vasculature (Averett, 1979). A trait unique to *Quincula*, and one often used to argue for the recognition of this genus, is its unique base chromosome number.

Though X=12 describes the other members of the Physalinae and the majority of the Solanoideae, *Quincula* has a base number of 11. While each taxon has similarities to *Physalis*, it is unclear whether *Quincula* or *Chamaesaracha* is sister to the *Rydbergis* clade. The inflated fruiting calyces of *Quincula* look much like those of *Physalis*, while the pale, spotted flowers of *Chamaesaracha* are more *Physalis*-like than those of *Quincula*.

Species relationships within subgenus *Rydbergis*: Morphological and geographical characters have been the main criteria for establishing sections within *Physalis* (Martinez, 1999; Menzel, 1951). Four of the 9 sections of subgenus *Rydbergis* are small, with only 1-2 species. With the exception of *P. carpenteri*, the species of subgenus *Rydbergis* form a monophyletic group, but relationships within this group are poorly resolved.

ITS data place *P. minimaculata*, one of two species in section *Rydbergae*, solidly within the main clade of *Physalis* species. Neither the placement nor the branch length of this species justifies separation from the larger sections of the genus. Martinez (1999) noted that there are several as yet unnamed species from Mexico which should belong to this section, and further molecular work with these species might help determine whether this section should be recognized.

The two species of section *Campanulae*, *P. campanulata* and *P. glutinosa*, are unique because of their very large flowers. Aside from that, they are rather distinct from one another, and the ITS sequence data provide no support for a sister taxon relationship,

suggesting that section *Campanulae* may be an artificial grouping. These species might better be recognized merely as distinctive members of one (or two) of the larger sections.

The remaining five sections of subgenus *Rydbergis* contain from 6-14 species, and are based, in large part, on the gross morphological characters traditionally used in *Physalis* taxonomy. Sections *Angulatae* and *Epeteiorhiza* contain mostly annual species, and are distinguished from one another largely on the basis of calyx angles. In fruit, members of section *Angulatae* have rounded or somewhat 10-angled calyces, while most members of section *Epeteiorhiza* have five. The type species of the two sections, *P. angulata* and *P. pubescens*, formed one of the most strongly supported pairs of sister taxa in the *Rydbergis* clade. Both species are weedy annuals, but are easily distinguished by several distinctive morphological characters. That they may be sister species is not particularly surprising, but their identical *waxy* and ITS sequences (results which hold after sampling multiple individuals) are an interesting example of morphological divergence without sequence differentiation.

A less surprising result among the annual taxa was that *P. minima,* a member of section *Angulatae*, and *P. lagascae*, currently treated as its synonym, did not group as sister taxa. However, section *Angulatae* has not been revised recently, and these two taxa may be distinct. The other relationships between the species of sections *Angulatae* and *Epeteiorhiza* are too weakly supported to determine how they might best be partitioned into monophyletic groups, though the two sections are distinctly paraphyletic, even when not considering the types.

The majority of the rhizomatous, perennial species of *Physalis* are placed in two sections: the ones with branched hairs in section *Viscosae*, and the ones with simple hairs in section *Lanceolatae*. Section *Viscosae* is monophyletic, and all but two of the United States species of section *Lanceolatae* form a monophyletic group sister to the *Viscosae*. *Physalis longifolia*, *P. hederifolia* and the non-U.S. species of section *Lanceolatae* are scattered throughout the rest of the *Rydbergis* clade. It is interesting to note that most species of section *Lanceolatae* s.s. occur in the southeastern United States, whereas *P. hederifolia* is endemic to the southwestern United States and *P. longifolia* ranges from the southeastern United States to northern Mexico. *Physalis longifolia* is morphologically distinct from the other members of section *Lanceolatae* and may be a species of southwestern United States or Mexican origin whose range has expanded northward.

The *Lanceolatae/Viscosae* clade represents about half of the species native to the United States, and is a striking example of a northern radiation into temperate habitats. *Physalis viscosa* is the only species in the United States perennial clade which is not common in the United States, and it is a wide ranging coastal species, occurring from Baja California down to the coasts of northern South America. Its current range might be explained by dispersal. Since *P. lanceolata*, a southeastern United States endemic species, typifies section *Lanceolatae*, the name should be applied to the United States members of the section, and the species which do not group with that clade should be assigned to other sections.

Section *Coztomatae* was erected to contain 11 distinctive Mexican species, characterized by dark corolla maculations formed from conglomerations of smaller spots.

Though represented by only three taxa in the two gene data set, there is no support for a monophyletic *Coztomatae*. However, resolution is poor, and neither species is strongly supported as being sister to taxa from another section. Though the sampling is slightly better, the results are similar in the 75-taxon ITS tree. Compound corolla maculations appear on and off throughout the Physalinae, and are either symplesiomorphic for the group or are easily evolved from other types of corolla spotting. It is unlikely that the *Coztomatae* is a monophyletic group, though further sampling may reveal some monophyletic subsets of species.

Morphological trends within the Physalinae: There are several morphological trends within the Physalinae. In general, there is a shift from a woody habit in the basal physaloids to a herbaceous habit in more derived taxa. Flower shape moves from a symplesiomorphic stellate form common throughout the Solanaceae to an unlobed form more typical of *Physalis* and its close relatives. Calyx shape and enlargement is more variable than flower form, but shape stabilizes to enveloping the fruit and amount of expansion generally increases as one moves up from the basal physaloids.

The basal four genera of the Physalinae (*Witheringia*, *Brachistus*, *L. viscosa* and *Tzeltalia*) are mostly woody perennials, with shrubby or sprawling habits. They have densely clustered flowers with deeply lobed corollas. Corolla color varies from white to greenish or yellowish. The amount of calyx expansion ranges from none in many species of *Witheringia*, to inflated and *Physalis*-like in *Tzeltalia*.

In comparison, the taxa in the physaloid grade are more morphologically variable. Though many of them are woody, herbaceous taxa form the bulk of the diversity. Most of

these taxa still have clustered flowers, but the clusters are often small, with only 2-4 flowers. Flower color varies from white to purple to yellow. The feature that all of these taxa share is calyx expansion in fruit, though some, like *Oryctes*, have deeply lobed, tightly clasping calyces which only partially cover the berry, while others, such as *Quincula*, have fully inflated, *Physalis*-like calyces.

The *Rydbergis* clade is morphologically homogenous, characterized by solitary, yellow flowers, with unlobed corollas, and inflated calyces in fruit. Most of the major variation occurs near the base of this clade. *P. crassifolia* is one of the few truly woody species, while its sister *P. acutifolia* is the one species with white flowers. *P. solanaceous* (=*Margaranthus*), also one of the basal-most branches in the clade, is the only species with urceolate flowers. Though all of the species in this clade have a very large amount of calyx expansion in fruit, there are a few species, such as the cultivated tomatillo, which have unusually large fruit that nearly fill the calyces, thus obscuring the calyx inflation.

Biogeography of the Physalinae: The Solanaceae, a family of 92 genera and 3500 species, is most diverse in South and Central America, though several genera co-occur or are endemic to the Old World (Hunziker, 1979; Gentry and D'Arcy, 1986). A Gondwanan origin has been hypothesized to explain this distribution (Hunziker, 1979; D'Arcy, 1991). However, Olmstead and Palmer (1992) found that the Cestreae, the basal-most lineage in the Solanaceae, is restricted to South America, and argued that a South American origin with subsequent dispersal is a better explanation. Additionally, a combination of fossil evidence and molecular clock methods have been used to date the divergence of the Solanales from the Boraginales to about 50 mya (Magallon et al.,

1999), which supports Olmstead and Palmer's hypothesis, since it is estimated that by 85 mya, Gondwana had split, and up to 800 km of ocean separated South America from Africa (Burnham and Graham, 1999). Dating the actual emergence of the Solanaceae is complicated by the fact that fossil Solanaceae are apparently uncommon or difficult to identify.

The earliest diverging lineages in the Physalinae are *Witheringia* and *Brachistus*, both of which are predominantly Central American, though *Witheringia* is almost as diverse in northwestern South America. A grade of physaloid taxa, nearly encompassing the geographical extent of the subtribe, falls between these genera and the *Rydbergis* clade. These physaloid taxa are predominantly North American, though they range from Mexico to the southeastern United States and the Great Lakes region of Canada. The majority are from the southwestern United States and northern Mexico, and *P. microphysa*, *Quincula*, and *Chamaesaracha* are all from the Chihuahuan and Sonoran desert regions. The basal-most branch of the *Rydbergis* clade also appears to be of Chihuahuan/ Sonoran origin. This pattern suggests that physaloid ancestors may have migrated north from Central America, spread widely over North America and Europe during the milder periods of the Tertiary, and receded during the ice ages to the deserts of the southwestern United States and northern Mexico, where they diversified into several small genera, one of which was *Physalis*. The early *Physalis* species then underwent an adaptive radiation, moving southward back into Mexico, diversifying as they encountered the wider variety of habitats available to the south of the deserts.

Avenues for future study: There are many aspects of the Physalinae that merit further study. D'Arcy (personal communication) suggested that there are at least two species under the auspices of *P. alkekengi* , and indeed, multiple *P. alkekengi*-like species have been named, though they have often been treated as horticultural variants, and not as distinct species. *Archiphysalis* and *Physalisatrum* are two poorly-known Asian genera whose species have occasionally been included within *Physalis* and *Leucophysalis*, and are currently considered to belong to the Withaninae. However, this determination is based solely on morphological data, and it would be interesting to see if molecular data place these species closer to *Withania* or *P. alkekengi*.

CONCLUSIONS:

The genus *Physalis* is highly paraphyletic, with the morphologically typical species forming a very well supported monophyletic group, and the morphologically atypical species, including the type, intermixing with the physaloid genera in a grade at the base of this clade. *Margaranthus* was the only physaloid genus to fall into the main *Physalis* clade, and its transfer to *Physalis* seems appropriate. The genera *Chamaesaracha*, *Leucophysalis*, *Quincula* and *Tzeltalia* were all well separated from the main *Physalis* clade, leaving little doubt that they are distinct.

Though morphological characters have generally proved to be reliable in delimiting monophyletic physaloid genera, they are not particularly useful for delimiting monophyletic species groups within *Physalis*. Most of the sections of *Physalis* appear to be paraphyletic, though the species relationships within the *Rydbergis* clade were for the

most part poorly supported. However, the monophyly of section *Viscosae* was well supported, and it proved to be sister to a monophyletic subset of species from section *Lanceolatae*. This *Lanceolatae/Viscosae* group is made up of mostly United States species, and represents a northward radiation from Mexican stock.

Given the other options (renaming the 75+ species of New World *Physalis*, or subsuming the majority of the Physalinae into one genus), the least taxonomically disruptive approach for dealing with the paraphyly of the genus would be to re-typify *Physalis* using a Linnaean species which is a member of the monophyletic, morphologically typical clade, such as *P. pubescens*. The atypical species could be recognized as small genera. This strategy would result in four new genera (for *P. carpenteri*, *P. alkekengi*, *P. microphysa* and subgenus *Physalodendron*), and a morphologically homogenous *Physalis*.

Chapter 2.

Interspecific ITS sequence variation in the U.S. perennial *Physalis* (Solanaceae)

INTRODUCTION:

The internal transcribed spacer region of the nuclear ribosomal DNA (ITS) is one of the most widely used tools for phylogenetic reconstruction, particularly at or below the genus level (Baldwin, 1995). Several factors contribute to its high utility. Currently, it is one of the most variable regions available for inferring phylogeny among closely related species. The two spacers are located between the 18S, 5.8S and 26S rRNA genes, providing highly conserved primer sites. The multicopy nature of the rDNA makes PCR amplification easy, and its relatively short length allows for direct sequencing of the ITS1, 5.8S, ITS2 array in its entirety.

Though the multicopy nature of the rDNA repeats could potentially be problematic when using ITS for phylogenetic reconstructions, concerted evolution is expected to homogenize copies within individual plants (Zimmer et al., 1980). Many studies involving allopolyploid taxa (*Paeonia*, *Amelanchier*, *Silene*) have shown that the homogenizing force of concerted evolution is not as universally effective as might be hoped (Sang et al., 1995; Campbell et al., 1997; Popp and Oxelman, 2001), but the atypical genetic makeup of such taxa may be reassuring to workers studying diploids and non-hybrid species. However, a growing number of studies has demonstrated that even these taxa are not immune to problems. Putative pseudogenes have been found in a wide

array of diploid taxa, such as *Lophocereus, Aeschynanthus* and *Quercus* (Hartmann et al., 2001; Denduangboripant and Cronk, 2000; Muir et al., 2001), and multiple functional copies have been isolated from *Nicotiana*, as well (Buckler and Ippolito, 1997). Despite these potential problems, phylogenetic hypotheses based exclusively on ITS are commonly reported, and often used to guide taxonomic revisions, biogeographical surveys, and population genetic studies.

Unlike concerted evolution, lineage sorting is a purifying force expected to function at the species level, eventually producing monophyletic species through the gradual loss of alleles (Avise, 1994). Though many studies have focused on concerted evolution, fewer have focused on the issue of lineage sorting and sample size, since surveying multiple individuals is costly, time consuming and potentially uninteresting under the assumption of low within-species variation. Also, for workers intent on generating an exemplar-based phylogeny, both monophyletic and paraphyletic species would result in similar sister-taxon relationships.

Structural constraints represent another problematic issue rarely considered in phylogenetic studies using ITS sequences. Because the RNA transcribed from the ITS region is neither long-lived nor protein-coding, it is often thought of as non-functional, thus freeing the transcribed spacers to evolve at potentially high rates. The fact that both spacers are highly variable and prone to indels generally supports this view. However, both ITS1 and ITS2 play a role in the processing of ribosomal RNA (Baldwin et al., 1995; Lalev and Nazar, 1999), most of which is dependant not upon the primary structure of their sequences, but on the secondary structures, or three dimensional shapes, formed

by their transcribed RNAs. Experimental studies in yeasts have shown that processing of ribosomal RNAs can be blocked by changes to the secondary structure of either transcribed spacer (Baldwin et al., 1995). It is known that both ITS1 and ITS2 have conserved motifs, found throughout the angiosperms (and even in some green algae), which emphasizes that these regions are not free of constraints (Liu and Schardl, 1994; Hershkovitz and Zimmer, 1996; Mai and Coleman, 1996). Such constraints on sequence evolution could have profound phylogenetic effects. Compensatory changes to maintain double-stranded regions would violate the assumption that characters change independently (Wheeler and Honeycutt, 1988; Hillis and Dixon, 1991), and would act as an internal weighting scheme on such characters. Regions which converge due to structural constraints could also be a source of homoplasy in an analysis.

An excellent study group for examining some of the assumptions made about ITS and its phylogenetic utility are the U.S. perennial *Physalis* species. The approximately 13 species in the group are diploid (2n=24) (Menzel, 1951; Hinton, 1976), thus avoiding the potential complications of multiple genomes, and their biology is relatively well known. The taxonomy is stable, and section *Viscosae* has been recently revised (Sullivan, 1985). Also, this complex of closely related species is typical of the type of group to which ITS studies are often applied.

A recent phylogeny of *Physalis* has shown that 11 of the U.S. perennial *Physalis* species, in two sections of the genus, *Viscosae* and *Lanceolatae*, form a monophyletic group (Whitson, Chapter 1). A subgroup of these species, the majority from section *Viscosae*, also forms a monophyletic group within the perennial clade, and shares a

common morphology, having narrow, running rhizomes and highly branched trichomes. The remaining species, all from section *Lanceolatae*, are generally distinguished by their stout, non-running rhizomes and unbranched trichomes.

These species have interesting reproductive biology, the majority being obligate outcrossers through self-incompatibility (Sullivan, 1984b; Menzel, 1951; Hinton, 1975, 1976). Crossing studies have shown that species within sections may hybridize, commonly resulting in reduced fruit set, seed set, and seed viability, though viable seeds can be obtained. Crosses between species of different sections are typically much less successful than within-section crosses, and seeds are rarely produced. Examples of hybridization in the wild are often based on anecdotal evidence, but one case from section *Lanceolatae* has been documented (Hinton, 1975), and morphological evidence strongly suggests another in section *Viscosae* (Sullivan, 1985).

A small data set from a section of the nuclear gene *waxy* was used to provide an independent phylogenetic estimate of the relationships among the southeastern perennial *Physalis*. *Waxy* is a protein-coding gene, and thus less variable than ITS. However, it has several advantages over ITS. Indels are rare and alignment is straightforward, as a single copy or low copy number gene, concerted evolution or lack thereof is not really an issue, and variation within species is low. The one drawback is that while the level of variation in *waxy* is useful for clarifying sectional level relationships within this species complex, there is not enough resolution to determine species relationships within the sections. However, the crossing studies done between many of these taxa provide some information about which pairs of species may be most closely related.

In addition, the *waxy* data set, a large ITS data set was generated by sampling from 13 species in sections *Viscosae* and *Lanceolatae*, and from multiple individuals of four species: two widespread species, one with an approximately three state range, and one narrow endemic. Eleven artificial hybrids were also sampled. These data were used to address several questions:

1) Will increased sampling within species reveal para- or polyphyletic groups of ITS sequences?
2) How will new combinations of characters, as seen in hybrids, effect the ITS gene tree?
3) How well does secondary structure of the ITS region help explain the patterns of character variation seen within this species complex?

METHODS:

DNA extraction and PCR amplification: DNA extractions were done using DNeasy Plant Mini kits (Quiagen Inc., Valencia, California, USA) for fresh material, and using a miniprep modification of Doyle and Doyle's (1987) CTAB method for herbarium material. CTAB extractions were cleaned using Elu-quik DNA purification kits (Schleicher and Schuell, Keene, New Hampshire, USA). PCR amplifications and sequencing of *waxy* and ITS followed protocols described in Chapter 1. Most individuals were direct sequenced. Individuals which sequenced poorly were not used in the analysis, though a few of these were cloned to determine if the problem was due to the presence of pseudogenes or to multiple functional ITS copies. Cloning was done using the Topo TA

cloning kit (Invitrogen, San Diego, California, USA) and following manufacturer's instructions.

Taxon sampling: ITS was sampled from five species of section *Lanceolatae*: *P. heterophylla*, *P. lanceolata*, *P. virginiana*, and two new species, a narrow endemic from Florida which is closely related to *P. heterophylla* (designated Species A in the figures), and a species morphologically most similar to *P. virginiana*, from Texas, Arkansas and Louisiana (designated Species B in the figures). Eight species falling into the stellate-haired group were sampled: five belonging to section *Viscosae*, including *P. viscosa* (which is the only non-native species sampled – it ranges from Mexico to South America), *P. cinerascens*, *P. walteri*, *P. angustifolia* and *P. mollis*, and three species which commonly have branched trichomes, but have not formally been placed in section *Viscosae*: *P. arenicola*, *P. hispida* and *P. pumila*. Ten or more individuals were sampled from each of four species: *P. heterophylla*, *P. virginiana*, Species A and *P. arenicola*. ITS sequences from eleven hybrid plants, including ten from crosses between *P. heterophylla* and Species A and one from a *P. heterophylla* by *P. arenicola* cross, were also sampled. Three annual species, *P. angulata*, *P. pubescens* and *P. cordata*, were used as outgroups. The *waxy* data set included samples from all of the species above, with the exception of *P. hispida* and the hybrid plants.

Alignment, character coding, and data partitioning: Sequences were aligned by eye in Sequencher 3.1.1 (Gene Codes Corporation, Ann Arbor, Michigan). Secondary structure analysis of the ITS1 and ITS2 regions was done using the program RNAfold (Zuker et al., 1999) and was used to more accurately align indel regions. The borders of

ITS1, ITS2 and the 5.8 S region were delimited using a tomato sequence (GenBank #X52265: from Kiss et al., 1988) as a guide. Sequences from hybrids were coded as polymorphic at variable sites, and indel regions were coded as missing, since the chromatograms could not be reliably read or aligned in those regions.

Detection of contaminants: Two types of ITS sequences were sampled from *P. heterophylla*, and to ensure that only ITS sequences from *Physalis* DNA had been amplified, BLAST searches were run on both ITS types.

Identification of pseudogenes: Several methods were used to identify sequences from potential pseudogenes (Mayol and Rosello, 2001; Muir et al., 2001). The 5.8S regions of all sequences were examined, and sequences with substitutions were removed from the analysis. Then, the conserved regions from ITS1 (Liu and Schardl, 1994) and ITS2 (Hershkovitz and Zimmer, 1996; Mai and Coleman, 1996) were examined for substitutions and compared to other sequences to determine whether the same variable site was found in many individuals, or was unique to a single sample. These sites were also mapped onto a hypothesized RNA secondary structure to determine if they fell into stem or loop regions. Sequences which had multiple unique substitutions in the conserved regions, and whose substitutions occurred in stem regions, were removed from the analysis. Sequences which had unusually long branches in the phylogenetic analysis were examined again to determine the source of the variation, and the changes were mapped onto the RNA secondary structure to determine whether they fell into stem or loop regions. Sequences with multiple unique and non-compensatory mutations in stem regions were removed from the analysis.

Examination of chromatograms: Chromatograms of all taxa were carefully examined for evidence that the individual was polymorphic for ITS sequences with different indel types. Individuals which were polymorphic for such sequences were not used in the analysis. However, PCR products from three individuals of *P. heterophylla*, one with a common ITS type, one with a rare ITS type, and one that appeared to have both, were cloned to determine which ITS types could be found within an individual. Ten of the resulting cloned sequences were included in subsequent phylogenetic analyses. Individuals that had some polymorphic sites but that did not appear to have length variation between their different ITS copies were coded as polymorphic at those sites and left in the analysis on the assumption that similar sequences from the same individual would group together.

Phylogenetic analysis: All analyses were done using PAUP* version 4.0b8 (Swofford, 1998). Maximum parsimony was used as the optimality criterion and gaps were treated as missing data, though alignable regions from indel zones were included. Initial parsimony searches were done for each analysis to gauge the length of the shortest trees, then a heuristic search was run, with 100 replicates of random addition, TBR branch swapping, and each replicate limited to saving 1000 trees if equal to or longer than the shortest tree initially found. This allowed sampling from several different islands of most parsimonious trees. The set of trees from this search was then used for branch swapping in subsequent parsimony searches, with MAXTREES set to 50,000 and the initial trees being retained. Bootstrapping was used as the measure of branch support (Felsenstein, 1985), and was done using 100 replicates, with MAXTREES set to 100, and

one replicate of random addition per bootstrap replicate. Fast bootstrapping methods such as this provide conservative measures of support (DeBry and Olmstead, 2000).

Waxy: Five variable characters, which were potentially sequencing artifacts, were dropped from *waxy* data set prior to analysis. Parsimony and bootstrapping analyses were as described above.

ITS: Only the ITS1 and ITS2 regions were used in the ITS analysis. Aside from two samples which appeared to be pseudogenes, there was no variation in the 5.8 S region. Three types of ITS analyses were done: (1) The ITS data was initially analyzed with no constraints, and without including the sequences from hybrids. (2) The data set was analyzed with the constraint that the stellate-haired taxa would form a monophyletic group. (3) The 11 hybrid sequences were included and another constrained analysis was run.

Secondary structure analysis: The secondary structures of two types of *P. heterophylla* ITS sequences, as well as the ITS1 indel region of various samples, were examined using the program RNAstructure (Zuker et al., 1999). Since the program cannot use ambiguously coded bases (G/A = R, etc.), consensus sequences of each ITS type, not containing ambiguous bases, were used for the analysis. Folding of the ITS1, ITS2 and 5.8 S regions was compared to that of a putatively functional tomato sequence (GenBank #X52265) which has been used in other studies of ITS structure (Hershkovitz and Zimmer, 1996). Base substitutions and indels from the entire ITS alignment were then mapped onto these secondary structures.

RESULTS:

Detection of contaminants: BLAST searches using both types of *P. heterophylla* ITS sequences returned Solanaceae sequences as being the most similar to them.

Identification of pseudogenes: Examination of the 5.8S region of the aligned data set revealed only three sequences with substitutions (1-3 nt). All sequences were from clones and appeared as long branches in an initial parsimony tree. Those sequences were removed from subsequent analyses, though they did group with other sequences from the same individuals in the preliminary analysis.

Examination of chromatograms: *Physalis heterophylla* had at least two major ITS types, one rare and one common, distinguished by differences in a 10 nucleotide indel region about 50 bp into ITS1. Some individuals had both types, based on the chromatograms, as the ITS2 and ITS5 sequences would become unreadable in the same area of ITS1. Cloning out ITS sequences proved this. Both ITS types were cloned from such an individual. Clones from two other individuals which did not appear to be polymorphic for the large indels were very similar to the initial direct sequences. Ten clones from three individuals were included in the phylogenetic analysis, though the original sequences from uncloned PCR products were not. No other taxa showed this pattern of multiple ITS types.

Waxy: The final *waxy* data set included 15 taxa and 34 individuals. Special effort was made to include multiple samples from species which were intensively sampled for ITS, and five sequences from *P. arenicola*, four from *P. heterophylla*, four from *P. virginiana* and four from Species A were used. The aligned region was 618 bp long, with

26 parsimony informative characters. An initial parsimony search resulted in 8400 most parsimonious trees, and swapping on these resulted in over 50,000 trees, each 47 steps long, with a CI of 0.91 (Figure 2.1).

The stellate-haired clade, including all members of section *Viscosae*, was strongly supported as monophyletic (93% bootstrap) and appeared in the strict consensus. The species from section *Lanceolatae* formed a paraphyletic grade at the base of the stellate-haired clade. Two weakly supported clades appeared in the strict consensus. One clade was comprised of 2 sequences from *P. heterophylla*, all samples of *P. virginiana*, and the one sequence from Species B. A second clade was formed by all sequences of *P. lanceolata* and Species A, and two *P. heterophylla* sequences. The *waxy* data set was not variable enough to resolve species relationships within any of the three clades.

ITS: The final ITS data set included 90 samples from 16 taxa, and 11 samples from artificial hybrids. *Physalis heterophylla* was represented by 25 samples from 16 individuals ranging through 13 states, and *P. virginiana* by 13 samples from eight states. Fourteen individuals of *P. arenicola* and 15 of Species A were also sampled.

In the aligned data set, ITS1 was 230 bp long, the 5.8 S region was 153 bp long, and ITS2 was 218 bp long. There were 99 variable characters, 68 of which were parsimony informative. The ITS1 region contained 43 of these; 25 were from ITS2.

ITS1+ITS2 unconstrained (hybrids out): The most parsimonious trees found were 190 steps long, and 2700 were saved from the initial parsimony search (data not shown). Swapping on these trees resulted in 50,000+ trees. The node separating the perennial species from the three outgroup taxa had 96% bootstrap support. Most species were either

monophyletic or paraphyletic, though often with little bootstrap support. However, the monophyly of *P. arenicola* was well supported (89%), as was that of *P. lanceolata* (82%). Sequences from *P. heterophylla* appeared in three different clades: one with *P. virginiana*, one isolated clade containing only *P. heterophylla* T-type sequences, and one clade of *P. heterophylla* A-type sequences with Species A.

There was no support for the monophyletic group of stellate-haired species. Three stellate-haired taxa, *P. cinerascens*, *P. walteri* and *P. angustifolia* grouped together with 66% bootstrap support, and formed the basal-most clade among the perennials. The rest of the perennial taxa formed a grade, with a clade of two stellate-haired species, *P. pumila* and *P. hispida*, at the base, followed by a clade of all *P. virginiana* sequences and four *P. heterophylla* sequences, then by a clade of the three Species B sequences. The next clade was formed by *P. mollis* and *P. viscosa*, two stellate-haired taxa. This clade was followed by two clades of *P. heterophylla* T-type ITS sequences, then by a grade of *P. heterophylla* A-type ITS sequences and Species A sequences. The grade terminated with a clade of *P. lanceolata* sequences sister to *P. arenicola* at its tip.

ITS1+ITS2 constrained (hybrids out): The most parsimonious trees found when the search was constrained to have a monophyletic clade of stellate-haired taxa were 194 steps long, which was four steps longer than the shortest trees from the unconstrained analysis (Figure 2.2). The initial parsimony search resulted in 8200 trees, and swapping on these trees resulted in over 50,000 trees. The structure of the majority rule consensus tree was similar to that of the unconstrained analysis, in that there was a clade consisting of *P. virginiana* and four *P. heterophylla* sequences and a clade consisting of *P.*

heterophylla ITS type A sequences, Species A and *P. lanceolata*, with the *P. heterophylla* T-type sequences at its base.

The placement of the ten cloned *P. heterophylla* sequences was determined by their ITS types. Clones with T-type sequences fell in the T-type clade, while clones with the A-type fell at the base of the Species A/*P. lanceolata* clade. Clones with the same ITS type from the same individual often grouped together. The majority of cloned sequences had relatively long branches compared to non-cloned sequences. As polymorphism was considered uncertainty in the analysis, this may be an artifact of their lack of polymorphic sites as compared to sequences from uncloned PCR products.

Constraining section *Viscosae* to be monophyletic did not have major topological effects, other than removing the *Viscosae* clades from among the *Lanceolatae* taxa. There were slight topological changes among the *Viscosae* taxa. *Physalis mollis* and *P. viscosa* moved to the base of the *P. arenicola* clade, and the *P. pumila*/*P. hispida* clade became the basal-most clade in the *Viscosae*, rather than the *P. cinerascens*/*P. walteri*/*P. angustifolia* clade.

ITS1+ITS2 constrained (hybrids in): Comparing the majority rule tree from this analysis to the one from the analysis without hybrid sequences shows that no major topological changes resulted from including the hybrid sequences (Figure 2.3). Though one hybrid was the product of a cross made between *P. heterophylla* and *P. arenicola*, that sequence grouped with its *P. heterophylla* parent. Thus, adding the hybrid sequences had no effect on the *Viscosae* clade. This hybrid also appeared with its *P. heterophylla* parent in an unconstrained analysis (data not shown).

The ten hybrid sequences resulting from crosses between *P. heterophylla* (ITS T-type) and Species A (ITS A-type) appeared as follows: four with the Species A clade, five with the *P. heterophylla* T-type clade, and one near its *P. heterophylla* parent in the *P. virginiana* clade. Hybrid sequences either grouped directly with one parent sequence, or fell at the base of a clade containing a parental sequence. Only one hybrid sequence failed to follow this pattern. Species A #96 x *P. heterophylla* #86 fell to the base of the *P. virginiana* clade, which contained no sequences from either of its parents. However, in some analyses, *P. heterophylla* #86 falls at the base of this clade or the base of the main *P. heterophylla* T-type clade, so the behavior of the hybrid sequence is similar to that of its *P. heterophylla* parent. Behavior of the hybrid sequences was similar in an unconstrained analysis, as well (data not shown).

Secondary structure analysis: For the *Physalis* sequences examined, the ITS1 region was the most variable in terms of nucleotide substitutions, and had the most variable secondary structure, making it difficult to determine which of the low energy secondary structures was most likely to be correct. All structures were characterized by the possession of 4-5 double stranded stem regions. The lowest energy structures of both *P. heterophylla* sequences were more similar to each other than either was to that of tomato. The conserved motif noted by Liu and Schardl (1994) was present, but was generally spread between two stem regions. Liu and Schardl's predictions were that it would fall within a single stem region, as was the case in the tomato ITS1 structure, and in slightly higher energy *Physalis* structures.

Physalis heterophylla had two distinct types of ITS sequences, characterized by variation at a 10 bp indel region which started near nucleotide 56 of ITS1 (Figure 2.4). The 10 nucleotide indel region in ITS1 mapped to the tip of a stem and loop structure in both *P. heterophylla* and tomato. The structures of the indel regions from several different taxa were compared, and involved the loss or gain of sets of base pairs at the end of the stem region and/or insertions and deletions in the single stranded loop at its tip (Figure 2.5). Changes at the A/T variable site (nucleotide #51) prior to the indel area resulted in the opening or closing of a small loop in the center of the stem region. Sequences with an adenine at that site are much more common than those with a thymine and represent structures with long stems, whereas those with a thymine have a central loop breaking the stem into two sections. The A-types generally have a lower free energy than the T-types.

All *Physalis* sequences used for phylogenetic analyses had identical 5.8S sequences. The 5.8S regions assumed a cruciform structure, with four stem regions connected to a central loop region, identical to that seen in tomato (data not shown). The one nucleotide which differed from the tomato sequence fell in the interior loop region. Three cloned sequences, which were not used in the parsimony analysis, had substitutions in the 5.8S region. Three of these occurred in stem regions directly adjoining loops, one occurred in the internal loop, and one occurred in a stem region one base pair away from a loop.

The ITS2 region was far more conserved than the ITS1 region, allowing greater confidence in the putative secondary structure (Figure 2.6). The *P. heterophylla* ITS2

sequences were very similar to each other and to that of tomato, all having a series of four long stem regions. The six conserved regions delimited by Hershkovitz and Zimmer (1996) were present, and mapped onto the secondary structure as predicted.

According to the putative ITS2 secondary structure for the *P. heterophylla* T-type ITS, there were 212 nt, 72 of which fell into single stranded regions and 140 of which were involved in double stranded regions. Variable sites from the 90 sequence alignment were mapped onto the secondary structure, showing that a total of 52 sites were subject to base substitutions. Single stranded sites at the tips of loops were the most likely to vary, while double-stranded sites within stem regions were the most conserved. See Table 2.1 for details. In the final ITS alignment, there were two indel regions in ITS2, both of which mapped symmetrically to loops at the ends of stem regions (Figure 2.6).

DISCUSSION:

In general, increased sampling within species resulted in finding a rare ITS type in *P. heterophylla*, and individuals with unique, relatively large indels in several species. As might be expected, sampling multiple ITS sequences from a single species did often result in a paraphyletic assemblage of sequences. However, it did not appear to result in increased phylogenetic signal. As many parsimony informative sites varied within species, adding samples often increased homoplasy while adding yet another individual to a data set which already had a high number of samples paired with a relatively low number of parsimony informative characters. One unexpected result of increased sampling was the discovery of what appear to be natural, backcrossed hybrids between *P.*

heterophylla and *P. virginiana*, which provided the unique opportunity to compare the phylogenetic placement of old versus new hybrids.

Using secondary structure to further examine patterns of variation within the ITS sequences of this species complex did a great deal to explain both why large indels occur in certain areas, and which structural features can effect the frequency of base substitutions. Structural considerations may help explain why large indels might be homoplasious more often than one would expect.

Effects of increased within species sampling: Three of the four species which were most intensively sampled proved to be paraphyletic for their ITS sequences, though for different reasons. Only *P. arenicola* was strongly supported as being monophyletic. The case of paraphyly between sequences of *P. heterophylla* was especially interesting. The presence of two distinct types of ITS sequences within *P. heterophylla* resulted in the division of the species into two clades, depending on the ITS type of the individual sequenced. Two other species within section *Lanceolatae*, Species A and *P. lanceolata*, have A-type ITS sequences very similar to those found in *P. heterophylla*, with an insertion of "TGCGTCCGCT" at the loop tip. The tendency of the *P. heterophylla* A-type sequences to cluster with sequences from these two taxa causes the paraphyly within *P. heterophylla*.

The most common ITS type in *P. heterophylla* is the rarest type among the U.S. perennial taxa and *Physalis* as a whole (Whitson, Chapter 1). Most taxa have an adenine paired to a thymine prior to the indel site at the tip of the second ITS1 stem and loop region. In *P. heterophylla*, substitution of a thymine for the adenine causes a loop to open

within the stem, potentially reducing the stability of the stem region. This T-type ITS is shared with *P. virginiana* and Species B, two species which from which no examples of A-type ITS sequences were found. Though a few other species within *Physalis* have substitutions at this site, a guanine replaces the adenine and thus maintains the double stranded structure of the region (Whitson, Chapter 1). Direct sequencing individuals of *P. heterophylla* resulted in mostly T-type sequences, several obviously polymorphic A/T mixtures, and a few A-type sequences. The chromatograms for the A-types generally showed low background typical of T-type sequences, suggesting that most individuals have at least some T-type sequences.

Hybridization is an unlikely explanation for the presence of A-type sequences within *P. heterophylla*. The widespread taxa which have similar sequence types are all members of section *Viscosae*. Artificial hybrids between *P. heterophylla* and these taxa are difficult to produce, and typically have reduced vigor (Sullivan, 1985). No reports of natural hybridization have been made. The two section *Lanceolatae* taxa which have sequences most similar to those found in *P. heterophylla* are both species with narrow ranges and specific habitat types. Species A is endemic to three counties in north Florida, and like *P. heterophylla*, prefers woodland habitats. However, *P. heterophylla* is very rare in north Florida and has not been reported from any of the areas where Species A has been collected. *Physalis lanceolata* occurs in the sandhill region where northern Georgia meets the Carolinas, and though *P. heterophylla* also occurs there, it does not prefer the sandhill habitats favored by *P. lanceolata*. Additionally, there is no geographical pattern to the presence of A-type ITS sequences in *P. heterophylla*, as might be expected if

hybridization with a narrow ranging species was responsible for the presence of this variation.

As a widespread species with weedy tendencies, it seems likely that incomplete lineage sorting explains the maintenance of two ITS types within *P. heterophylla*. Lineage sorting, which occurs as the gradual loss of different alleles from a species, is dependant on both time and population size (Doyle and Davis, 1998). The U.S. perennial taxa are one of the more derived groups of species within *Physalis*, and probably represent a relatively recent radiation of species (Whitson, Chapter 1). The large range and population size of *P. heterophylla* could additionally act to maintain allelic diversity. The four other *Lanceolatae* species for which multiple ITS types were not seen either have vastly smaller range sizes (*P. lanceolata*, Species A, Species B) or generally smaller population sizes (*P. virginiana*).

Different ITS types have also been found within species of the genus *Cucurbita*, but are thought to be the result of hybridization and/or polyploidization events during the domestication of these species (Jobst et al., 1998). The maintenance of multiple ITS types is found in a variety of allopolyploid taxa, including *Amelanchier*, *Draba* and *Silene*, (Campbell et al., 1997; Widmer and Baltisberger, 1999; Popp and Oxelman, 2001). However, *P. heterophylla* is an unusual example in that it is a diploid species.

The placement of *P. lanceolata* within Species A is surprising, since the two taxa have nonoverlapping ranges. *Physalis lanceolata* is strongly supported as being monophyletic, but all of the A-type ITS sequences found within section *Lanceolatae* seem to be very similar, as the inclusion of several A-type *P. heterophylla* sequences

within the Species A/ *P. lanceolata* clade demonstrates. If the three taxa are all derived from the same widespread, polymorphic ancestral taxon, it is possible that Species A and *P. lanceolata* fixed the same sequence type, and have not yet diverged enough to form distinct clades.

Characterization and phylogenetic placement of sequences from hybrids: The paraphyly seen in *P. virginiana* has a different origin than that seen in *P. heterophylla*, though the two species have similar ranges and habitat preferences. All 13 sequences sampled from *P. virginiana* formed a clade, but four sequences from *P. heterophylla* nested within it. Constraining *P. virginiana* to monophyly added six steps to the tree, and simply shifted the four *P. heterophylla* sequences to the base of the clade. Constraining the four T-type *P. heterophylla* sequences to join the T-type *P. heterophylla* clade added ten steps.

Interspecific hybridization can explain why sequences from these two taxa appeared in the same clade, as *P. virginiana* and *P. heterophylla* have been known to hybridize in the wild, though only under unusual circumstances (Hinton, 1975). Artificial crosses between the two taxa do not have a high rate of success, but when fruit do develop, seed set is similar to that seen in non-hybrid fruits (Menzel, 1951). Individuals morphologically intermediate between the two species are occasionally seen in herbarium material, but are rare. However, the individuals sampled for this study were all morphologically typical of their respective species.

Both the ITS and *waxy* gene trees show some *P. heterophylla* sequences grouping with *P. virginiana*. However, even in the large ITS data set, no *P. virginiana* sequences

group with other clades of *P. heterophylla*. About 25% of the *P. heterophylla* individuals sampled had sequences which appeared to be of hybrid origin, possessing mostly *P. virginiana*-like characters with a few potentially recombinant sites. Given a similar percentage of hybrid individuals in *P. virginiana*, the 13 individuals sampled should have produced about 3 sequences with *P. heterophylla*-like characters. This was not the case. Thus, if the presence of *P. heterophylla* sequences in the *P. virginiana* clade is a result of hybridization, then concerted evolution seems to be biased in favor of *P. virginiana*-type sequences.

Gene conversion is a mechanism which could explain this result. Unlike other methods of concerted evolution such as unequal crossover, gene conversion can result in the biased conversion of multicopy genes in hybrids to one parental type (Hillis et al., 1990). Such biased conversion of ITS sequences has been noted in other angiosperms, including *Cardamine* and *Armeria* (Franzke and Mummenhoff, 1999; Augilar et al., 1999). The process can be rapid: almost total conversion to one parental type was seen in F2 *Armeria* hybrids (Augilar et al., 1999), and in the case of allopolyploid *Cardamine* hybrids, populations of these plants arose within the last century, when man-made disturbance brought the parental taxa into contact (Franzke and Mummenhoff, 1999).

The *P. virginiana* sequences group into two major clades, which seem to be loosely correlated with geography. One clade consists of individuals from the Carolinas and Florida, whereas the other clade contains individuals from Minnesota, Oklahoma, and Illinois, along with some individuals from Southeastern states. The four *P. heterophylla* sequences which group with *P. virginiana* also follow this pattern. The one

sequence from South Dakota falls into the wide-ranging Midwestern/Southeastern clade, whereas the other three sequences, from Mississippi and Tennessee, group with the Carolinas clade.

The three *P. heterophylla* sequences in the Carolinas clade form a monophyletic group on the basis of one shared character -- a thymine at nucleotide 167 of ITS1. This is a feature unique to *P. heterophylla*, and though it varies between thymine and guanine within the species, all other species surveyed are fixed for guanine. Sequences most similar to one parental type but retaining a few characters from the other have also been seen in *Armeria* (Aguilar et al., 1999) and *Cardamine* hybrids (Franzke and Mummenhoff, 1999).

That sequences from *P. angustifolia* and *P. walteri* form a paraphyletic group in the *Viscosae* clade can also be explained by hybridization. These species form a hybrid zone where their ranges overlap in Florida, and experimental crosses have resulted in highly viable seeds (Sullivan, 1985). These species are closely related, as shown by the *waxy* tree (Figure 2.1), and interbreeding, so their resultant paraphyly is not unexpected.

Whereas sequences from individuals which arose from a hybridization event followed by backcrossing to one or both of the parental taxa were nearly identical to sequences from one of the parental species, sequences from F1 hybrids followed a different pattern. Chromatograms of direct sequenced F1 hybrid individuals showed additivity at the sites where the two parental taxa differed. When one parental sequence showed a polymorphic site and the other did not, the hybrid often inherited that polymorphism, as well. Additivity of ITS sequences has also been seen in F1 hybrids

between diploid species of *Armeria*, though one variable site was already biased toward one parental sequence (Aguilar et al., 1999). Inheritance of indel characters from both parents resulted in unreadable chromatograms at the ITS1 indel site. When one parent had a unique indel which the other parent lacked, the gap appeared in the direct sequenced product from the hybrid, as was predicted by Baldwin et al. (1995).

When used in the phylogenetic analysis, hybrid sequences grouped directly with one of their parental sequences or at the base a clade containing one of their parental sequences. This is similar to results seen by McDade (1992), who examined the effects of including data from hybrid individuals of *Aphelandra* (Acanthaceae) in a morphological cladistic analysis. Hybrid *Physalis* sequences generally seemed to be more similar to one or the other of their parental sequences, as hybrids resulting from crosses between Species A (A-type ITS) and *P. heterophylla* (T-type ITS) did not obscure the A and T ITS clades by falling out between them, but grouped with one or the other clade. This agrees with other workers who have found that inclusion of hybrids between closely related lineages had little phylogenetic impact (McDade, 1992; Rieseberg and Ellstrand, 1993), and that these hybrids did not behave in a way which distinguished them from non-hybrid taxa. All in all, sequences from both new and old hybrids fit seamlessly into the phylogenetic analysis, and were only identifiable as potential hybrids either by their many polymorphic sites or by the mismatch between the ITS sequence and the morphology of the individual under study.

Secondary structure and patterns of character variation: ITS1 was longer and more variable than ITS2. Because ITS1 has fewer conserved regions than ITS-2 (Liu and

Schardl, 1994; Hershkovitz and Zimmer, 1996), it was difficult to gauge which of the low energy secondary structures was most likely to be correct, though most of them shared the same 4-5 stem and loop regions. Multi-stemmed ITS1 structures have also been reported in other taxa, such as *Arabidopsis, Aeschynanthus,* and *Quercus* (Liu and Schardl, 1994; Denduangboripant and Cronk, 2000; Mayol and Rosselo, 2001).

While ITS1 had two indel sites, one large and one small, ITS2 had three, two large and one small. Many individuals shared the indels found in ITS1, whereas the indels found in ITS2 appeared in only 1-3 individuals and were not characteristic of any species as a whole. All of the large indels, whether found in ITS1 or ITS2, were located at the tips of stem structures. Both of the small 1 bp indels were found in strings of cytosines within stem structures.

Mapping the ITS1 indels onto the constrained phylogeny indicates that homoplasy is an issue, regardless of whether one invokes several losses of base pairs resulting in taxa with gaps, or several gains of base pairs resulting in the insertions seen in taxa such as *P. lanceolata* (Figure 2.7). However, with this large data set and only low levels of variation, there is little support for any specific topology. Constraining all of the taxa with gaps to form a monophyletic group only adds three steps to the tree.

Considering ITS sequence variation in this indel region over *Physalis* as a whole shows that most taxa have the full 10 base pairs. The *P. walteri* type indel is found in many different taxa, and could represent the retention of an ancestral ITS type in the *P. cinerascens/ P. walteri/ P. angustifolia* clade. Other *Physalis* taxa which do not belong to

the U.S. perennial clade also have gaps in this region, suggesting that gaps have arisen multiple times.

If all of the perennial taxa lacking gaps do form a monophyletic group, then lineage sorting has resulted in the taxa from different sections fixing a wide variety of ITS types, which seems unlikely since the group is so strongly supported as monophyletic according to the *waxy* data. It seems just as likely that the indel variation in the perennial taxa has resulted from multiple losses of bases from the ancestral ITS type. It is quite likely that structural constraints limit both how many bases are lost and how many bases are gained. All taxa with the 10 nucleotide indel had a 2 nucleotide loop, then 2 sets of base pairs, then four nucleotide loop conformation (Figure 2.5). If some amount of base pairing is necessary to keep the indel loop region stable, then perhaps the base composition of insertions is constrained to favor ones with high G/C contents, which can maintain base pairing.

The ITS2 region was much more conserved than ITS1, and the folded sequence adopted a four stemmed conformation that is conserved from the Volvocalean green algae to the angiosperms (Mai and Coleman, 1996). Mapping of base substitutions on this structure showed that more than half of the sites at the tips of loops were subject to base substitutions, followed by other single stranded sites and the double stranded sites adjoining them (Table 2.1). Double stranded sites which did not adjoin single stranded regions were the least likely to have base substitutions. In ITS1, the situation was less conclusive, and base substitutions seemed to be most concentrated in the third and fifth stem and loop regions. Within those regions, base substitutions appeared to occur in

single stranded areas just as often as they occurred in double stranded ones. The first, second and fourth stem and loop regions were more conserved, and seemed to more closely follow the patterns found in ITS2, with many changes at the loop tips, some changes in other single stranded areas or in the stem regions directly adjoining them, and fewer changes in well-confirmed double stranded regions.

CONCLUSIONS:

Sampling of ITS sequences from multiple individuals within four closely related species of perennial *Physalis* resulted in three cases of paraphyly, one due to putative hybridization and biased gene conversion, one due to polymorphism for two different ITS sequence types and incomplete lineage sorting, and one due to the recent divergence of two taxa sharing similar ITS types. Including artificial hybrids in the phylogenetic analysis had little topological effect, and the hybrid sequences did not behave in a way which distinguished them from their parents. These results suggest that using ITS alone to determine species relationships between closely related taxa may not be desirable, and that two-gene analyses are likely to be both more resolved and potentially less misleading.

Structural analysis of these ITS sequences indicated that most large indel regions occur at the looped tips of stem regions, and may be subject to various structural constraints, including constraints on size and maintenance of some base pairing within large insertions. Though they do provide some phylogenetic information, these large

indels are often homoplasious, and mapping of the characters onto a phylogeny is probably a better approach than coding them and including them in the analysis.

Chapter 3.

Biosystematics of a new taxon of *Physalis* and related species from the Apalachicola river bluffs

INTRODUCTION:

About 20 species of ground-cherries (*Physalis*) are native or naturalized to the southeastern United States, and half of them can be found in northern Florida. An especially interesting area for study of the genus *Physalis* is the Apalachicola river bluffs in northwestern Florida. Considered a glacial refugium, this two-county region has long been known for its unique flora (James, 1961; Wolfe et al., 1988). Along with a diversity of common southeastern woodland taxa, the limestone bluffs harbor a rich array of temperate taxa found nowhere else in subtropical Florida, relict species such as *Torreya taxifolia*, and many species and subspecies endemic to the region (James, 1961; Mitchell, 1963; Platt and Schwartz, 1990).

Eight species of *Physalis* are reported to occur there. Five are common throughout Florida, but three fit the endemic floristic pattern of the river bluff region. Two of the species are temperate taxa that are locally rare. *Physalis virginiana* and *P. heterophylla* are common species in the Northeastern United States, but reach the southern limits of their ranges in the western Florida panhandle. The third species is endemic to the Southeast, and found only in four states: Florida, Alabama, Mississippi and Louisiana. *Physalis carpenteri* is morphologically unusual, uncommon throughout its range, and

considered to be a relict species, with closer affinities to the woody Mexican taxa than to any of the United States species (Menzel, 1951).

An interesting new addition to these examples are several recently discovered populations of a diminutive, unidentified perennial *Physalis* found while searching for the elusive *P. heterophylla*. These plants are morphologically suggestive of both *P. arenicola* and *P. heterophylla*, two species that are taxonomically problematic in this region – a situation which could explain why the river bluff *Physalis* populations have heretofore escaped botanical notice.

Physalis heterophylla, commonly known as the 'clammy ground-cherry' is a large, somewhat weedy and wide-ranging perennial species, and has often been reported to occur in Florida, though few specimens have been cited to document this claim. *Physalis arenicola*, the 'sand ground cherry,' is a smaller species, most common in Florida and occasionally found along the southern edges of adjoining states. Both taxa are vegetatively variable, making it difficult to simply define either species on the basis of leaf shape or hair characters. Though these species can be vegetatively similar, rhizome morphology clearly differentiates them (Sullivan, 1984a). *Physalis heterophylla* is a species adapted to a wide variety of soil types, and has short, stout, deeply-buried rhizomes. *Physalis arenicola* is a species adapted to loose, sandy soils, and has slender, running rhizomes which parallel the soil surface and can extend for several feet from the parent shoot. While rhizome morphology is an easy way to distinguish these two taxa, it is only rarely used. When pulled, the stems of both species tend to snap off where they join the rhizomes, meaning that most herbarium material consists solely of aerial shoots.

The unidentified *Physalis* found along the river bluffs are smaller plants than *P. arenicola*, and superficially resemble young shoots of that species. However, unlike *P. arenicola*, their rhizomes are stocky and deeply-buried.

Though several areas in northern Florida, and especially Jackson and Liberty counties, were searched, populations of the river bluff *Physalis* were found in only three places: along the Apalachicola river bluffs at Torreya State Park, and at Three Rivers State Recreation Area, and along the Chipola river bluffs at Florida Caverns State Park. The plants are locally common, favoring disturbed, partially shady woodland sites such as old road cuts, trail sides, and washing bluff tops. Their distribution is patchy, with isolated clusters of shoots scattered over many square meters, and ranging in size from groups of 10-20 shoots to groups of 100 or more. The plants seemed generally healthy, but flowering, and especially fruiting, individuals were less common than sterile shoots.

In addition to the river bluff taxon, *Physalis arenicola* is found in all three parks, and was occasionally seen growing intermixed with populations of the river bluff plants, though it prefers open, sandy areas to the wooded, calcareous river bluffs. Only a few individuals of *Physalis virginiana* were seen, all at Three Rivers State Recreation Area, and no representatives of *P. heterophylla* were found at all. As *P. virginiana* and *P. heterophylla* are closely related species, with similar ranges and habitat preferences, the dearth of *P. heterophylla* was somewhat surprising.

The focus of this study was twofold. The first goal was to taxonomically characterize the river bluff taxon. The rhizome morphology of the plants was similar to that of *P. heterophylla*, which was expected to occur in the area, but had not been

observed. Although their vegetative morphology is distinctive, *P. heterophylla* is known to be vegetatively variable. Thus, the river bluff taxon might represent a locally adapted variant of *P. heterophylla*. Alternatively, these plants might represent a previously undescribed taxon.

Two types of data were used to address the question of identity. First, DNA sequence data from two regions of nuclear DNA from a wide sampling of the U.S. perennial *Physalis* species in sections *Lanceolatae* and *Viscosae* were used to address the broad question: To which of U.S. perennial species is the river bluff taxon most closely related?

After using the phylogenetic analysis to identify the closest relative of the river bluff taxon, morphological data in the form of leaf-shapes was compared between the two taxa to help address a second, more specific question, which was: Is the river bluff taxon a locally adapted variant of its closest relative, or do these plants represent a new species? Finally, crossing studies were used to help determine the amount of compatibility between the river bluff plants and their closest relative, as well as to help address some questions from the next part of this study.

The second goal of this study was to learn more about how these rare river bluff plants persist in the wild. They obviously have a limited range, specific habitat requirements, and are outnumbered by their more common relative *P. arenicola*. Field observations also indicated that their opportunities for sexual reproduction might be rare. Isozyme data from multiple local populations and greenhouse crossing studies were used to address the following questions: (1) How are the populations of the river bluff taxon

plants reproducing? (2) Are they able to take advantage of potentially rare reproductive events by self-pollinating, or are the majority of these rhizomatous plants simply vegetative clones of a few successful individuals? (3) Is their any evidence that they are in danger of being genetically swamped out through hybridization with the more common *P. arenicola*?

MATERIALS AND METHODS:

DNA extraction, amplification and sequencing: DNA was extracted from fresh leaves or flowers using a miniprep modification of Doyle and Doyle's CTAB procedure (1987), or using DNeasy Plant Mini kits (Quiagen Inc., Valencia, California, USA).

Two regions of nuclear DNA were initially amplified, the ITS region of the nrDNA and an approximately 614 bp segment of the nuclear gene for granule bound starch synthase (*waxy*), from exon eight to ten. When the first two regions of DNA sequenced did not provide enough resolution to determine the closest relative of the river bluff taxon, a small additional data set was generated by amplifying and sequencing an upstream segment of *waxy*. Two primers specifically designed for *Physalis*, WTP F (5'-ATGGCAAGCATCACAGCTTC-3') and PHY R (5'-GTCATTACTCGATGTCCGCG-3'), were used to amplify and sequence an approximately 500 bp fragment of the gene, consisting mostly of exon one, which codes for the transit peptide, one of the most variable regions of the immature *waxy* protein (Salehuzzaman et al., 1993). PCR and sequencing protocols generally followed Whitson (Chapter 1).

Taxon sampling: Earlier phylogenetic work using ITS and *waxy* sequences has shown that most of the U.S. perennial species of *Physalis*, which fall into two sections of the genus, *Viscosae* and *Lanceolatae*, form a monophyletic group (Whitson, Chapter 1). Members of section *Viscosae* typically have stellate hairs, while members of section *Lanceolatae* have simple hairs. The river bluff taxon is a simple-haired species, but some members of the *Viscosae* have simple hairs as well, so both sections were well-sampled. The final combined data set included 21 individuals, representing six of the seven species from section *Viscosae*: *P. angustifolia*, *P. pumila*, *P. cinerascens*, *P. mollis*, *P. walteri* and *P. viscosa*, all but one of the United States native members of section *Lanceolatae*: *P. heterophylla*, *P. lanceolata*, *P. virginiana*, and an undescribed species from Texas (labeled Species B in figures), and *P. arenicola*, which is not currently assigned to a section. *Physalis longifolia*, technically a member of section *Lanceolatae*, is more closely related to several Mexican taxa, and thus was not included (Whitson, Chapter 1). Three species were used as outgroups, one common weedy annual, *P. angulata*, and two cultivated taxa *P. philadelphica*, and *P. peruviana*.

Phylogenetic analysis: Analyses were done using PAUP* version 4.0b8 (Swofford, 1998). The data sets were analyzed separately and in combination using heuristic search methods and maximum parsimony as the optimality criterion. Gaps were treated as missing data. Bootstrapping (Felsenstein, 1985) was used as the measure of relative support. For analysis of the individual data sets, bootstrapping was done using 100 replicates and TBR branch swapping. Starting trees for each replicate were generated with random addition, and MAXTREES was set to 10,000.

ITS: A heuristic search with 100 replicates of random addition and TBR branch swapping was used to generate an initial set of most parsimonious trees. This set of trees was then used for branch swapping in another heuristic search, with both the initial trees and new trees being retained.

Waxy, exons eight to ten: Five variable characters, which appeared to be sequencing artifacts, were removed from the *waxy* data set. An initial parsimony search was done to gauge the length of the shortest trees, then a heuristic search was run, with 100 replicates of random addition, TBR branch swapping, and each replicate limited to saving 1000 trees if equal to or longer than the shortest tree initially found.

Waxy, exon one: A heuristic search with 100 replicates of random addition and TBR branch swapping was used to generate the set of most parsimonious trees.

Combined (3-region) data set: A heuristic search with 100 replicates of random addition and TBR branch swapping was used to generate the set of most parsimonious trees. The bootstrap analysis consisted of 100 replicates, each with ten replicates of random addition, and no MAXTREES limit.

Common Garden Experiments: Wild-collected individuals of *P. heterophylla*, *P. arenicola* and the river bluff taxon were planted in one gallon pots with the same soil mixture, and grown under the same greenhouse conditions. This was to provide fresh material for morphological examination, crossing studies, and allozyme extractions, and to determine the extent of morphological changes induced by the change in growing conditions.

Crossing: Three taxa were used in crossing studies: *P. heterophylla*, *P. arenicola* and the river bluff taxon. The majority of crosses were done in the greenhouse, and reciprocal crosses were made when possible. Three types of hand pollinations were done: self pollinations, crosses between individuals of the same species, and crosses between individuals of different species.

When available, young flowers whose anthers were either unopened or just opening were used. The anthers were removed, and the stigma examined for the presence of pollen. Given that no pollen was found on the stigma, the emasculated flower was then pollinated by using tweezers to brush an open anther against the stigma. The corollas were removed to make the flowers unattractive to any pollinators which may have entered the greenhouses. Flowers on plants outside were also bagged. Each flower was marked by hanging a numbered, color-coded paper clip at that node. Because many of the plants were infected by downy mildew and were prone to drop their leaves and fruit, all crosses were kept under observation for calyx enlargement, so that some measure of compatibility could be gained from crosses which did not result in the production of fruit.

Fruit resulting from hand pollinations were collected and the seeds were counted and qualitatively observed for signs of viability (size, thickness, color). A maximum of 25 seeds per fruit were planted in labeled flats and kept on a mist bench until germination had occurred. Seeds which did not germinate after four weeks (by which time germination from fruit with highly viable seeds had stopped) were considered inviable.

Leaf shape analysis: Three leaves each from 15 individuals of the river bluff taxon and 20 of *P. heterophylla* were pressed and then digitized using the program

Image-1 (Universal Imaging Corp., 1991). Leaf selection was designed to result in sets of leaves characteristic of the individuals sampled. The largest healthy leaves were taken from the third node from the stem tip. When no such leaves were available, three of the most standard looking, healthy leaves were chosen.

Prior to imaging, the petioles of the leaves were hidden using white tape, and they were placed adaxial side up under a sheet of overhead film to keep them flat. For leaves with decurrent bases, the blade was considered to end where the amount of blade running along each edge of the petiole fell below 1 mm wide. Leaves were oriented horizontally across the field of view and edgelists generated such that the first point was the leaf apex.

The edgelists were then run through the program NUKEDUPS to remove duplicated points, and the resulting files used for elliptic Fourier analysis (EFA) using the program written by Ferson et al. (1985). For each leaf, size and location were factored out and 25 harmonics were generated. The number of harmonics was decided upon by taking the edgelist of one of the most complex leaves and running it through the EFA program numerous times, generating 5 harmonics, then 10, then 15 and etc., to 50 harmonics. Sets of 75 and 100 harmonics were also generated. These analyses were then used to regenerate edgelists for the leaf, and all of the edgelists were plotted out and compared for detail.

Average leaf shapes for the two taxa were generated by averaging the Fourier harmonics of each set of samples. The resulting harmonics were plotted in Mathematica 4.0 (Wolfram Research, Inc., 1999.) using a function written by J. Mercer (Duke University, Biology Department).

The Fourier harmonic data were used for principal components analysis. The Wilks Lambda Criterion (a type of MANOVA) was also used to determine if the means of the two sets of leaf-shape samples were significantly different. Both analyses were run using the program Data Desk 5.1 (Data Description, Inc., 1995.).

Allozymes: A total of 129 individuals from seven populations of the river bluff taxon and 86 from six populations of *P. arenicola* were screened for variability at four loci: 6PGD, PGI. PGM and IDH. The majority of collections were made at three parks: Torreya (Liberty Co., FL), Florida Caverns (Jackson Co., FL) and Three Rivers (Jackson Co., FL). Isolated populations of the river bluff taxon were sampled from each park, as well as populations found growing close to or intermixed with populations of *P. arenicola*. Isolated populations of *P. arenicola* were sampled from Three Rivers and two areas in Liberty County a few miles from Torreya State Park. (See Table 3.1 for a list of locations.) Allozyme data were analyzed using the program BIOSYS-1 (Swofford, 1989).

Leaves were ground on ice in the following extraction buffer, which was designed to minimize oxidation due to secondary compounds: 100 mM Tris, 1 mM Na4EDTA, 10 mM KCl, 10 mM anhydrous MgCl2, 50 mM NaAscorbate. The pH was adjusted to 7.4 with HCl, and the solution stored in the dark at 4° C. Just prior to use, the final ingredients were added: 0.1% of 10% Triton X-100, 5.0% PVP-40, 0.2% B-mercaptoethanol, 10% DMSO and 10 mM DIECA. Wicks not used immediately were stored at –80° C.

Protocols for horizontal starch gel electrophoresis and staining generally followed Soltis and Soltis (1989). Wicks were run on 11% starch gels using a pH 5.7 Histidine-

citrate system, with the electrode buffer being a 1/6 dilution of the gel buffer. Gels were run in a cold room at 4° C for 3.5-4 hours. Determination of number of loci per enzyme and number of alleles per locus was based on information about the inheritance of these enzymes in other plant species. Alleles were generally scored using the following system: the slowest band on the gel was designated allele A, the next slowest designated B, etc.

RESULTS:

ITS data set: The aligned ITS data set consisted of 676 characters, including 38 variable but parsimony uninformative characters and 58 that were variable and parsimony informative. It included 35 sequences, representing four outgroup taxa, seven species from section *Viscosae*, and a broad sampling of sequences from species in section *Lanceolatae*, including six from the river bluff taxon, four from *P. heterophylla*, five from *P. virginiana*, four from *P. arenicola*, three from *P. lanceolata*, and one from Species B.

Examination of the ITS alignment showed that *P. heterophylla* was represented by two ITS types, marked by an indel region in ITS1 (Figure 3.2). The most common type had a thymine just prior to a 6 bp gap. This T-type indel was shared by *P. virginiana* and species B. The rarer type had an adenine followed by a 6 bp insertion. The A-type indel was also found in the river bluff taxon and *P. lanceolata*, and was similar to the indel types found in the stellate-haired taxa.

The initial heuristic search resulted in 804 most parsimonious trees, each of 172 steps. Swapping on those trees resulted in a final set of 807 most parsimonious trees. In

the strict consensus (not shown), both stellate and simple-haired taxa were intermixed, and *P. arenicola* was the only ingroup species supported as being monophyletic (95%). Sequences of *P. heterophylla* were highly paraphyletic, and several grouped with sequences of *P. virginiana* or with the river bluff taxon. Though there was bootstrap support for some clades, support for relationships within clades and for the backbone of the tree was weak or absent.

There were three major clades of the perennial species. One clade, with 51% bootstrap support, consisted of sequences from *P. cinerascens*, *P. walteri* and *P. angustifolia*, all species from section *Viscosae*. This clade was sister to the other two clades, one of which included all of the *P. virginiana* sequences and three *P. heterophylla* sequences. There was 83% bootstrap support for grouping all of the *P. virginiana* sequences with two of the *P. heterophylla* sequences. The largest clade was split into two sister groups, a small group with no bootstrap support consisting of Species B, *P. viscosa* and *P. mollis*, and a large clade with 65% bootstrap support including all sequences from the river bluff taxon, *P. lanceolata*, and *P. arenicola*, and one from *P. heterophylla*. The position of *P. pumila*, a stellate-haired taxon, was unresolved.

Waxy, exons eight to ten: This aligned *waxy* region was 613 bp, with 26 parsimony informative characters. The data set included 33 sequences, and was similar to the ITS data set, but only included three sequences from the river bluff taxon. A set of 91,000 most parsimonious 44 step trees was saved.

The strict consensus (not shown) was less resolved than that of the ITS data set, and had one major topological difference. The stellate-haired species were strongly

supported (96%) as being monophyletic. This clade included a monophyletic *P. arenicola* (81%), a monophyletic *P. cinerascens* (61%), *P. viscosa*, *P. pumila*, *P. mollis*, *P. angustifolia* and *P. walteri*. There were two weakly supported clades of simple-haired taxa, including one clade (53%) with Species B, all of the *P. virginiana* sequences, and two sequences from *P. heterophylla*, and a smaller clade consisting of two sequences from the river bluff taxon and two from *P. heterophylla*. The placement of the *P. lanceolata* sequences and the two other river bluff taxon sequences was unresolved.

Waxy, exon one: The aligned data set was 487 bp long, with 24 variable, but parsimony uninformative, characters and 22 which were variable and parsimony informative. The data set included 21 sequences, with the same taxa as the other data sets, but only one *P. arenicola* sequence, two *P. heterophylla* and *P. virginiana* sequences, and three river bluff taxon sequences. There were 1734 most parsimonious trees of 54 steps.

The strict consensus tree (not shown) was less resolved than those of either of the earlier analyses. Only three clades had bootstrap support. The three *P. lanceolata* sequences formed a strongly supported (100%) monophyletic group. There was 78% support for *P. pumila* being sister to *P. mollis*, and the large clade formed by Species B, and all of the sequences from *P. heterophylla*, *P. virginiana*, and the river bluff taxon had 62% support. The sequence from Species B was sister to the *P. heterophylla*/ *P. virginiana*/ river bluff taxon clade, which was strongly supported (96%) as forming a monophyletic group.

Combined analysis: The combined, three-region data set had 85 parsimony informative characters and sequences from 21 individuals of 14 taxa. Three samples were

missing some data. For *P. angulata*, the *waxy* sequence from one individual was paired with the ITS sequence from another. Since ITS varies little within this species (Whitson, unpublished) and *waxy* is less variable than ITS, this was considered acceptable. *Physalis angulata* was also represented by a partial *waxy* exon one sequence. Two of the three individuals (underlined in Figure 3.3) of the river bluff taxon had been sequenced for ITS and *waxy* exon one, but not the other segment of *waxy*. While other individuals of the river bluff taxon had been sequenced for the larger segment of *waxy* alone, it was considered more conservative to leave the data missing than to use *waxy* sequences from other individuals.

The combined analysis resulted in four most parsimonious 288 step trees (Figure 3.3). Within the perennial taxa, four clades of taxa had bootstrap support. The stellate-haired taxa formed a monophyletic group, with 64% bootstrap support. The three sequences from *P. lanceolata* were strongly supported (100%) as forming a monophyletic group. In three of the four most parsimonious trees, *P. lanceolata* was sister to a clade formed by the river bluff taxon sequences and one *P. heterophylla* sequence. The river bluff taxon/*P. heterophylla* clade had 73% bootstrap support. Finally, the clade consisting of two *P. virginiana* sequences and two *P. heterophylla* sequences had 79% bootstrap support. Species B was sister to this clade in three fourths of the most parsimonious trees.

Common garden results: Though the plants were not subjected to quantitative measurements, several observations were made. The vegetative parts of cultivated plants were generally larger than those of field-collected plants, though characters such as leaf shape and the relative sizes of each species did not change. For example, field-collected

plants of *P. heterophylla* are generally much larger than those of *P. arenicola*, and this remained true under greenhouse conditions. Hair types, which are crucial characters for identifying these taxa, did not change in response to cultivation. Cultivated plants were generally more floriferous than field-collected specimens, but unlike leaf and shoot size, flower size did not increase.

Crosses: The three species responded differently to self pollination, though none produced fruit (Table 3.2). *P. arenicola* was the least tolerant of self pollen. In all 15 crosses, the offending buds were dropped before any calyx enlargement took place. Self pollinations of the river bluff taxon resulted in calyx enlargement 33% of the time. Slight ovary enlargement followed in a few cases, but fruit never developed. It is likely that *P. heterophylla* is also self-incompatible. Only three self pollinations were performed with *P. heterophylla*, since both Menzel (1951) and Hinton (1975) reported that the species is self-compatible. However, all three buds were dropped, and no calyx enlargement was seen. Similar results have been seen by other workers, and some self-incompatible taxa will accept self-pollination if the flower is pollinated before the bud opens (personal communication, Y. Lu). In this study, flowers were pollinated after they opened, but bud pollination was the technique used by Menzel (1951) and Hinton (1975), which may explain why they considered *P. heterophylla* self-compatible.

The results of crosses made between different individuals within a species initially differed based on which species was involved. Pollination caused calyx enlargement 58% of the time in the river bluff taxon, as opposed to 39% in *P. arenicola* and 31% in *P. heterophylla*. However, the percentage of mature fruit resulting from hand

pollinations was similar in all three species, ranging from 28%-33% (Table 3.2). At about 30 seeds per fruit, the river bluff taxon had the lowest average seed set, while *P. heterophylla* had the highest, at 66. Of the three taxa, *P. heterophylla* generally had the largest fruit.

Seeds planted from four *P. heterophylla* fruit and five from the river bluff taxon resulted in about 40% germination for both taxa. Seed viability varied widely from fruit to fruit, ranging from 60% to 11% in the river bluff taxon, and 64% to 6% in *P. heterophylla*. One *P. heterophylla* fruit was left out of the viability calculations because it was picked early, due to the parent plant being damaged. All 153 seeds were tiny, flat and obviously immature. Surprisingly, a few germinated, though seedlings were exceedingly small, with cotyledons under 2 mm long when they first emerged.

Fifty crosses (including reciprocals) were made between the river bluff taxon and *P. heterophylla*. Mature fruit resulted 35% of the time. When the river bluff taxon was the maternal plant, mature fruit resulted 54% of the time, as opposed to 14% with *P. heterophylla*. Seed set averaged about 40 per fruit in the river bluff taxon, and about 62 per fruit in *P. heterophylla*.

Seeds planted from three *P. heterophylla* fruit and 11 the river bluff taxon fruit resulted in approximately 61% germination for the river bluff taxon and 52% for *P. heterophylla*. Seed viability varied widely from fruit to fruit, ranging from 0%-96% in the river bluff taxon, and 24%-100% in *P. heterophylla*. In one fruit from the river bluff taxon, none of the 25 seeds planted germinated, though they looked healthy. Another fruit from the same plant had 87% seed germination. This was one of only two sets of seeds

which did not exhibit any germination. The other example was from a between species cross that made only a few seeds.

Crosses made between the river bluff taxon and *P. arenicola* were rarely successful. Only 6.5% resulted in mature fruit, and only two of the four fruit produced seeds. All of the successful crosses were made with the river bluff taxon as the maternal plant. Seed set was low as well, averaging about 7 seeds per fruit. A total of 29 seeds were planted, from 2 different fruit, but none germinated.

Crosses made between *P. heterophylla* and *P. arenicola* followed a pattern similar to those made between the river bluff taxon and *P. arenicola*. Mature fruit resulted from14.3% of the crosses, and *P. heterophylla* was the maternal plant in all of these. Seed set was low to moderate, ranging from 12-22 seeds. Of the 34 seeds planted, four germinated.

Seeds resulting from the hand pollinations were all planted in September, and not exposed to artificial light. It is possible that germination would have been higher if they had been planted during the summer. Seeds from the *P. heterophylla* x *P. heterophylla* and *P. heterophylla* x *P. arenicola* crosses were planted 8 days later than seeds from the other crosses, which again could have lowered the amount of germination seen. The river bluff taxon x *P. arenicola* seeds were the last batch planted, potted up 12 days after the first sets of seeds, and none germinated. However, this result matches the result from another set of river bluff taxon x *P. arenicola* seeds which were planted during the summer.

Leaf shape analysis: Plotting the averaged Fourier harmonics for each species resulted in a pair of simplified leaf shapes, as fine variations in toothing of margins and base shapes cancelled each other out. The major differences between the taxa were overall blade shape and type of leaf base (Figure 3.4). The average *P. heterophylla* leaf was broadly ovate with a truncate base, while the average leaf of the river bluff taxon was lanceolate with an attenuate base. The Wilks Lambda test agreed with these results, indicating that the means of the two sets of samples were significantly different ($P<0.0005$).

Principal components analysis: The first three principal components were responsible for 85.3% of the variance, with the first component alone explaining 64.8%. A plot of these three components showed two overlapping clusters of points: a larger, diffuse cloud consisting of *P. heterophylla* samples and a smaller, denser cloud of samples from the river bluff taxon (Figure 3.5). Leaves from two individuals of *P. heterophylla* considered to be morphologically atypical and more similar in morphology to the river bluff taxon were sampled, but there was no particular pattern to their placement. Some did fall among the river bluff taxon samples, and others fell well within the cloud of typical *P. heterophylla* samples.

Allozymes: In general, *P. arenicola* and the river bluff taxon had the same sets of alleles for each enzyme. There were two common alleles for 6PGD (A-slow and B-medium), three for PGI (A-slow, B-medium, and C-fast), one for PGM (B-fast), and one for IDH (B-fast). The river bluff taxon had three unique alleles, a fast 6PGD allele (C), a slow PGM allele (A), and slow IDH allele (A). *Physalis arenicola* had two unique alleles,

a very fast PGI allele (D) and a very slow PGM allele (C).

Of the seven populations sampled for the river bluff taxon, mean heterozygosity was generally higher than what would be expected based on the Hardy-Weinberg genotype frequencies (Table 3.3), though sample sizes were small, and these differences were generally not significant. However, the number of genotypes seen typically matched the number of isolated clusters of shoots sampled. Each population sampled included several different clusters of shoots with different genotypes, and in the final analysis this may have served to average out the differences seen between clusters.

The situation was similar for the six populations of *P. arenicola* sampled (Table 3.3). Most of the *P. arenicola* populations were not very variable. All populations were fixed for IDH and all but one for PGM.

For populations of the river bluff taxon, Nei's genetic identity was, on average, 0.854, and ranged from 0.674 to 0.941. For *P. arenicola*, the average genetic identity was 0.831, and it ranged from 0.593 to 0.989. The average genetic identity between the two species was 0.583, ranging from 0.390 to 0.836. (Table 3.4).

DISCUSSION:

Species relationships: Phylogenetic techniques are often seen as the cure-all for clarifying relationships within taxonomically difficult species groups, but although the southeastern Unites States perennial species form a relatively small complex, there is no simple phylogenetic answer for the placement of the river bluff taxon. One reason for this is the rather low interspecific DNA sequence variation. The southeastern perennials

probably represent a relatively recent radiation of species, which means that even rapidly evolving regions like ITS have not had time to diverge. Conversely, the other problem encountered was too much molecular variation, in the form of significant intraspecific ITS variation, which presents a different set of challenges for accurate phylogenetic reconstruction.

ITS was the first region of DNA to be sequenced, since it is well known for its high variability between closely related species. Initial analyses indicated that sequences from the river bluff taxon were most similar to those from *P. arenicola* and *P. lanceolata.* However, as sampling increased, the pattern changed, and a sequence from one individual of *P. heterophylla* also grouped with the river bluff taxon sequences. Examination of the *P. heterophylla* sequences revealed that *P. heterophylla* had two ITS types, a rare one which is one most similar to the type found in the river bluff taxon and *P. lanceolata*, and a common one which is most similar to they type seen in *P. virginiana* (Figure 3.2).

To help clarify the situation, several samples were also sequenced for a segment of the *waxy* gene between exons eight and ten. The *waxy* gene tree strongly disagreed with the ITS gene tree, in that all of the stellate-haired taxa, including *P. arenicola*, formed a monophyletic group. This suggests incomplete lineage sorting for ITS copies among these closely related taxa. Unfortunately, the *waxy* data were not variable enough to resolve the relationship between *P. heterophylla*, *P. lanceolata*, and the river bluff taxon.

Finally, data from another small segment of *waxy*, mostly from exon 1, was gathered for several taxa. Though this data set alone was not variable enough to resolve many relationships between species, it did indicate that *P. lanceolata* was a rather divergent species, and that all of the sequences examined formed a clade. Combining the three data sets (Figure 3.3) finally provided enough information to resolve the relationship as follows: ((river bluff taxon + *P. heterophylla*) *P. lanceolata*).

An interesting question is the placement of *P. lanceolata*. Like the river bluff taxon, it is a species adapted to a very specific habitat and with a limited range. Also like the river bluff taxon, it is a species closely related to *P. heterophylla*. One could explain the close relationship between the three taxa by hypothesizing that both *P. lanceolata* and the river bluff taxon are peripheral isolates of a widespread ancestor of *P. heterophylla*, and speciated as they adapted to specific habitats. However, that does not explain why the most common ITS variant found in *P. heterophylla* is not seen in either the river bluff taxon or *P. lanceolata*. An additional complexity is that ITS is a multicopy region, and concerted evolution may have unexpected effects on which ITS copy is maintained. For example, biased gene conversion could result in rarer ITS copies being preferentially fixed (Hillis et. al., 1990; Aguilar et al., 1999).

In agreement with the combined molecular analysis, morphology and range information also suggest that *P. heterophylla* is the closest relative to the river bluff taxon. *Physalis lanceolata* is a species endemic to the sandhills of northern Georgia and the Carolinas, and is thus totally disjunct to the river bluff taxon. It has morphological features distinct from both *P. heterophylla* and the river bluff taxon, including very large

fruit and antrorse trichomes. *P. heterophylla* is a widespread species, reaching the southern extent of its range in northern Florida, and though the river bluff taxon is vegetatively distinct from most individuals of *P. heterophylla*, there are occasional atypical individuals of *P. heterophylla* which are vegetatively very similar to individuals of the river bluff taxon.

Gross morphology: Plants of the river bluff taxon have a sprawling habit, with low growing, arching branches that spread out over the ground. In the field, these plants are extremely small, rarely exceeding 1.5 dm in height and diameter. Even in cultivation, they never become especially large plants. *Physalis heterophylla* generally has an erect habit, and is one of our larger perennials, regularly surpassing 3 dm in height.

While shoots of the river bluff taxon are generally covered with 2-4 mm long, non-glandular trichomes, shoots of *P. heterophylla* are typically covered with 1-2 mm long glandular hairs, with the older stems developing a longer, wooly pubescence. However, two plants of *P. heterophylla* collected in Mississippi have the long, non-glandular pubescence typical of the river bluff taxon, and have a somewhat sprawling habit. Though they are still larger than individuals of the river bluff taxon, they are smaller than the typical individual of *P. heterophylla*.

In *P. heterophylla*, morphology is not correlated with ITS type. The one individual which had the A-type ITS (Whitson, Chapter 2), most similar to that of the river bluff taxon, was morphologically typical, being a large plant covered in glandular trichomes. The two individuals of *P. heterophylla* which were morphologically most

similar to the river bluff taxon had T-type ITS sequences, which are most common in *P. heterophylla* and are not seen at all in the river bluff taxon.

Leaf shape: Leaf shape is a character which differentiates *P. heterophylla* from the river bluff taxon (Figure 3.4). Standard *P. heterophylla* leaves are broadly ovate, with truncate to cordate bases and dentate margins. Leaves of the river bluff taxon are narrowly ovate, with truncate to attenuate bases, and sparsely dentate to entire margins. However, leaves of the morphologically atypical *P. heterophylla* individuals are similar in shape to those of the river bluff taxon, though generally larger.

The mathematical techniques used to quantify these shape differences verified what can been observed with the naked eye, but did not really provide any further information. The means of the two sets of samples were significantly different, the averaged Fourier harmonics produced shapes which agreed with the original visual observations, and the principal components analysis resulted in two, somewhat overlapping clouds of points. Thus, the leaf shapes of the two taxa are definitely useful in differentiating the plants, though they would best be used in conjunction with other characters, such as plant size and hair types.

Compatibility: All three taxa tested, *P. arenicola*, *P. heterophylla* and the river bluff taxon, appear to be self-incompatible, which is common among the species of both section *Lanceolatae* and section *Viscosae* (Menzel, 1951; Sullivan, 1984b). Within species pollinations resulted in a similar amount of fruit set for all three taxa, about 30%.

Crosses between the taxa produced more variable results. *Physalis arenicola* did not set fruit from crosses with either *P. heterophylla* or the river bluff taxon. A few fruit

resulted when *P. heterophylla* and the river bluff taxon were pollinated by *P. arenicola*, but both the number of seeds set and germination were low. This result generally agrees with the rather distant phylogenetic placement of *P. arenicola* in relation to the two other taxa.

The river bluff taxon and *P. heterophylla* proved to be highly compatible, producing fruit with relatively high seed set and seeds with high viability. All three of the molecular data sets agree that these two taxa are closely related, so their interfertility is not particularly surprising. One interesting feature of these crosses was the difference in their relative success rate. When *P. heterophylla* was pollinated with other *P. heterophylla* pollen, fruit resulted 31% of the time, whereas when it was pollinated by the river bluff taxon, fruit resulted about 14% of the time. On the other hand, the river bluff taxon made fruit 33% of the time when given other river bluff taxon pollen, but produced fruit 54% of the time when it was pollinated by *P. heterophylla*.

Of the three taxa surveyed, the river bluff taxon was the one most likely to attempt to accept any pollen given – a surprising result because these were plants most strongly affected by the downy mildew. In *Physalis*, pollination of a flower results in calyx enlargement first, and then the ovary begins to enlarge soon after. A comparison of the percentage of crosses which resulted in some calyx enlargement shows that pollination of the river bluff taxon typically resulted in more buds showing some calyx enlargement than in either of the other species, regardless of the type of cross being done. For example, when the plants were self-pollinated, none of the *P. arenicola* pollinations showed any calyx enlargement before the bud was dropped, but 33% of those on the river

bluff taxon did. Plants of the river bluff taxon generally inhabit shady, woodland areas, which may reduce their opportunities to sexually reproduce. If this is the case, perhaps selection has favored the ability to accept whatever non-self pollen is presented.

Allozymes: In the river bluff taxon, the genotype of a shoot was often identical to the genotypes of other shoots sampled from the same cluster, even in cases where individuals were heterozygous at one or more loci, which suggests that these plants are often clonal. Though the plants are rhizomatous, this result was somewhat surprising. Unearthing the rhizomes of these plants generally shows that one shoot has a single rhizome which terminates 4-6" under the soil surface, and does not appear to connect with other rhizomes. It is likely that the connections between rhizomes between shoots are short-lived, resulting in clusters of isolated, but genetically identical, shoots. These plants require disturbed habitats to grow, so the ability to reproduce vegetatively would allow fast colonization of newly disturbed areas, such as gaps or eroding river bluffs. Populations of *P. arenicola* also proved to be highly clonal, though this was expected, since the long, running rhizomes can often traced to several different shoots.

There was no evidence that the river bluff taxon and *P. arenicola* were hybridizing, which agrees with the fact that crosses between the two taxa were rarely successful. Populations of both the river bluff taxon and *P. arenicola* tended to be genetically more similar to each other than to populations of the other species (Table 3.2), even when collected from areas where both species occurred. While the average within species genetic identity was over 80% for both taxa, the average between species genetic

identity was 58%. Though both species shared similar sets of alleles, each species had at least two unique alleles.

Status of the river bluff taxon: Considering all of the data, it is unlikely that the river bluff taxon is a hybrid between *P. arenicola* and *P. heterophylla*. Populations of the river bluff taxon were not highly morphologically variable, as might be expected among backcrossing hybrid plants. No individuals intermediate between *P. arenicola* and the river bluff taxon were seen in the field, even though mixed populations were found, and no Floridian individuals of *P. heterophylla* were seen at all.

The DNA results, though made more difficult to interpret by the ITS variation, showed that *P. arenicola* is distinct and monophyletic for the ITS sequences examined. This species is a member of the stellate-haired clade, and not closely related to either the river bluff taxon or *P. heterophylla* (Figure 3.3). The river bluff taxon and *P. heterophylla* appear to be sister taxa. All of the PCR products generated for the river bluff taxon were direct sequenced, and there was no evidence that the river bluff taxon was either polymorphic for *P. heterophylla* and/or *P. arenicola* sequence types, or that it had recombinant sequence types.

Crossing studies proved that while it is possible to cross *P. arenicola* to both *P. heterophylla* and the river bluff taxon, the crosses only work in one direction, fruit set is rare, and seed set and seed viability are low. All of this combined with the fact that the plants generally prefer different habitats means that the chance of such an event occurring in the wild and resulting in a hybrid plant is extremely unlikely.

Though the allozyme data was most useful in showing that populations of the river bluff taxon are largely clonal, it did also show that these plants share several unique alleles that were not found in *P. arenicola*, even when the plants sampled were from mixed populations. This supports the argument that the river bluff taxon and *P. arenicola* are not hybridizing.

Relation to *P. heterophylla*: Morphology, crossing studies and phylogenetic information all make it clear that the river bluff taxon is closely related to *P. heterophylla*. The question of at which taxonomic rank to recognize the river bluff taxon is perhaps the most difficult to address with the techniques at hand. However, current species concepts in *Physalis* include both species which hybridize naturally and species which are more morphologically similar than this pair of taxa.

The phylogeny of *Physalis* sections *Viscosae* and *Lanceolatae* demonstrates that all of these species are closely related. Crossing studies done by various workers have also shown that most of these taxa are at least somewhat compatible in the lab, though generally, they don't seem to cross in the wild (Menzel, 1951; Hinton, 1975 and 1976; Sullivan, 1984a). However, in section *Viscosae*, *P. walteri* and *P. angustifolia* are notorious for their complete compatibility in the field, and plants collected in Florida, where the ranged of the two species overlap, often have morphology intermediate between the two very distinct parental species (Menzel, 1951; Sullivan, 1984b). In section *Lanceolatae*, *P. heterophylla* and *P. virginiana* have also been shown to occasionally hybridize in the wild (Hinton, 1975), though the two species remain morphologically distinct, and morphologically intermediate individuals are rare.

The woodlands along the Apalachicola river bluffs have been characterized as 'relict forests' (Platt and Schwartz,1990), containing many northern taxa which are disjunct from the nearest Appalachian populations (Parks et al., 1994). Endemic taxa found in these areas are often explained as populations of northern temperate taxa which were forced south by glaciation and never moved north again when the glaciers retreated, thus leaving them isolated from the rest of their species and allowing them to speciate as they adapted to the warmer Floridian climate. However, this situation is hard to distinguish from another scenario: that some of the endemics were actually distinct when they arrived in Florida, and that they adapted to the river bluff habitat during glacial maxima, and then were unable to later recolonize their former habitats. If the river bluff taxon was distinct from *P. heterophylla* prior to its arrival, it might explain why a species with a small range and relatively small populations sizes has fixed an ITS type which is quite rare in one of its closest relatives.

Morphologically, the river bluff taxon is very distinct from most individuals of *P. heterophylla*. It is similar to the occasional, morphologically unusual individual of *P. heterophylla*, but even these plants are not identical to individuals of the river bluff taxon. They remain larger, even when grown under the same greenhouse conditions. Also, the one *P. heterophylla* which was vegetatively most similar to the river bluff taxon had flowers with darkly feathered maculations which were very unlike anything seen in the river bluff taxon.

CONCLUSIONS:

The river bluff taxon does not appear to be a hybrid between *P. arenicola* and *P. heterophylla*, nor does it seem to be hybridizing with *P. arenicola*. Though it is closely related to *P. heterophylla* and to *P. lanceolata*, it is as distinct as many other *Physalis* species. It can be successfully crossed with *P. heterophylla*, but no individuals of *P. heterophylla* were seen along the river bluffs, or even in northern Florida, so the plants do not have the opportunity to cross in the wild. Morphologically, the two taxa are generally easy to distinguish, and plants of *P. heterophylla* are always larger than those of the river bluff taxon. In conclusion, I propose to recognize the river bluff taxon as a new species, and to name it *Physalis diminuta*, due to the extremely small stature of these plants in the field.

Chapter 4.

A new species of *Physalis* (Solanaceae) from the Apalachicola river bluffs of northern Florida

CURRENTLY, *PHYSALIS DIMINUTA* IS A PROVISIONAL NAME, AND PUBLICATION OF THE SPECIES DESCRIPTION IN THIS THESIS IS NOT MEANT TO CONSTITUTE PUBLICATION OF THE SPECIES NAME.

Physalis diminuta Whitson, sp. nov. Figure 4.1

P. heterophyllae Nees affinis sed parvior, 1-2 dm alta, pubescentia hirsuta, pilis simplicibus, eglandulatis. Folia anguste ovata, basi truncata vel attenuata.

Perennial herbs, generally 1.5 dm tall and wide, reaching 2 dm tall by 2.5 dm wide under exceptionally good growing conditions (cultivation). Rhizome extending vertically down into the soil, stouter than the aerial stems, reaching 5 mm in diameter, brittle and easily broken. Most vegetative growth is in the form of new shoots emerging from the ground. Stems unbranched to sparsely branched, 2-3 times, spreading out over the ground. Unbranched stems have a zigzag morphology, with the leaves paired at oblique angles. Internodes short, generally (1.0) 1.5-3.0 (4.5) cm long, stems to 4 mm wide. Stems hirsute to hirtellous with simple, non-glandular, divergent hairs, 1-4 mm

long. Most individuals have occasional hairs with uneven branching at the tip. Rarely, a plant will have mostly branched hairs.

Leaf blades lanceolate to ovate, (2.0) 4.0-7.0 (9.3) cm long, (1.6) 2.0-3.0 (3.6) cm wide. Adaxial surfaces hirtellous with non-glandular hairs 1-2 mm long, abaxial surfaces with hairs mostly along the veins. Apices acute to acuminate, margins entire to undulate or sparsely, often unequally, 1-5 dentate. Bases typically attenuate or truncate with a winged petiole and often unequal. Petioles hirsute, generally 1/5 to 1/3 of blade length.

Flowers solitary in the leaf axils. Pedicels (5) 10-18 mm long, hirsute and non-glandular. Flowering calyces 8-10 mm long, 5-9 mm wide, hirtellous and non-glandular, lobes acute to slightly acuminate, 3-4 mm long. Corollas campanulate, yellow, (1.6) 2.0-2.5 cm wide, with 5, olive green to blackish-brown, more-or-less feathered spots at the base. Anthers yellow or tinged with purple, 2-3 mm long, 1 mm wide. Filaments purple, clavate at the apex, about 5 mm long, no wider than the anthers. Filaments elongate and anthers open over several days, in a 2+2+1 series. Gynoecium bicarpellate, ovary rounded, style 8-12 mm long, excerted 2-4 mm beyond the stamens, purple, stigma flat, slightly bi-lobed, bright yellow-green. Fruiting pedicels 2.0-2.5 cm long. Fruiting calyx 10-angled, 2.0-2.5 cm long by 1.0-2.0 cm wide, yellowing at maturity. Berry round, ca. 1 cm diameter, greenish-yellow to yellow, viscid and sweet smelling, with 20-50 (100) seeds. Seeds reniform, tan to yellow, 1-1.5 mm long.

Blooming from late spring to fall. Disturbed woodland habitats near the Apalachicola and Chipola river bluffs: gaps, trailsides, washing bluffs, old roadbeds.

Calcareous or clayey soils. Liberty, Jackson and Gadsden counties, Florida. Decatur county, Georgia.

Type: Cultivated material, 08/17/2000, Maggie Whitson, 1203a. Living plant originally collected from: U.S.A., Florida, Liberty County, Torreya State Park, River Trail, site 1: wooded trailside along the bluffs of the Apalachicola river. (Holotype: DUKE).

ADDITIONAL SPECIMENS EXAMINED: All specimens are deposited in the Duke University herbarium (DUKE). Florida Caverns State Park, Liberty Co., FL: *Whitson* 1239, 1240, 1241, 1275. Three Rivers State Recreation Area, Jackson Co., FL: *Whitson* 2072, 1250, 1249, 1248. Torreya State Park, Liberty Co. FL: *Whitson* 1203a, 1276, 1268.

Physalis diminuta is a member of *Physalis* subgenus *Rydbergis* (Hendrych, 1989) section *Lanceolatae* (Rydb.) Menzel (Menzel, 1951). The plants occur along the bluffs of the Apalachicola river and its tributary, the Chipola river, in Jackson and Liberty counties, Florida. A new population was recently discovered along Lake Seminole (part of the Apalachicola river) in Decatur county, Georgia.

This species is closely related to *P. heterophylla* Nees., but can be distinguished on the basis of its very small stature, non-glandular pubescence, and narrowly ovate leaves. The divergent trichomes seen in *P. diminuta* also distinguish it from *P. longifolia* Nees. (antrorse trichomes) and *P. virginiana* Miller (retrorse trichomes) (Sullivan, 1984).

The stocky, non-running rhizomes of *P. diminuta* distinguish it from *P. arenicola* Kearney, which has running rhizomes and is very common in north Florida.

Chapter 5.

Proposal to conserve the name *Physalis* L. (Solanaceae) with a conserved type

Physalis L., Sp. Pl.: 183. (1753)

Type: *Physalis pubescens* L., *typ. cons. prop.*

Physalis is a genus of 75+ species of herbs almost entirely limited to the New World. The one exception is *P. alkekengi* L., a cultivated plant naturalized widely in Europe, but probably native to China. *Physalis alkekengi* L. was designated as the type of *Physalis* L. by Britton and Brown in 1913. Though this typification was made under the American Code, it has remained uncontroversial, being accepted by Hitchcock and Green (1929) and later by the Linnaean generic types project (Jarvis et al., 1993).

Though *P. alkekengi* was probably well known to Linnaeus, and fits the generic protologue, it is not morphologically representative of the genus. *Physalis alkekengi* possesses two major features that make it unique within *Physalis*. This species has somewhat lobed, white flowers in a genus otherwise characterized by unlobed, yellow corollas, and its fruit and fruiting calyces are a brilliant red-orange. Though a few species of *Physalis* have tangerine orange calyces, none approach red. Noting its morphological distinctiveness, Rydberg, who monographed the genus in 1896, stated, "If...*P. alkekengi* could be also removed, the genus would be a very natural one." It is ironic to note that he wrote the original treatment for Britton and Brown (1898).

Within the last decade, cpDNA phylogenies focusing on related genera (Mione et. al., 1994; Olmstead et al., 1999) have suggested that *P. alkekengi* may be divergent from the rest of *Physalis*. Now, a two-gene phylogeny of the subtribe Physalinae (Whitson, Chapter1) has shown conclusively that the morphologically typical New World species of *Physalis* form a strongly supported group, of which *P. alkekengi* is not a member.

As *P. alkekengi* fits the generic protologue, and has been accepted as the type of *Physalis* by workers after Britton and Brown, its supersession under article 10.5 of the present Code (Greuter, 2000) is not possible. Several other nomenclatural options are available, but all have some undesirable consequences. Providing a new generic name for the New World taxa would require the transfer of over 70 species, would outdate multiple floras, and would have a profound taxonomic effect on New World botany, as *Physalis* is one of the larger genera in the Solanaceae, and makes up a significant part of the Solanaceous flora in the U.S., Mexico, and Central America.

The two-gene phylogeny indicates that several small genera fall between *P. alkekengi* and the New World *Physalis* species. These genera could all be transferred to *Physalis*, making the genus monophyletic. However, this solution would require doing away with some genera which have long been considered distinct from *Physalis*, such as *Chamaesaracha* and *Oryctes*. The enlarged *Physalis* would be morphologically heterogeneous, defined by few characters other than some amount of calyx enlargement in fruit.

The least disruptive solution is to invoke article 56.1 of the Code (Greuter, 2000), conserve *Physalis* using a Linnaean New World species as the type, and place *P.*

alkekengi in its own genus -- a strategy which has been successfully used to solve similar problems in *Petunia* A. L. Jussieu and *Chrysanthemum* L (Wijnands and Bos, 1986; Trehane, 1995). Though Linnaeus named nine species when he described *Physalis*, several points argue for selecting *P. pubescens* L. as the conserved type. It is a well-known, morphologically typical, widespread species and has been typified with a Linnaean specimen (LINN 247.11).

Retypifying *Physalis* will have the undesirable effect of outdating many floras and re-naming a horticulturally important species, as *P. alkekengi*, the Chinese lantern plant, is grown for its showy red-orange fruiting calyces. However, there are economically important species among the New World taxa which would otherwise have to be transferred, such as *P. philadelphica* Lam. (the tomatillo) and *P. peruviana* L. (the cape gooseberry). Additionally, any of the weedy New World annuals which are included in Old World floras would still have the correct names. One advantage of placing *P. alkekengi* in a new, monotypic genus is that it will emphasize the distinctiveness of the single, Old World member of the Physalinae.

Though multiple superfluous names exist for *Physalis*, *P. alkekengi* (typified by LINN 247.5) has never been placed in its own genus. Therefore, given that *P. alkekengi* is rejected as the type of *Physalis*, the name *Sinophysalis* Whitson, gen. nov., is provided for it, with *Sinophysalis alkekengi* (L.) Whitson being the type and only species. This genus may be diagnosed as follows: *A Physalem L. corallis albis leviter lobatis, calycibus rubroaurantiacis conspicuis baccas rubras maturitate includentibus differt. Species typica et singularis, S. alkekengi, gerontogaea est.*

Considering all of the options, conserving *Physalis* L. with *P. pubescens* L. as the conserved type is the best way to preserve nomenclatural stability, requiring only that a single species be renamed.

Chapter 6.

Conclusion

Though *Physalis* is a morphologically distinctive genus, defining its taxonomic limits has long been a challenge, because the unusual, expanded fruiting calyx characteristic of *Physalis* also occurs in several other small genera, called physaloids. However, phylogenetic analysis of molecular data, in the form of DNA sequences from the ITS region and the nuclear gene *waxy*, has definitively solved this problem. The morphologically typical, New World species of *Physalis* form a strongly supported, monophyletic group, which includes only one of the traditionally recognized physaloid genera, the monotypic *Margaranthus*. This taxon has already been transferred to *Physalis*, and should be recognized as *P. solanaceous* (Schlecht.) Axelius. The other North American physaloid genera, including *Quincula*, *Chamaesaracha*, *Leucophysalis* and *Oryctes*, were all distinct from *Physalis*, and thus merit continued recognition at the genus level.

One surprising result of this study was that several morphologically unusual species of *Physalis* were also distinct from the main clade of *Physalis* species. These taxa included *P. arborescens*, *P. melanocystis*, *P. carpenteri*, *P. microphysa*, and *P. alkekengi*, the type species and only Eurasian-native member of *Physalis*. Removing these taxa from *Physalis* and recognizing them as monotypic or small genera would result in a morphologically homogenous and monophyletic *Physalis*, but this is not an option in the

case of *P. alkekengi*, since it is the generic type. However, this nomenclatural problem can be dealt with by conserving *Physalis* with a conserved type. Consequently, a proposal was made to conserve *Physalis* with *P. pubescens* L. as the type, and to provide a new generic name, *Sinophysalis*, for *P. alkekengi*.

While ITS sequence data was useful in clarifying the relationships between the physaloid genera, examination of patterns of ITS variation within a closely related complex of U.S. perennial *Physalis* species indicates that between these very closely related taxa ITS gene trees do not necessarily reflect species relationships. One species, *P. heterophylla*, appeared to have at least two ITS types, one of which was more similar to that of *P. virginiana*, and one of which was more similar the ITS sequences of *P. lanceolata*. In addition, some individuals of *P. heterophylla* had ITS sequences nearly identical to those of *P. virginiana*, a species with which it is known to hybridize. This suggests either hybridization followed by biased gene conversion or the presence of a third ITS type in *P. heterophylla*. Thus, incomplete lineage sorting and putative hybridization are forces acting to obscure the phylogenetic relationships estimated by ITS gene trees.

Secondary structure of the ITS regions also appears to have an impact on patterns of sequence variation. In ITS2, base substitutions were most common in areas which were single-stranded in the putative secondary structures, and less common in double-stranded regions. Indels were found to occur at the single-stranded tips of loop regions.

Careful study of the U.S. perennial *Physalis* resulted in the discovery of several unusual populations of an unidentified perennial taxon along the Apalachicola river

bluffs in northern Florida. A variety of techniques, including phylogenetic analysis of DNA sequences, elliptic Fourier analysis of leaf shapes, and crossing studies, were used to determine whether these populations of plants were of hybrid origin, or represented a new species. *Physalis arenicola* is one of the most common perennial species in northern Florida, and can be found growing intermixed with populations of the river bluff taxon, but crosses between these two taxa only rarely produce seeds, and the molecular data did not support a close relationship between the two taxa. There was no evidence that the river bluff taxon was naturally hybridizing with *P. arenicola*, or that it had arisen as a hybrid between *P. arenicola* and one of the other perennial taxa. The phylogenetic analyses indicated that *P. heterophylla*, a species more common in the northeastern U.S., is probably the sister species of the river bluff taxon. Crosses between these two taxa often resulted in seeds. Leaf shape analysis indicated that there were differences between the leaf shapes of *P. heterophylla* and the river bluff taxon, in that the river bluff taxon generally had narrower leaves, with fewer teeth and attenuate or truncate leaf bases, as opposed to the more cordate leaf bases in *P. heterophylla*. Other morphologically distinctive features of the river bluff taxon include its strictly non-glandular pubescence and very small stature. Because the river bluff taxon could be distinguished from *P. heterophylla* both on the basis of molecules and morphology, and because these two taxa do not grow in the same areas, and thus do not naturally hybridize, it was determined that the river bluff taxon was a new species. The name *P. diminuta* was proposed, and a formal species description and photos of the type specimen were provided.

Table 1.1. Genera of the Physalinae and Infrageneric Classification of *Physalis*.

The Physalinae:

Brachistus (3 spp., Central America)
Chamaesaracha (10 spp., SW U.S. to Central America)
Leucophysalis (2 sp. U.S., 1+ sp. Central America)
Margaranthus (1 sp., SW U.S. to Central America)
Oryctes (1 sp., California and Nevada)
Quincula (1 sp., SW U.S. and N Mexico)
Tzeltalia * (2 spp., Guatemala and S Mexico)
Witheringia (20 sp., Central and South America)

* *Tzeltalia* was segregated from *Physalis* after Olmstead et al. (1999) provided a classification for the Physalinae. However, it can be placed in this subtribe on morphological grounds.

Infrageneric Classification of *Physalis*:

Physalis subgenus *Physalis* (1 sp., China)
    P. alkekengi (Type)
Physalis subgenus *Physalodendron* (2 spp., Mexico and Central America)
    P. arborescens, *P. melanocystis*
Physalis subgenus *Quincula* (1 sp., SW U.S. and N Mexico)
    P. (=*Quincula*) *lobata*
Physalis subgenus *Rydbergis* (60+ spp., 9 sections, New World, mostly Mexico)
    section *Epeteiorhiza* (14 spp.)
    section *Campanulae* (2 spp.)
    section *Coztomatae* (11 spp.)
    section *Lanceolatae* (14 spp.)
    section *Carpenteriae* (1 sp.)
    section *Angulatae* (10 spp.) (includes *Margaranthus)*
    section *Viscosae* (6 spp.)
    section *Rydbergae* (2+ spp.)
    section *Tehuacanae* (1 sp.)

Table 1.2. Voucher information for taxa from which DNA was extracted. Underlined taxa were sequenced for both ITS and *waxy* , other taxa were sequenced only for ITS.

Taxon:	Voucher:	Collection Information:
Brachistus stramonifolius	L. Williams 41524, DUKE	Solola and Chimaltenango, Guatemala.
Brachistus stramonifolius	Cochrane 2018, DUKE	Jalisco, Mexico.
Chamaesaracha comopus	B. L. Turner 15854, TEX	Native to the SW U.S. and N Mexico.
Chamaesaracha sordida	R. G. Olmstead s.n., WTU	Native to the SW U.S. and N Mexico.
Chamaesaracha sordida	Turner 97-0413, LL-TEX	Crockett Co., TX, USA.
Leucophysalis grandiflora	R. G. Olmstead S-30, WTU	USA.
Leucophysalis nana	Bartholomew 5994, MO	Modoc Co., CA, USA.
Leucophysalis nana	M. Williams, 82-108-1, MO	Douglas Co., NV, USA.
Leucophysalis viscosa	Torres 7932, MO	Oaxaca, Mexico.
Margaranthus solanaceous	R. G. Olmstead S-37, WTU	Native to the SW U.S. and Mexico.
Oryctes nevadensis	Tiehm 11982, LL-TEX	Churchill Co., NV, USA.
Physalis acutifolia	Nijmegen Accession # 974750059	SW USA.
Physalis aff. heterophylla	M. K. Whitson s.n., DUKE	Liberty Co., FL, USA.
Physalis aff. heterophylla	M. K. Whitson s.n., DUKE	Liberty Co., FL, USA.
Physalis alkekengi	M. K. Whitson 1280, DUKE	Cultivated source.
Physalis alkekengi	M. K. Whitson 1283, DUKE	Cultivated source.
Physalis angulata	J. Horn 1284, DUKE	Worth Co., GA, USA.
Physalis angustifolia	M. K. Whitson, no voucher	N Florida, USA.
Physalis angustiphysa	Ton 9286, LL-TEX	Chiapas, Mexico.
Physalis arborescens	Jimenez 454, LL-TEX	Tamaulipas, Mexico.
Physalis arborescens	Nee 28700, MO	Veracruz, Mexico.
Physalis arenicola	M. K. Whitson 987, DUKE	Polk Co., FL, USA.
Physalis arenicola	M. K. Whitson, no voucher	Florida, USA.
Physalis campanulata	Ventura 4882, MO	Veracruz, Mexico.
Physalis carpenteri	M. K. Whitson 1133, DUKE	Suwanee Co., FL, USA.
Physalis carpenteri	W. J. Dunn 201, FLAS	Alachua Co., Florida, USA.
Physalis caudella	Quintana 3075, TEX	Chihuahua, Mexico.
Physalis chenopodifolia	M. K. Whitson 1287, DUKE	Cultivated source.
Physalis chenopodifolia	Ventura 4402, TEX	Mexico, Mexico.
Physalis cinerascens	M. K. Whitson, no voucher	Kaufman Co., TX, USA.
Physalis cinerascens	M. K. Whitson, no voucher	Kaufman Co., TX, USA.
Physalis cordata	M. K. Whitson s.n., DUKE	Gadsden Co., FL, USA.
Physalis coztomatl	Ventura 1006, MO	D.F., Mexico.
Physalis coztomatl	Garcia 264, MO	Mexico, Mexico.
Physalis crassifolia	Richmond, no voucher	California, USA.
Physalis crassifolia	Panero 2824, MO	Baja California Norte, Mexico.
Physalis glutinosa	Sikes 375, TEX	Durango, Mexico.
Physalis greenmanii	Nee 22432, MO	Veracruz, Mexico.
Physalis hederifolia	Van Devender 85-36, LL-TEX	Brewster Co., TX, USA.
Physalis hederifolia var. *puberula*	Henrickson 5869, TEX	Chihuahua, Mexico.
Physalis heterophylla	M. K. Whitson, no voucher	Caswell Co., NC, USA.
Physalis hintonii	Villarreal 4909, MO	Nuevo Leon, Mexico.
Physalis hintonii	Luckow 3050, NCU	Veracruz, Mexico.
Physalis ignota	Breedlove 52891, MO	Chiapas, Mexico.
Physalis lagascae	Flores 1810, MO	Nayarit, Mexico.
Physalis lanceolata	J. Horn 1133, DUKE	Scotland Co., NC, USA.
Physalis lassa	Sanders 11807, MO	Comala, Mexico.
Physalis longifolia	Mona Whitson s.n., DUKE 358627	Riley Co., KS, USA.
Physalis longifolia	M. K. Whitson s.n., DUKE	USA.
Physalis melanocystis	M. Martinez 1940, MO	Tamaulipas, Mexico.
Physalis microcarpa	Lafernere 1661, MO	Chihuahua, Mexico.
Physalis microphysa	Henrickson 11850, TEX	Coahuila, Mexico.
Physalis minima	Nijmegen 974750167	Seeds from Thailand.
Physalis minimaculata	Torres 1595, TEX	Michoacan, Mexico.
Physalis minimaculata	Mayfield 986, TEX	Oaxaca, Mexico.
Physalis mollis	M. K. Whitson s.n., DUKE	Van Zandt Co., TX, USA.
Physalis nicandroides	L.G. Hernandez 2488, MO	Morelos, Mexico.
Physalis patula	Nee 32810, MO	Veracruz, Mexico.
Physalis peruviana	N. Pitman, no voucher	Cultivated, Ecuador.
Physalis philadelphica	M. K. Whitson s.n., DUKE	Cultivated source. Native to Mexico.
Physalis philadelphica	Nijmegen 894750257	Cultivated source. Native to Mexico.
Physalis pruinosa	Nijmegen 894750256	Cultivated source. Native to Mexico.
Physalis pubescens	M. K. Whitson 3, DUKE	Seeds from La Selva Biological Station, Costa Rica.

Table 1.2 continued. Voucher information for taxa from which DNA was extracted. Underlined taxa were sequenced for both ITS and *waxy* , other taxa were sequenced only for ITS.

Taxon:	Voucher:	Collection information:
Physalis pubescens	M. K. Whitson 3, DUKE	Seeds from La Selva Biological Station, Costa Rica.
Physalis pumila	M. K. Whitson s.n., DUKE	Van Zandt Co., TX, USA.
Physalis sordida	Hinton 18464, TEX	Nuevo Leon, Mexico.
Physalis virginiana	M. K. Whitson, no voucher	Orange Co., NC, USA.
Physalis virginiana	M. K. Whitson, no voucher	Orange Co., NC, USA.
Physalis viscosa	M. K. Whitson 1282, DUKE	Cultivated source. Native to Mexico and South America.
Physalis walteri	M. K. Whitson, no voucher	N Florida, USA.
Quincula lobata	R. G. Olmstead 93-74, WTU	Native to the SW U.S. and N Mexico.
Tzeltalia amphitricha	E. Martinez 20523, LL-TEX	Chiapas, Mexico.
Tzeltalia calidana	Lundell 19625, LL-TEX	Baja Verapaz, Guatemala.
Tzeltalia calidana	Matuda 5199, LL-TEX	Chiapas, Mexico.
Witheringia macrantha	Bohs 2512, UT	Monteverde, Costa Rica.
Witheringia meiantha	Bohs s.n., UT	Central America.
Witheringia solanacea	Bohs 2427, UT	Central America.

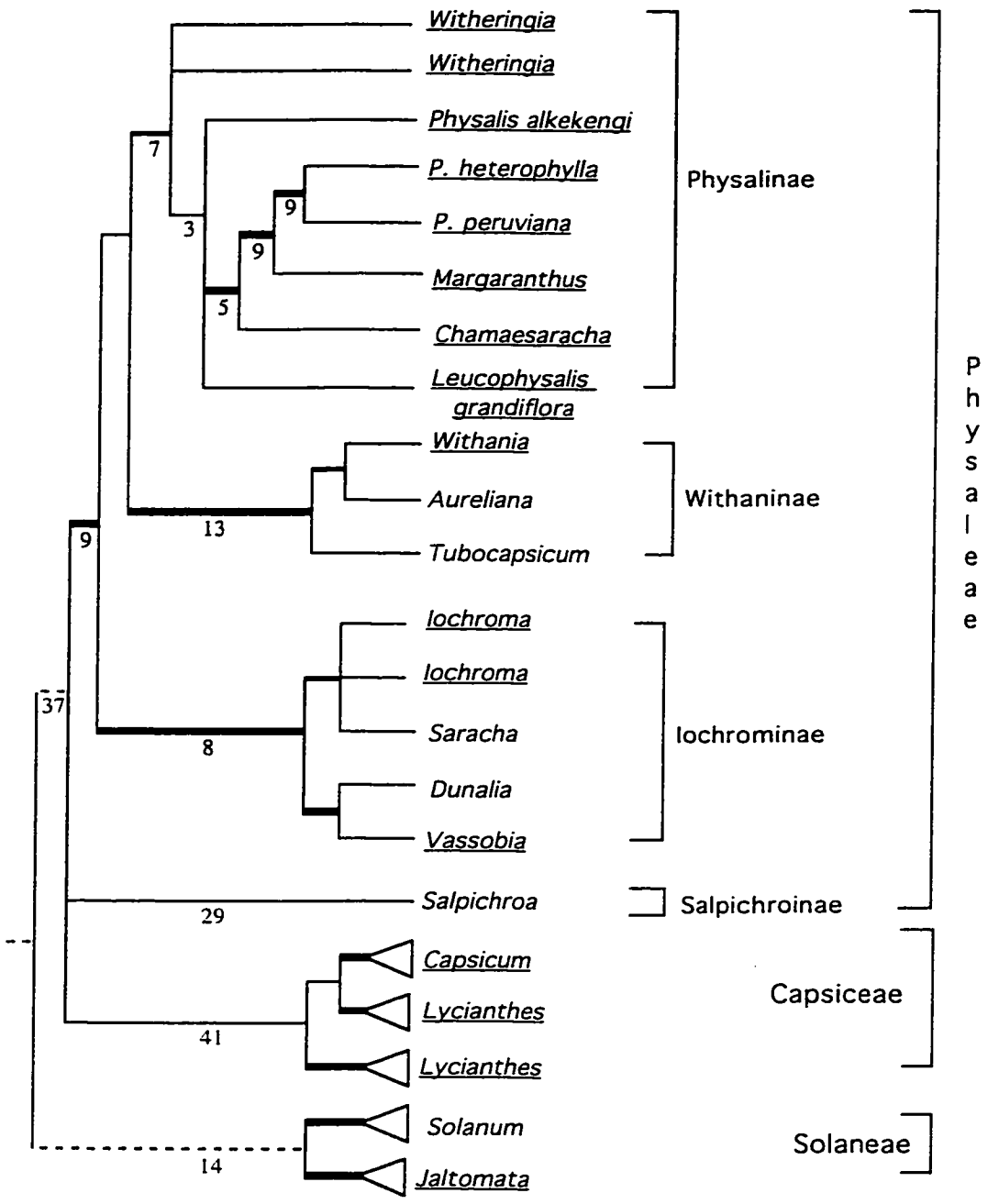

Figure 1.1. Chloroplast DNA phylogeny of the tribe Physaleae (Olmstead et al.,1999). Taxa which were examined as potential outgroups for the nuclear phylogeny of the Physalinae are underlined. Branch lengths appear below the branches. Bootstrap support indicated by branch width: heavy, 96-100%, midwidth, 72-84%, narrow 52-63%. Dashed branches have less than 50% bootstrap support.

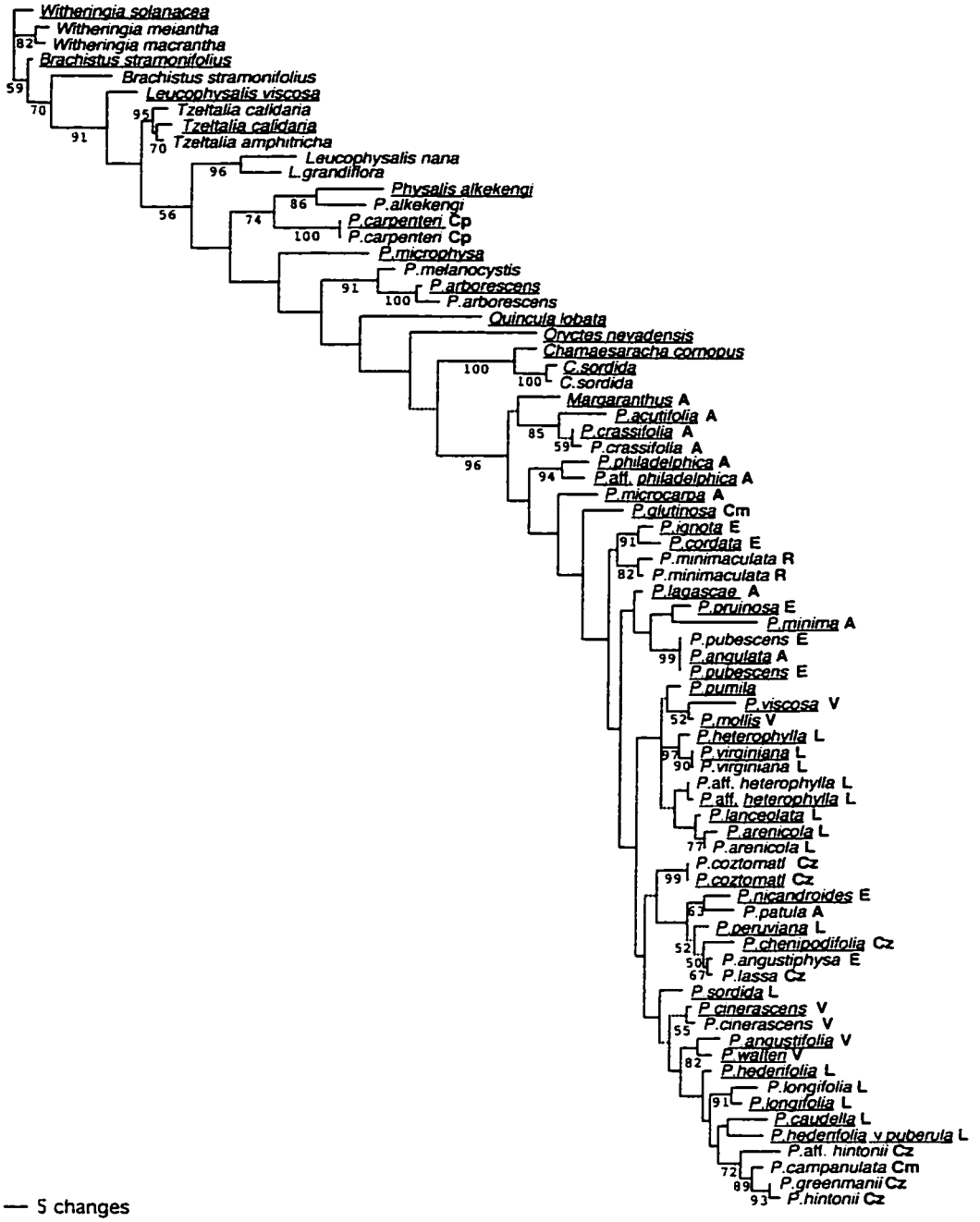

Figure 1.2. One of 6691 most parsimonious trees from the 75-taxon ITS data set. Length = 784 steps, CI = 0.4439, HI = 0.5561. Fast bootstrap values appear below each branch. Dashed branches collapse in the strict consensus. Underlined taxa were used in the two-gene analysis. Letters after taxon names represent sections of *Physalis* subgenus *Rydbergis*: A=*Angulatae*, Cm=*Campanulatae*, Cp=*Carpenteriae*, Cz=*Coztomatae*, E=*Epeteiorhiza*, L=*Lanceolatae*, V=*Viscosae*.

Figure 1.3. One of 220 most parsimonious trees based on the 50-taxon ITS data set. Length = 675 steps, CI = 0.4652, HI = 0.5348. Boostrap values from a 100 replicate heuristic search in bold face type. Fast bootstrap values appear below the branches. Dashed branches collapse in the strict consensus. Letters after taxon names represent sections of *Physalis* subgenus *Rydbergis*: A=*Angulatae*, Cm=*Campanulatae*, Cp=*Carpenteriae*, Cz=*Coztomatae*, E=*Epeteiorhiza*, L=*Lanceolatae*, V=*Viscosae*.

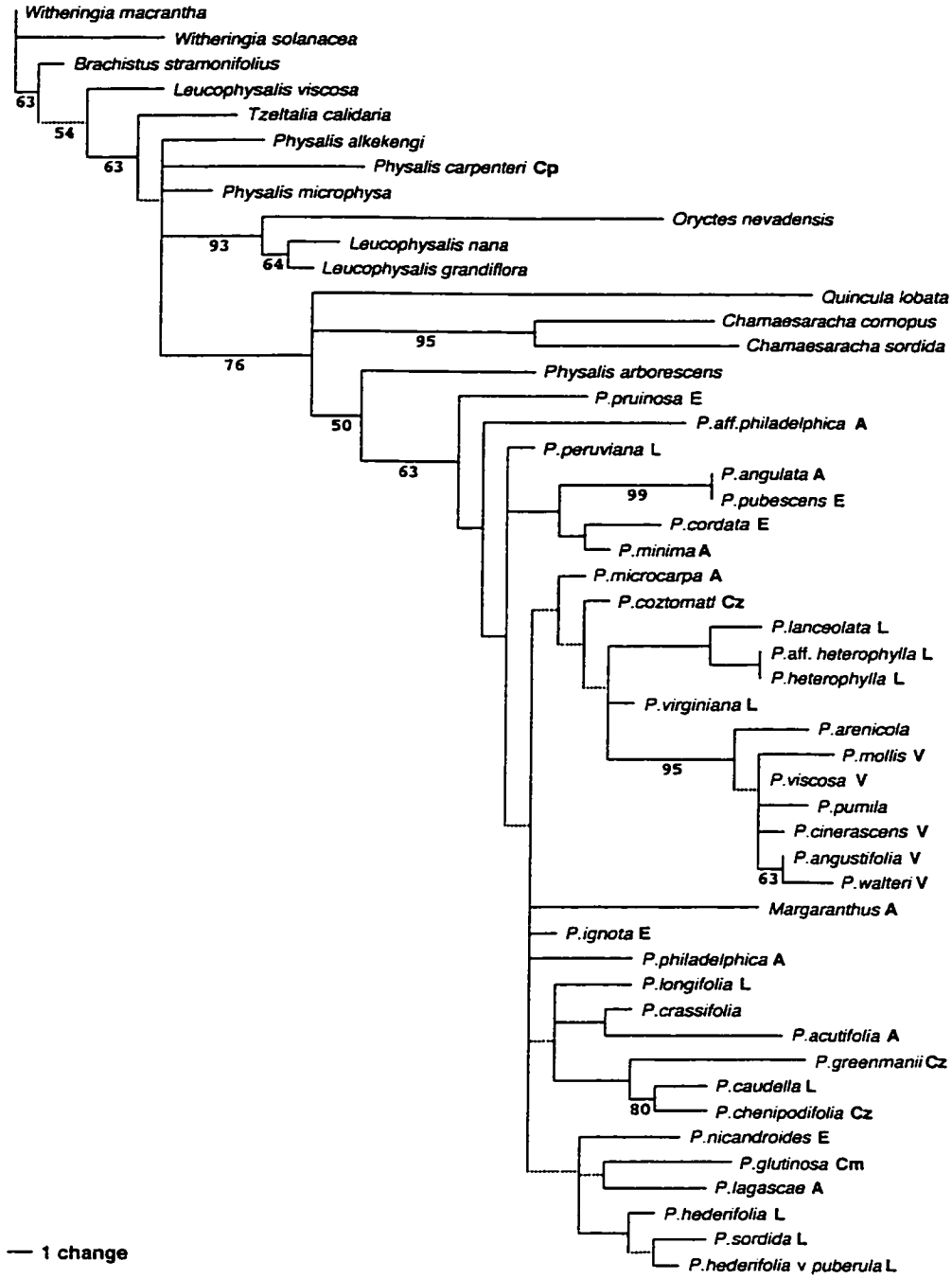

Figure 1.4. One of 85,000+ most parsimonious *waxy* trees. Length = 249 steps, CI = 0.7269, HI = 0.2731. Fast bootstrap values from a 10,000 replicate run appear below the branches. Dashed branches collapse in the strict consensus. Letters after taxon names represent sections of *Physalis* subgenus *Rydbergis*: A=*Angulatae*, Cm=*Campanulatae*, Cp=*Carpenteriae*, Cz=*Coztomatae*, E=*Epeteiorhiza*, L=*Lanceolatae*, V=*Viscosae*.

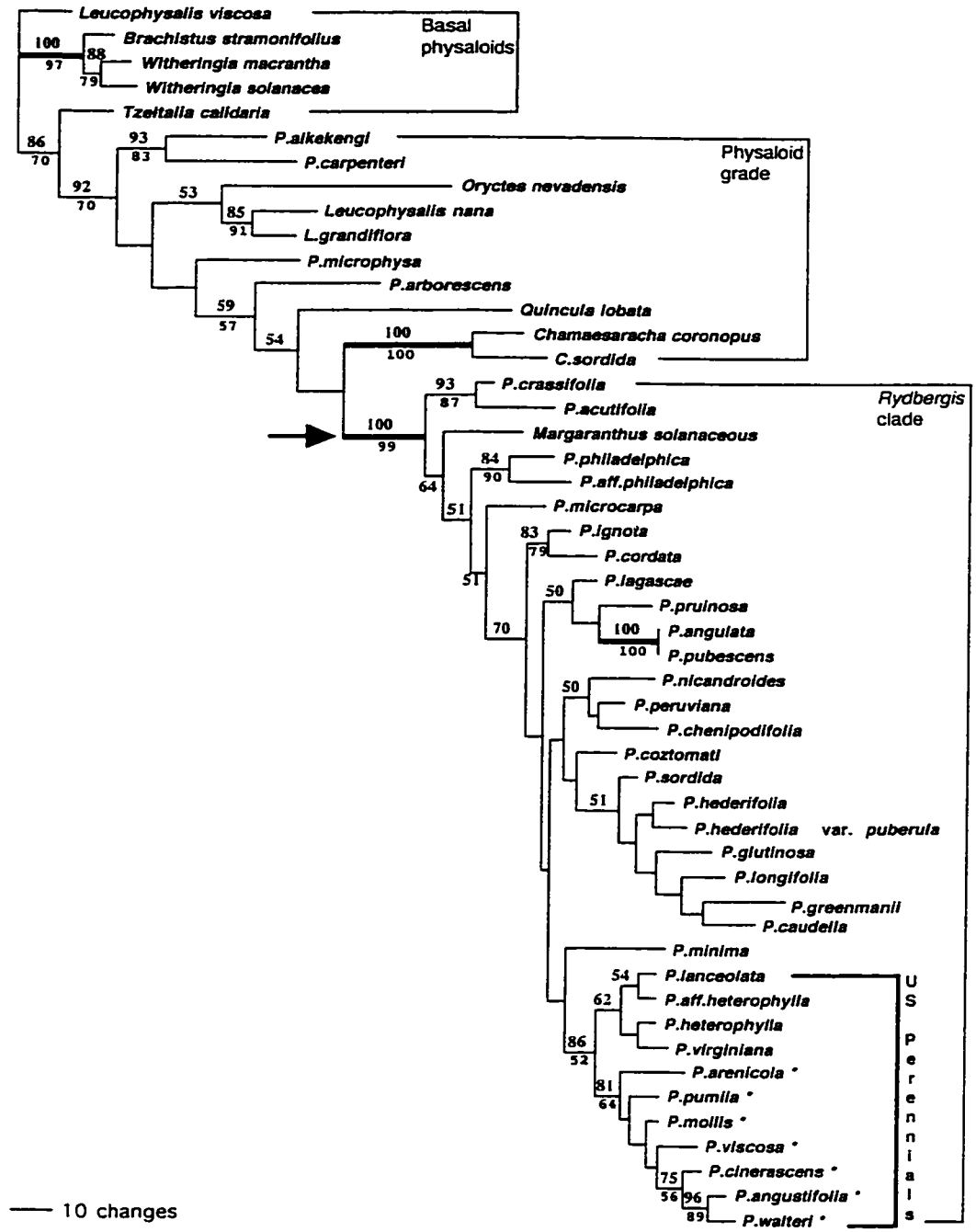

Figure 1.5. Two-gene phylogeny of the Physalinae. One of 170 most parsimonious trees, 966 steps long. Bootstrap values appear in boldface type above the branches, fast bootstrap values in plain text below. Arrow indicates the New World *Physalis* clade, which consists of morphologically typical species and is generally congruent with subgenus *Rydbergis*. The monophyletic clade of U.S. perennial species in sections *Lanceolatae* and *Viscosae* is outlined, and asterisks mark the stellate-haired taxa.

Table 2.1. Breakdown of sites subject to base substitutions in ITS2. Indels not included.

ITS2 regions:	Number of sites	Number of variable sites	Percent of sites varying
Single-stranded	72	23	31.94%
Loop tip	17	10	58.82%
Internal loop	55	13	23.64%
Double-stranded	140	29	20.71%
Adjoining same	80	14	17.50%
Adjoining loop	60	15	25.00%
TOTAL	**212**	**52**	**24.53%**

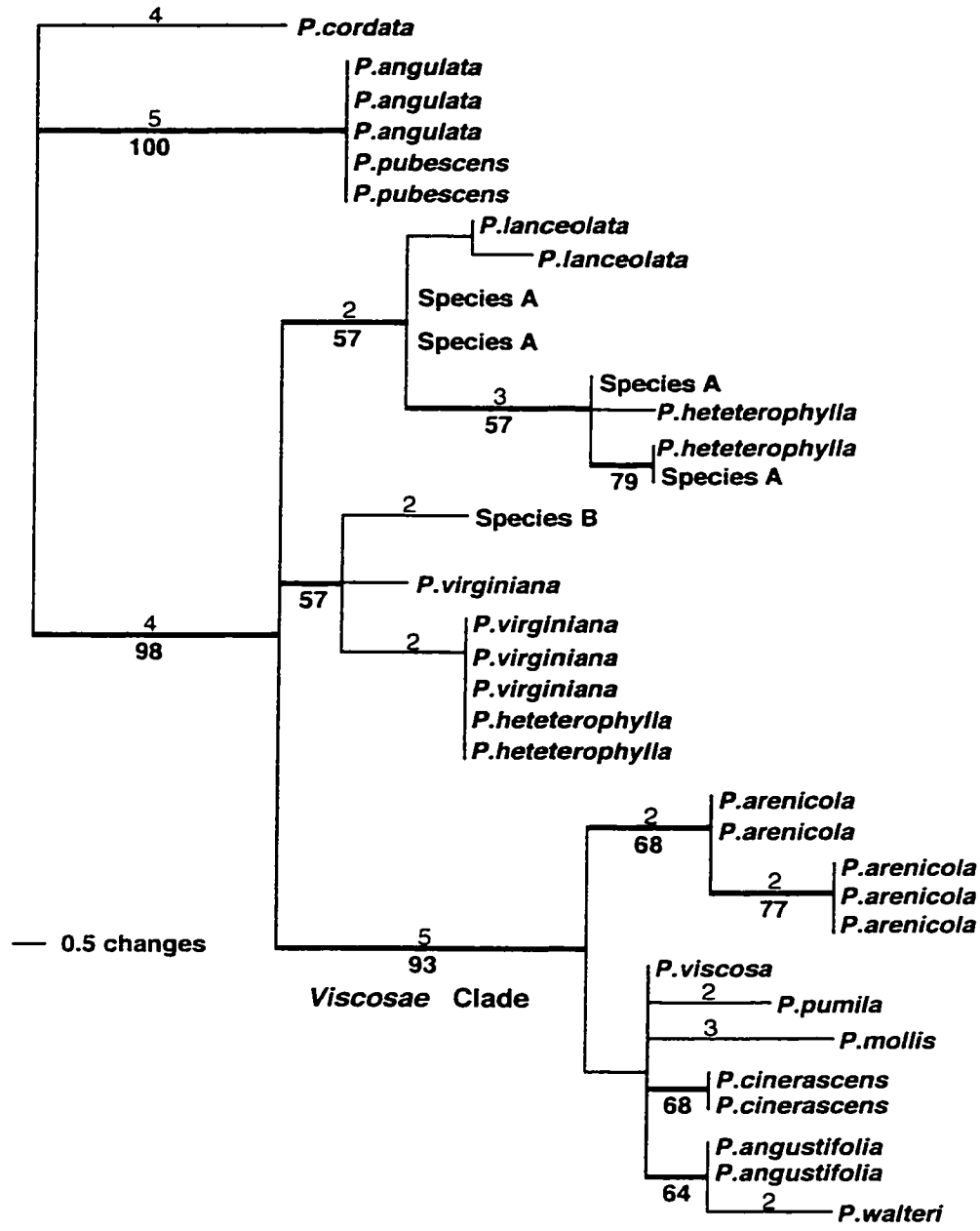

Figure 2.1. One of 50,000+ most parsimonious *waxy* trees. Length = 47 steps, CI = 0.91. Branch lengths above lines, bootstrap values below. Branches present in the strict consensus denoted by heavy lines.

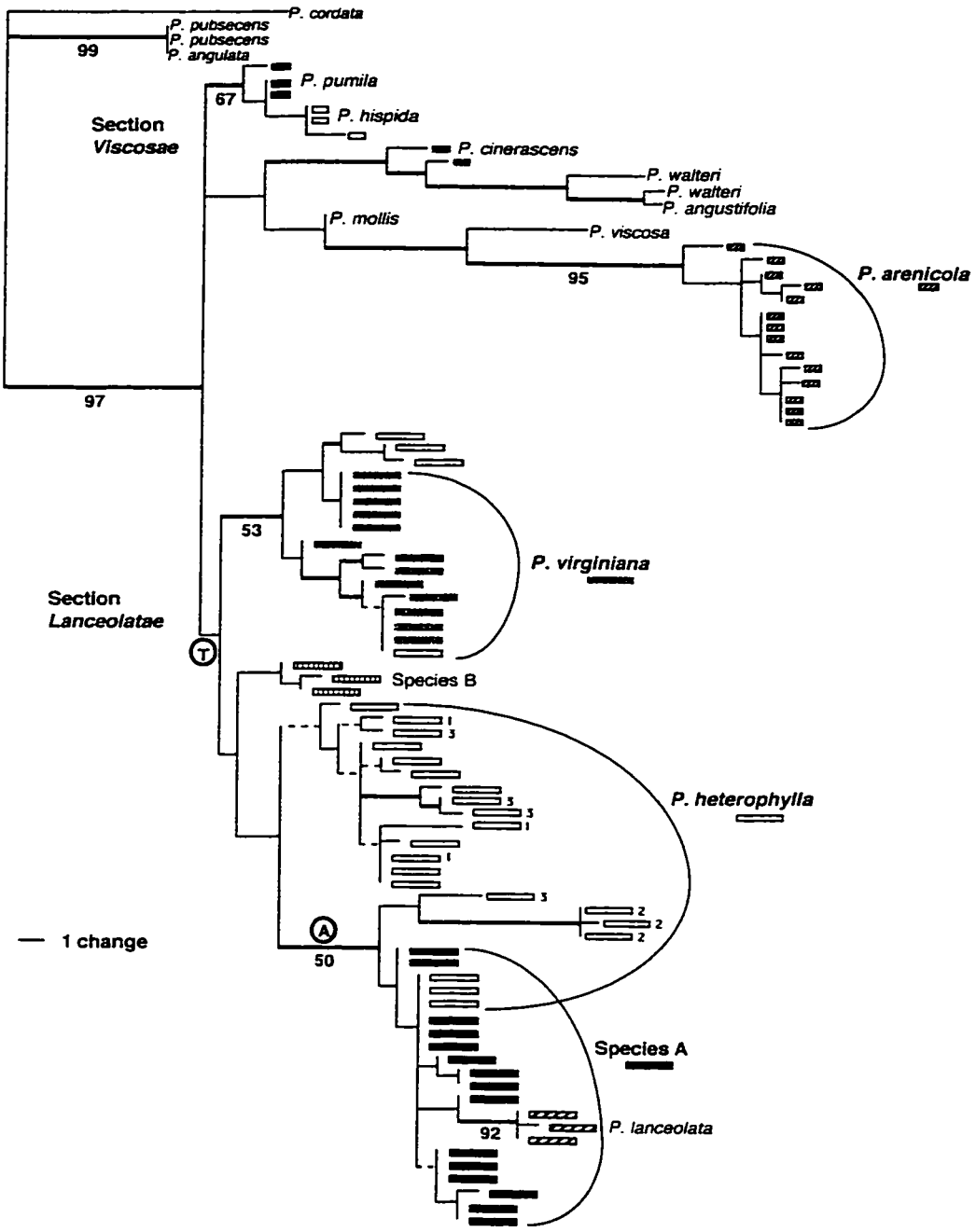

Figure 2.2. One of 50,000+ most parsimonious trees from the constrained ITS analysis. Length = 194 steps, CI = 0.5876. Heavy branches appear in the strict consensus, and dashed branched collapse in the majority rule consensus. Bootstrap values appear below the branches. Circled "T" and "A" denote nodes where the two *P. heterophylla* ITS types begin. The T-type is shared with *P. virginiana* and Species B, while the A-type is shared with Species A and *P. lanceolata*. Cloned *P. heterophylla* sequences are labeled with individual ID numbers (1, 2, 3).

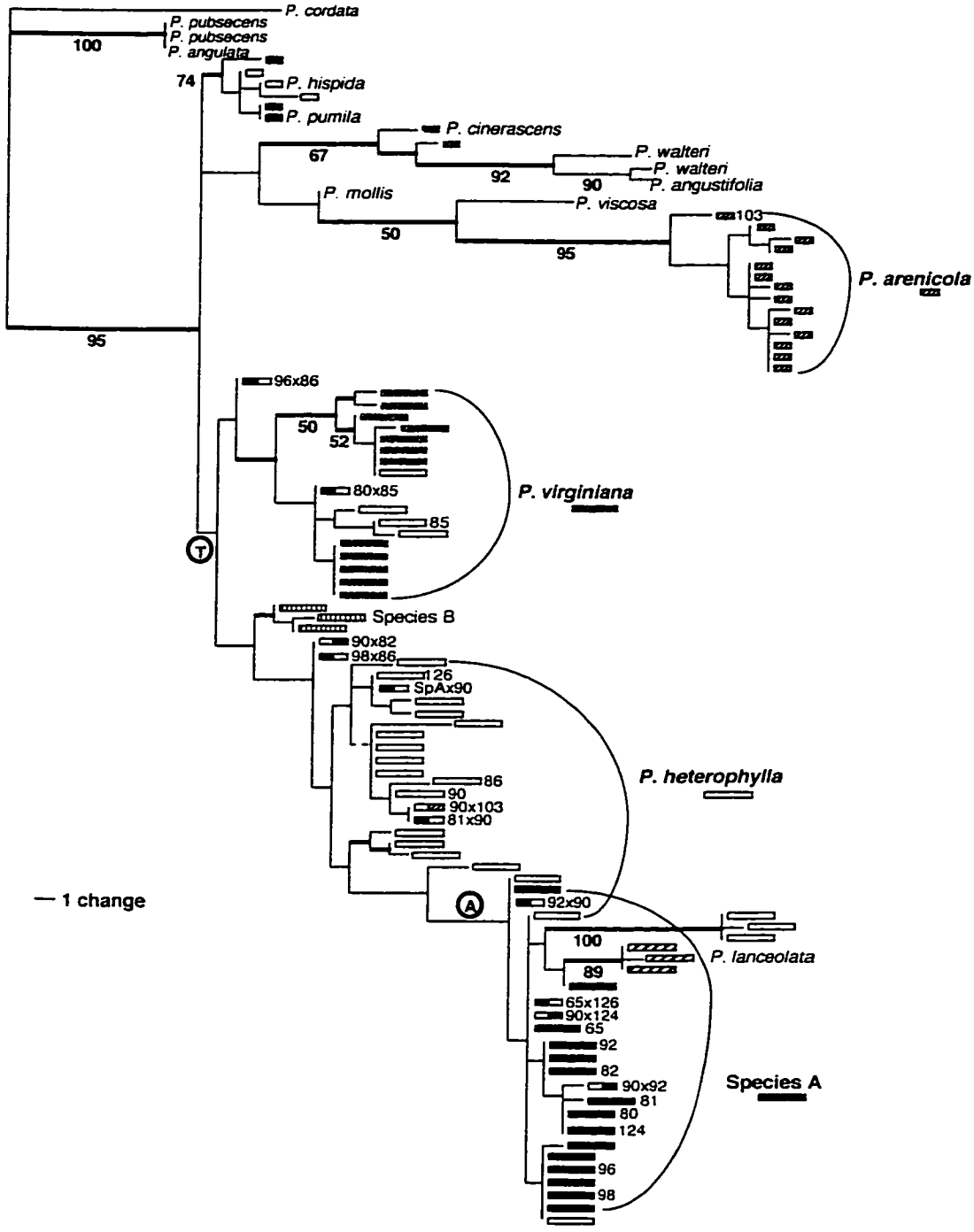

Figure 2.3. One of 50,000+ most parsimonious trees. Length = 196 steps, CI = 0.5867. Constrained ITS analysis including hybrid sequences. Parental sequences labeled with unique ID numbers, and hybrids labeled with both parental numbers. For hybrids, the maternal species color and number appears first, followed by the paternal information. Other labeling as in Figure 2.2.

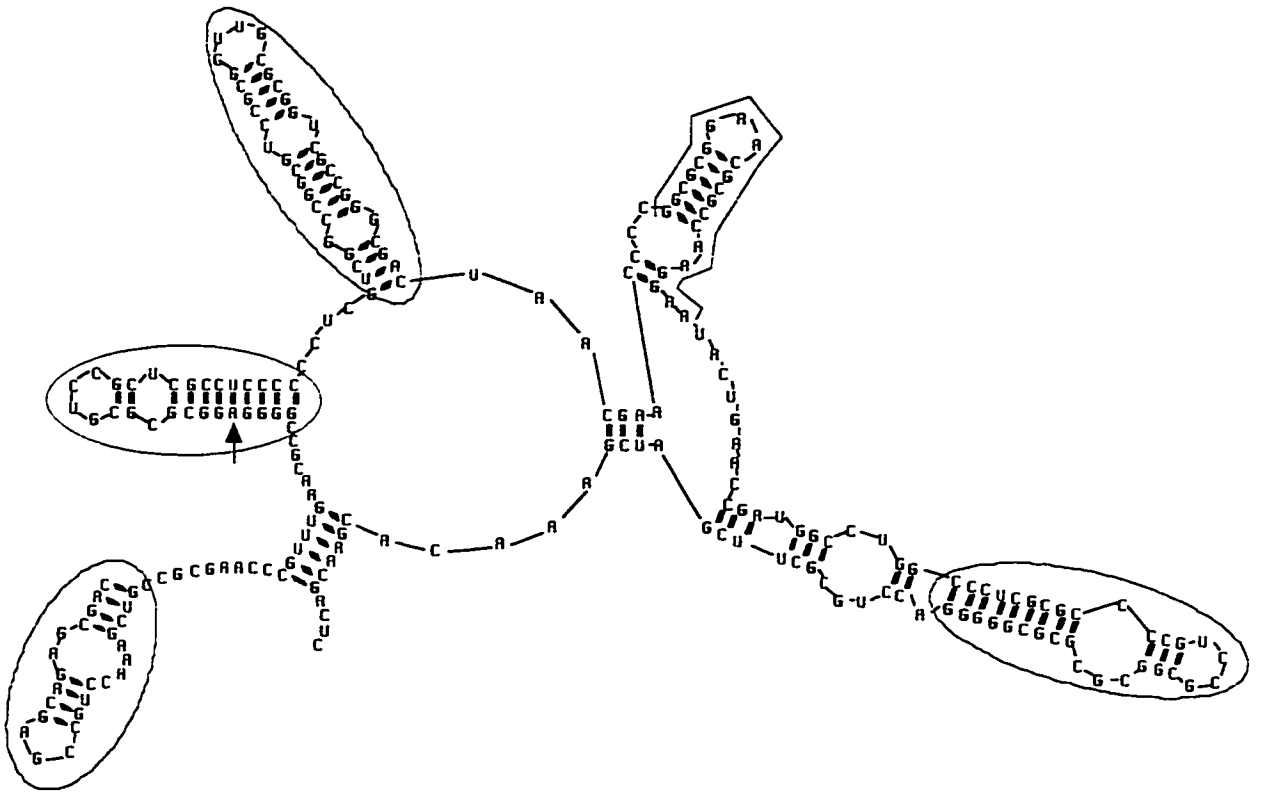

Figure 2.4. Putative secondary structure of the ITS1 region from a *P. heterophylla* A-type sequence. Circled regions represent stem and loop structures found in multiple low energy conformations. Outlined region is the conserved motif noted by Liu and Schardl (1994). The ITS-1 indel region is at the tip of the second stem and loop region, and the A/T marker site is noted with an arrow.

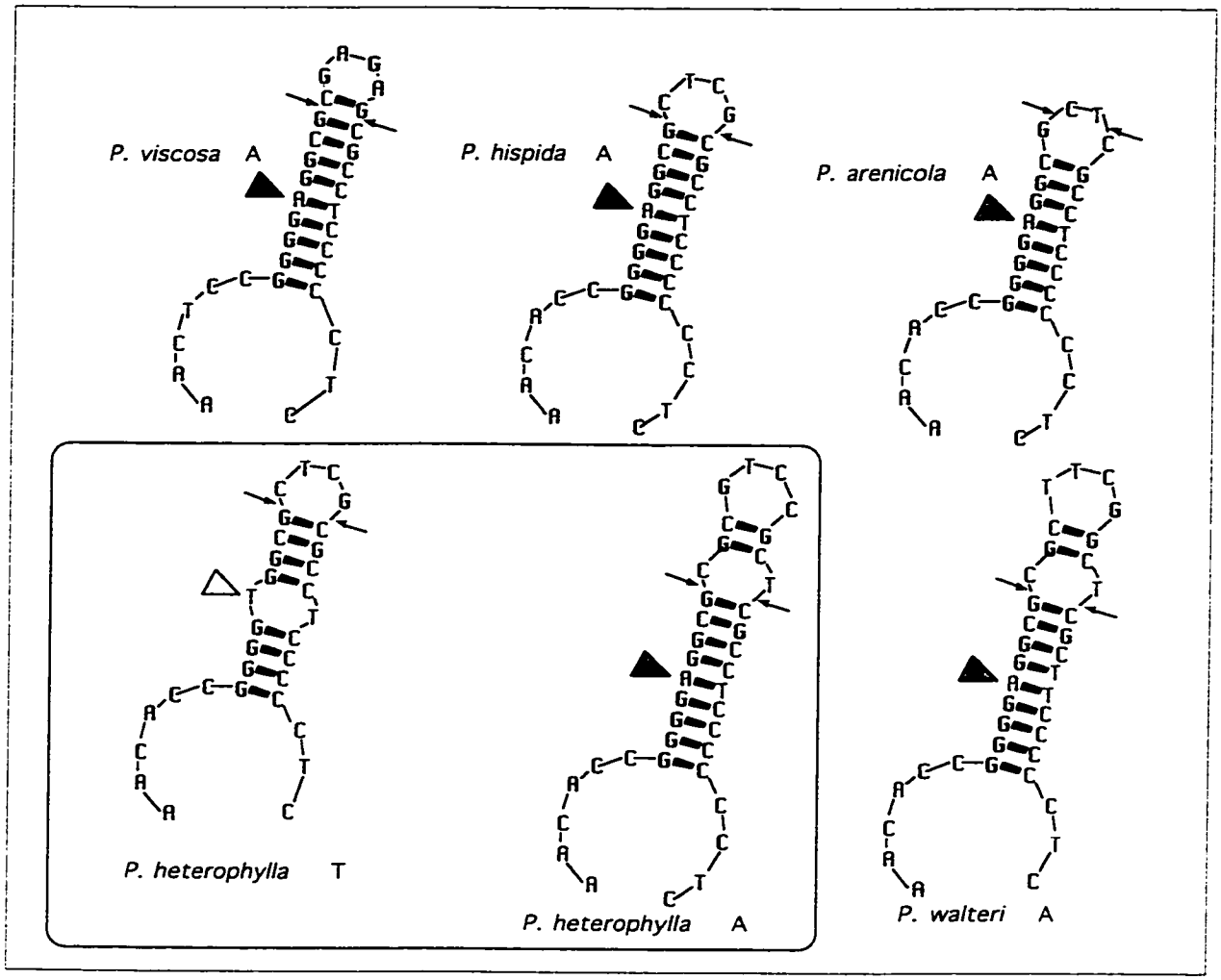

Figure 2.5. ITS1 indel regions from various taxa. The A/T marker site is indicated with large arrows. Small arrows mark where indel region starts and ends. The two ITS types from *P. heterophylla* are boxed.

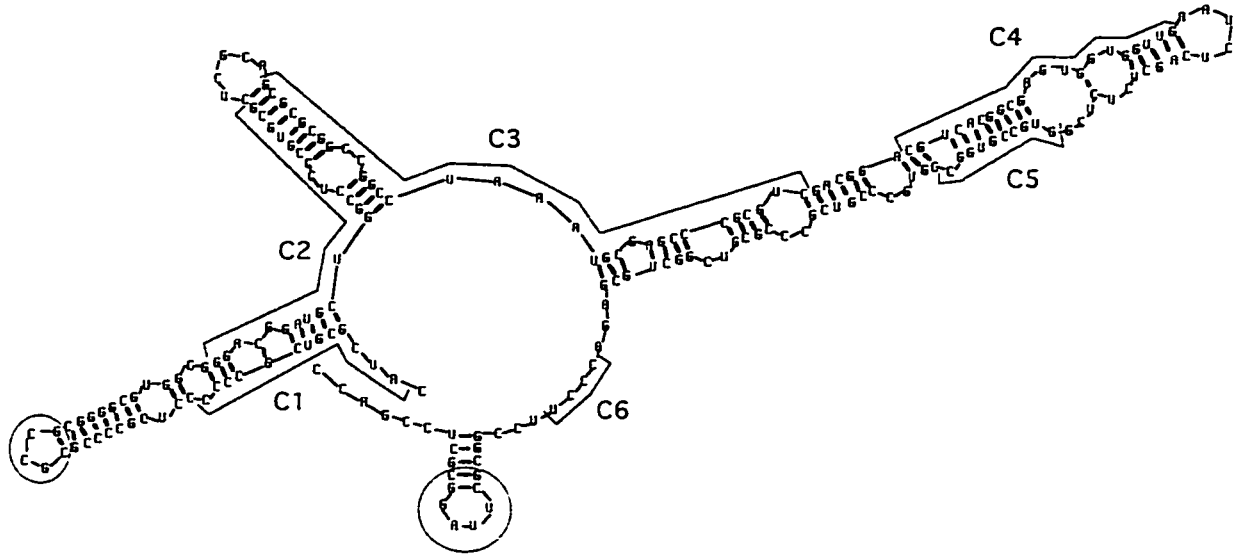

Figure 2.6. Secondary structure of the ITS2 region from a *P. heterophylla* T-type sequence. Conserved regions as seen in Hershkovitz and Zimmer (1996), labeled C1-C6. Indel regions are circled.

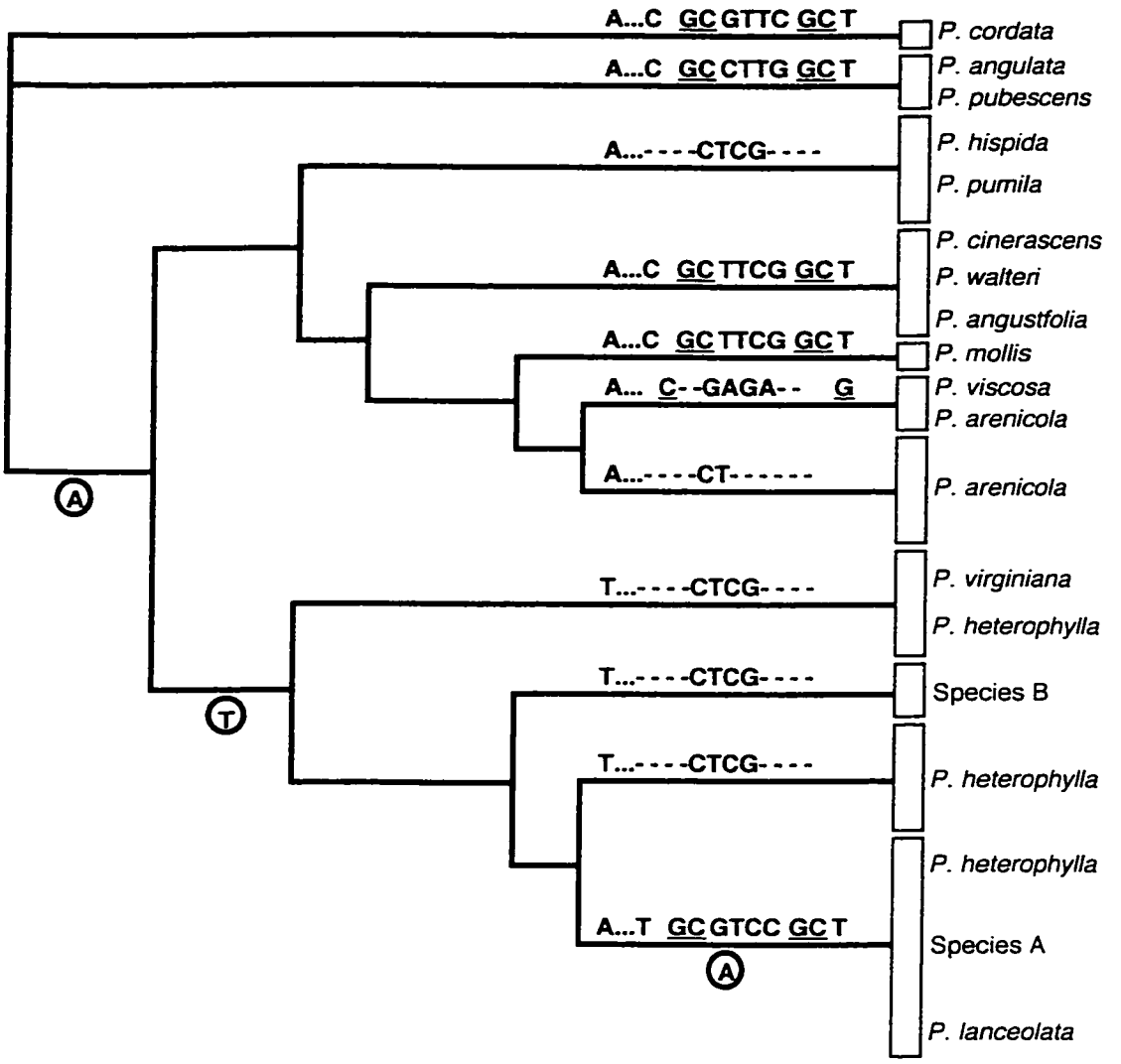

Figure 2.7. ITS1 indel types mapped onto a simplified, constrained ITS gene tree. Nucleotides involved in base-pairing are underlined.

Table 3.1. List of *Physalis* species, localities and populations sampled for isozyme analysis. Populations of different species with the same population names were sampled from the same general area. Populations marked with an asterisk had both species growing intermixed. A total of 13 populations and 215 individuals were sampled.

River bluff taxon:

Florida Caverns State Park (Jackson Co., FL)
- FLCAV-PS*
- FLCAV-FPT
- FLCAV-VC

Three Rivers State Recreation Area (Jackson Co., FL)
- 3RIVERS-LT
- 3RIVERS-LR

Torreya State Park (Liberty Co., FL)
- TORREYA-RT*
- TORREYA-TT

***P. arenicola*:**

Florida Caverns State Park (Jackson Co., FL)
- FLCAV-PS*

Three Rivers State Recreation Area (Jackson Co., FL)
- 3RIVERS-LT
- 3RIVERS-LR

Torreya State Park (Liberty Co.)
- TORREYA-RT*

Liberty Co., FL
- LIBERTY-P
- LIBERTY-LF

Table 3.2. Results of crosses within and between the river bluff taxon, *P. heterophylla* and *P. arenicola* . Notation of "na" means not applicable.

Type of cross:	Total:	Calyx enlargement	% Calyx enl.	Ovary Enlargement	% Ovary enl.	Young fruit	% Young fruit	Mature fruit	% Mat. fruit	Average seed # per fruit	% germination
P. arenicola self	15	0/15	0.0%	0/15	0.0%	na	...	na	...	0.0	na
River bluff taxon self	13	5/15	33.3%	4/15	26.7%	na	...	na	...	0.0	na
P. heterophylla self	3	0/3	0.0%	0/3	0.0%	na	...	na	...	0.0	na
P. arenicola x *P. arenicola*	43	17/43	39.5%	16/43	37.2%	13/40	32.5%	11/39	28.2%	46.2	not checked
River bluff taxon x River bluff taxon	26	14/26	53.8%	12/26	46.2%	12/26	46.2%	8/24	33.3%	29.9	39.6%
P. heterophylla x *P. heterophylla*	16	5/16	31.3%	5/16	31.3%	5/16	31.3%	5/16	31.3%	66.0	37.9%
River bluff taxon x *P. heterophylla*	27	18/27	66.7%	17/27	63.0%	17/27	63.0%	14/26	53.8%	40.3	60.9%
P. heterophylla x River bluff taxon	23	6/23	26.1%	4/23	17.4%	4/23	17.4%	3/22	13.6%	62.3	51.8%
P. arenicola x River bluff taxon	22	2/22	9.1%	1/22	4.5%	na	...	na	...	0.0	na
River bluff taxon x *P. arenicola*	26	19/26	73.1%	13/26	50.0%	3/24	12.5%	4/24	16.7%	7.3	0%
P. arenicola x *P. heterophylla*	7	2/7	28.6%	0/7	0.0%	na	...	na	...	0.0	na
P. heterophylla x *P. arenicola*	8	4/8	50.0%	3/8	37.5%	3/8	37.5%	2/7	28.6%	17.0	11.8%

Table 3.3. Nei's (1972) genetic identity for populations of the river bluff taxon and *P. arenicola* . Based on data from four enzymes: 6PGD, PGI, PGM and IDH. Numbers in parenthesis indicate the range of distance values between populations of the species. Higher values indicate more similar populations. Values range from zero to one.

Average genetic identities between populations of the two species.

Species:	Number of pops.:	1	2
1: River bluff taxon	7	0.854	
2: *P. arenicola*	6	0.583	0.831

Genetic identities between populations of the river bluff taxon.

Population:	1	2	3	4	5	6	7
FLCAV-PS	...						
FLCAV-FPT	.763	...					
FLCAV-VC	.740	.829	...				
TORREYA-RT	.674	.897	.932	...			
TORREYA-TT	.877	.802	.862	.721	...		
3RIVERS-LT	.864	.898	.941	.934	.857	...	
3RIVERS-LR	.780	.883	.939	.881	.926	.939	...

Genetic identities between populations of *P. arenicola* .

Population:	1	2	3	4	5	6	
FLCAV-PS	...						
TORREYA-RT	.869	...					
3RIVERS-LT	.989	.818	...				
3RIVERS-LR	.956	.940	.906	...			
LIBERTY-P	.692	.689	.593	.813	...		
LIBERTY-LF	.945	.710	.933	.876	.739	...	

Table 3.4. Genetic variability at four loci in 14 Florida *Physalis* populations. Standard errors appear in parentheses. Loci were considered polymorphic if more than one allele was found. Nei's (1978) unbiased estimate of expected Hardy Weinburg heterozygosity was used.

Population	Mean sample size per locus	Mean number of alleles per locus	Percentage of loci polymorphic	Mean heterozygosity	
				Direct count	Hardy-Weinberg expected
River bluff taxon: FLCAV-PS	18.0 (.0)	1.8 (.3)	75.0	0.250 (.108)	.207 (.083)
River bluff taxon: FLCAV-FPT	18.0 (.0)	2.3 (.5)	75.0	.722 (.242)	.429 (.145)
River bluff taxon: FLCAV-VC	20.0 (.0)	1.5 (.3)	50.0	.250 (.250)	.256 (.148)
River bluff taxon: TORREYA-RT	20.0 (.0)	2.3 (.3)	100.0	.525 (.139)	.398 (.087)
River bluff taxon: TORREYA-TT	14.0 (.0)	1.3 (.3)	25.0	.196 (.196)	.124 (.124)
River bluff taxon: 3RIVERS-LT	19.8 (.3)	2.0 (.4)	75.0	.372 (.181)	.321 (.152)
River bluff taxon: 3RIVERS-LR	19.8 (.3)	1.5 (.3)	50.0	.349 (.226)	.217 (.129)
P. arenicola : FLCAV-PS	18.3 (1.4)	1.8 (.5)	50.0	.084 (.062)	.160 (.099)
P. arenicola : TORREYA-RT	10.5 (.3)	1.3 (.3)	25.0	.175 (.175)	.120 (.120)
P. arenicola : 3RIVERS-LT	19.0 (.7)	1.3 (.3)	25.0	.000 (.000)	.029 (.029)
P. arenicola : 3RIVERS-LR	16.0 (.0)	1.5 (.3)	50.0	.078 (.078)	.221 (.130)
P. arenicola : LIBERTY-P	14.0 (.0)	2.0 (.4)	75.0	.339 (.153)	.282 (.113)
P. arenicola : LIBERTY-LF	8.0 (.0)	1.3 (.3)	25.0	.250 (.250)	.133 (.133)

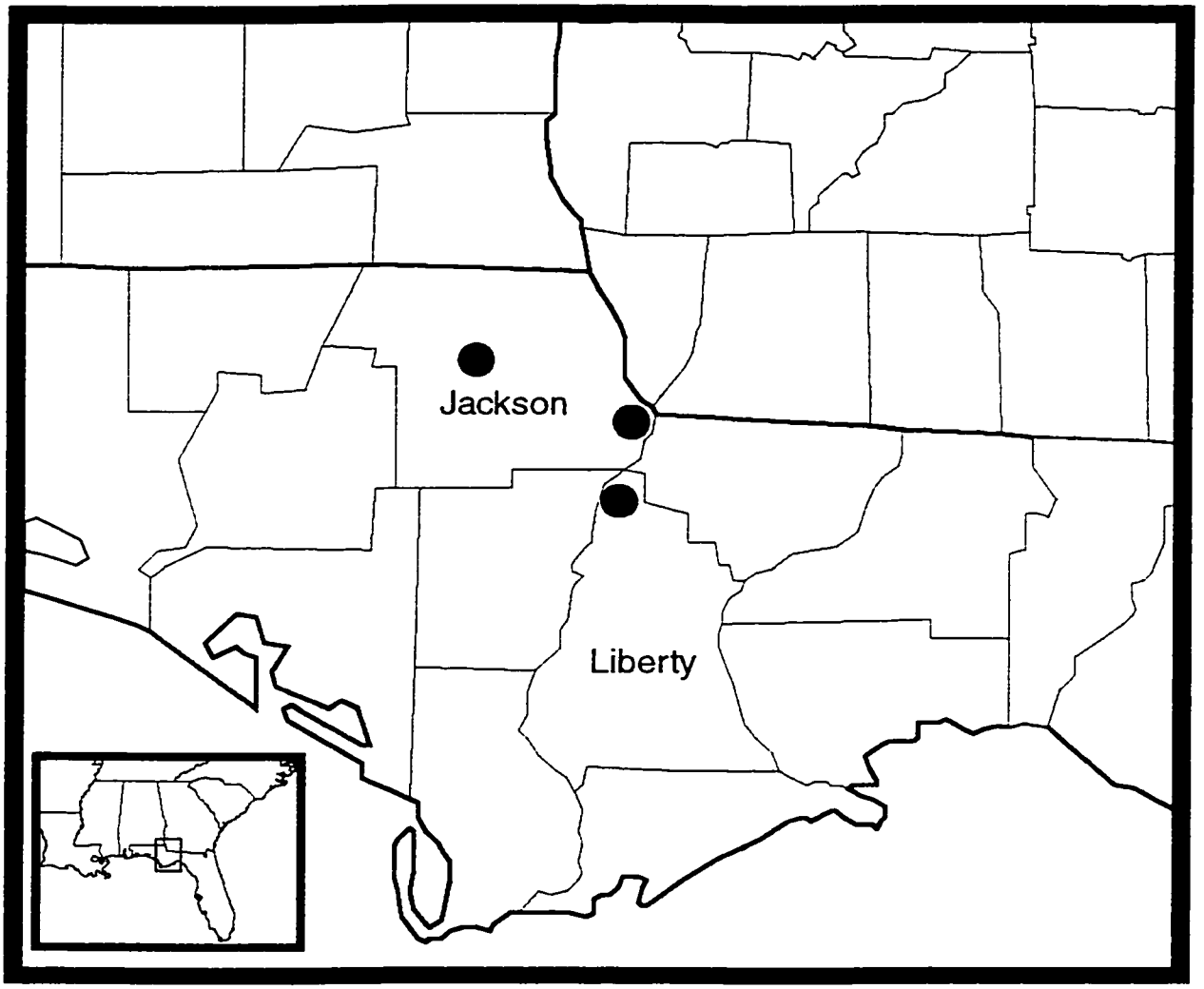

Figure 3.1. Map of the three areas of Florida where the river bluff taxon was found. Liberty county location is Torreya State Park. Eastern Jackson county location is Three Rivers State Recreation Area. Both locations are along the Apalachicola river bluffs. The central Jackson county location is Florida Caverns State Park, in the Marianna Lowlands, along the bluffs of the Chipola river, which is a tributary of the Apalachicola river.

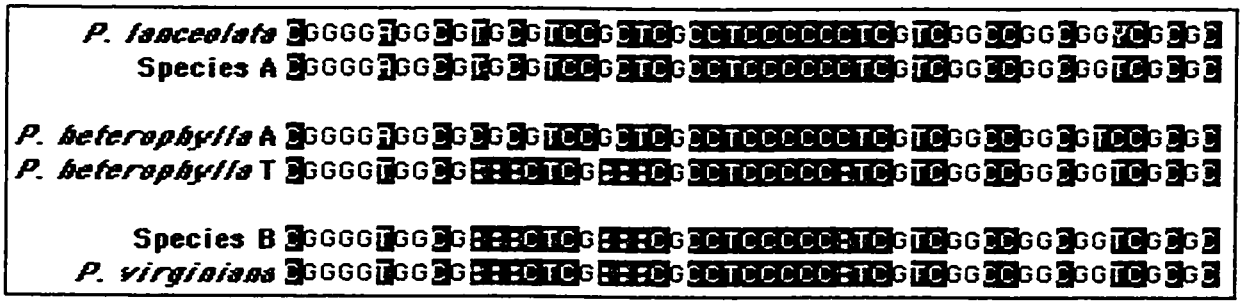

Figure 3.2. ITS1 indel region in the species of section *Lanceolatae*. The two ITS types found in *P. heterophylla* are more similar to the ITS types seen in other species than they are to each other. The A/T marker site, found four nucleotides prior to the indel region, occurs about 50 nucleotides into ITS1.

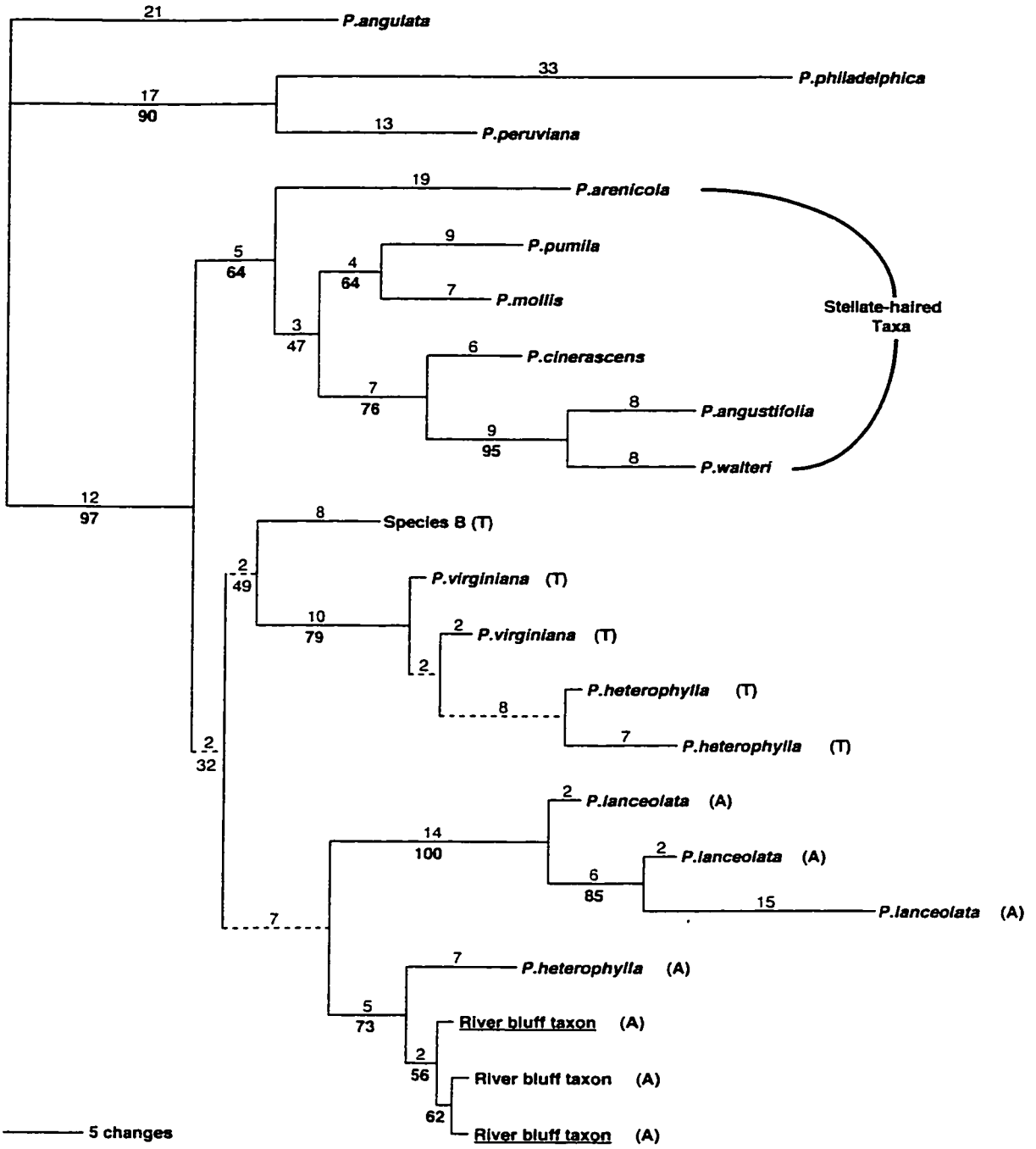

Figure 3.3. One of four most parsimonious trees from the combined analysis of ITS and two regions of *waxy*. Trees were 288 steps long. Branch lengths are above branches, bootstrap values below. Dashed branches collapsed in the strict consensus. Underlined individuals of the river bluff taxon were not sequenced for the 600+ bp *waxy* fragment between exons eight and ten. ITS types (T and A) are indicated for individuals from section *Lanceolatae*.

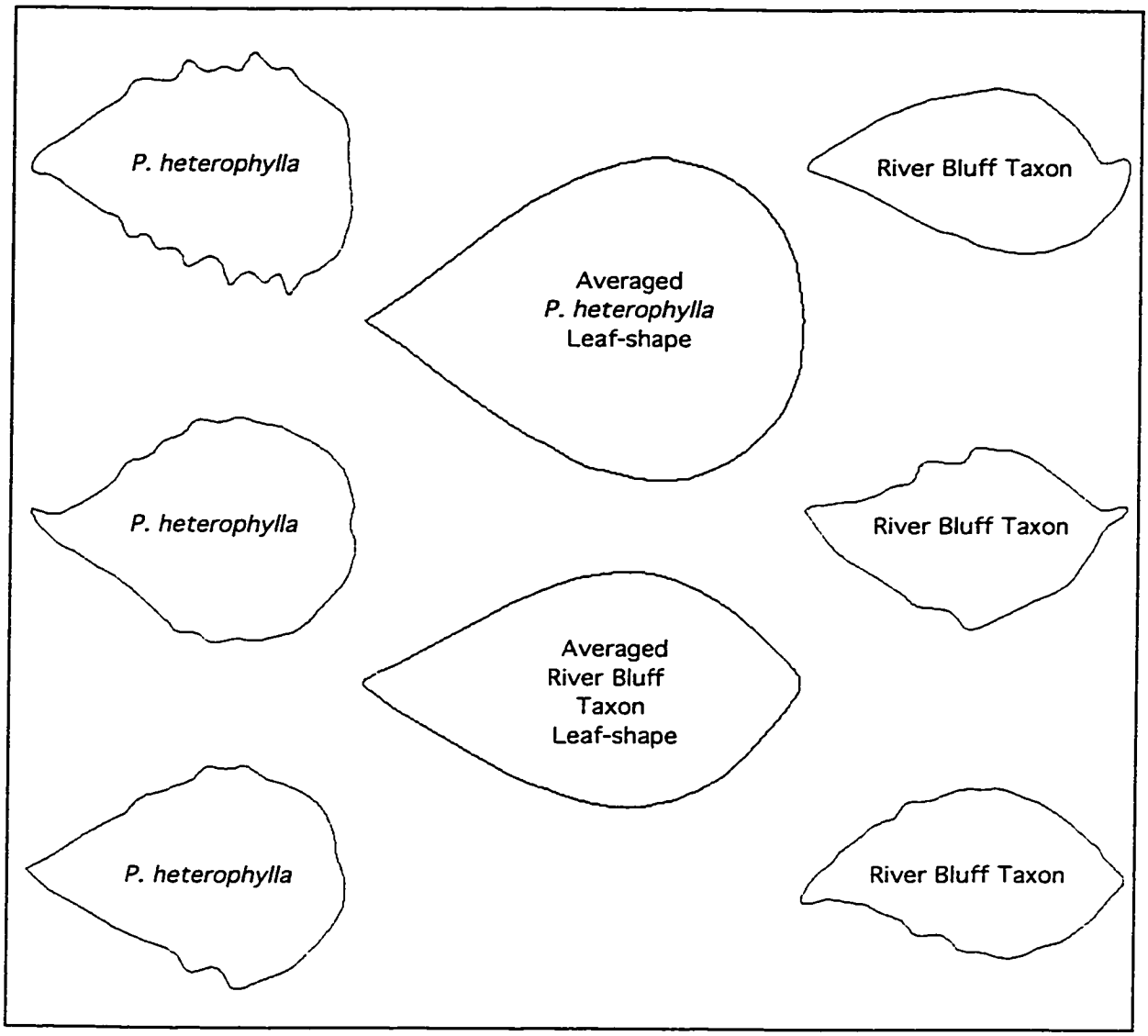

Figure 3.4. Comparison of average leaf shapes for *P. heterophylla* and the river bluff taxon. These were generated by averaging and then plotting the Fourier harmonics for all sampled leaves of a taxon. Three exemplar leaves from each taxon are included to illustrate features which are lost in the averaging process, such as toothing of the leaf margins and the cordate leaf bases of *P. heterophylla*.

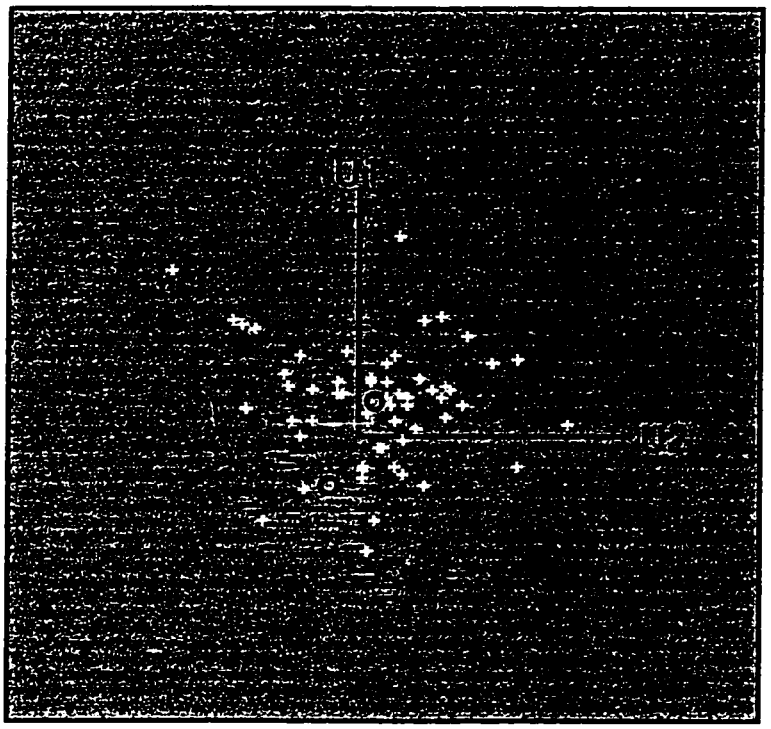

Figure 3.5. Plot of the first three principal components from the analysis of leaf shape between the river bluff taxon and *P. heterophylla*. Centered at the origin, scaled by the range. Lines represent leaves from the river bluff taxon. Points are leaves from *P. heterophylla*. Average leaf shapes of each taxon indicated by open circles.

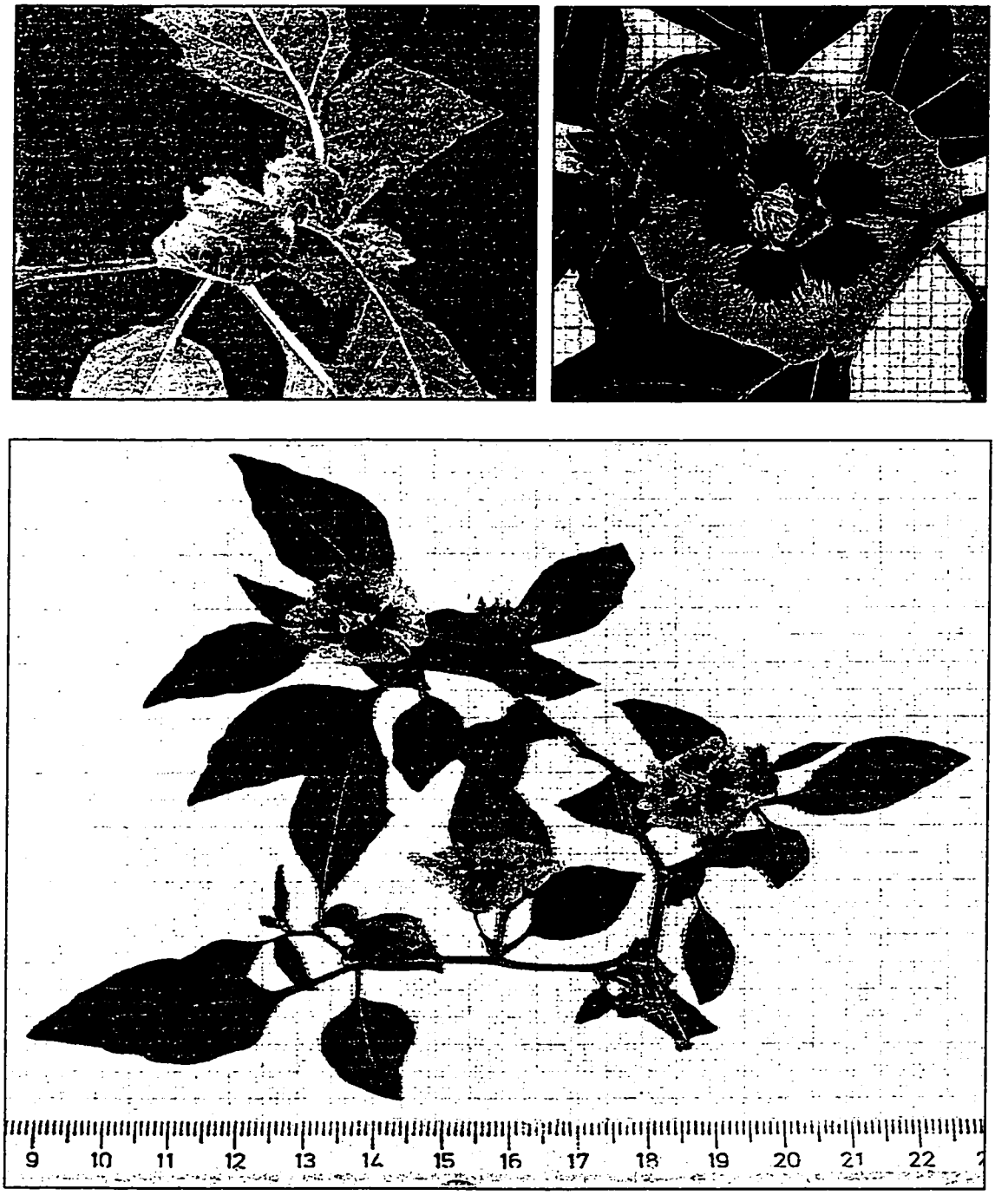

Figure 4.1. *Physalis diminuta*. Photo of the type specimen (Whitson #1203a) illustrating major morphological features of the species. Note the small size of the shoots, the narrowly ovate leaves with attenuate to truncate bases, and the sparse, hirtellous pubescence. The scale is in centimeters, and the smallest units of the background grid are in milimeters.

CHAPTER 1, WORKS CITED:

AVERETT, J. E. 1970. New combinations in the Solaneae (Solanaceae) and comments regarding the taxonomic status of *Leucophysalis*. Annals of the Missouri Botanical Garden 57(3): 380-381.

AVERETT, J. E. 1973. Biosystematic study of *Chamaesaracha* (Solanaceae). Rhodora 75: 325-365.

AVERETT , J. E. 1979. Biosystematics of the physaloid genera of the Solanaceae in North America. Pp. 493-503 in *The biology and taxonomy of the Solanaceae*, eds. J. G. HAWKES, R. N. LESTER, and A. D. SKELDING. London: Academic Press.

AVERETT , J. E. and W. G. D'ARCY. 1983. Flavonoids of *Oryctes*. Phytochemistry 22(10): 2325-2326.

AXELIUS, B. 1992. Testa patterns in some species of *Physalis* L. and some other genera in the tribe Solaneae (Solanaceae). International Journal of Plant Sciences 153(3): 488-502.

AXELIUS, B. 1995. A new combination in *Physalis* (Solanaceae). Phytologia 79(1):10-11.

AXELIUS, B. 1996. The phylogenetic relationships of the physaloid genera (Solanaceae) based on morphology. American Journal of Botany 83(1):118-124.

BALDWIN, B. G. 1992 Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the compositae. Molecular Phylogenetics and Evolution 1(1): 3-16.

BARBOZA, G. E. 2000. Rehabilitacion del genero *Quincula* (Solanaceae: Solaneae). Kurtziana 28(1): 69-79.

BUCKLER, E. S., A. IPPOLITO, and T. P. HOLTSFORD. 1997. The evolution of ribisomal DNA: divergent paralogues and phylogenetic implications. Genetics 145: 821-832.

BURNHAM, R. J. and A. GRAHAM. 1999. The history of Neotropical vegetation: new developments and status. Annals of the Missouri Botanical Garden 86: 546-589.

CHIANG, H. C., S. M. JAW, C. F. CHEN, and W. S. KAN. 1992. Antitumor agent, physalin F from *Physalis angulata* L. Anticancer Research 12(3): 837-843.

D'ARCY, W. G. 1986. The calyx in *Lycianthes* and some other genera. Annals of the Missouri Botanical Garden 73(1): 117-127.

D'ARCY, W. G. 1991. The Solanaceae since 1976, with a review of its biogeography. Pp. 75-137 in *Solanaceae II: taxonomy, chemistry, evolution*, eds. J. G. Hawkes et al. Kew: The Royal Botanic Gardens.

D'ARCY, W. G. and J. E. AVERETT. 1996. Recognition of tribes Capsiceae and Physaleae, subfamily Solanoideae, Solanaceae. Phytologia 80(4): 273-275.

D'ARCY, W. G, J. L. GENTRY, and J. E. AVERETT. 1981. Recognition of *Brachistus* (Solanaceae). Annals of the Missouri Botanical Garden 68: 226-227.

DEBRY, R. W. and R. G. OLMSTEAD. 2000. A simulation study of reduced tree-search effort in bootstrap resampling analysis. Systematic Biology 49(1):171-179.

DOYLE, J. J. and J. L. DOYLE. 1987. A rapid DNA isolation for small quantities of fresh tissue. Phytochemical Bulletin 19: 11-15.

ESTRADA, E. 1996. Systematic studies in *Physalis*. Ph.D. dissertation, Cornell University, Ithaca, New York.

ESTRADA, E., and M. MARTINEZ. 1999. *Physalis* L. (Solanaoideae: Solaneae) and allied genera: I. a morphology-based cladistic analysis. Pp. 139-159 in *Solanaceae IV: Advances in biology and utilization*, eds. M. Nee, D. E. Symon, R. N. Lester, and J. P. Jessop. Kew: Royal Botanic Gardens.

ESTRADA, E., and M. MARTINEZ. 1998. *Physalis* (Solanaceae) and allied genera: *Tzeltalia*, a new genus from the highlands of southern Mexico and northwestern Guatemala. Brittonia 50(3): 285-295.

GENTRY, J. L., and W. G. D'ARCY. 1986. Solanaceae of Mesoamerica. Pp. 14-26 in *Solanaceae II, biology and systematics*, ed. W. G. D'Arcy. New York: Columbia University Press.

GRAHAM, A. 1999. *Late Cretaceous and Cenozoic history of North American vegetation north of Mexico*. New York: Oxford University Press.

FARRIS, J. S. 1989. The retention index and rescaled consistency index. Cladistics 5: 417-419.

FARRIS, J. S., M. KALLERSJO, A. G. KLUGE, and C. BULT. 1994. Testing significance of incongruence. Cladistics 10: 315-319.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783-791.

HENDRYCH, R. 1989. *Physalis alkekengi*, in Europa und in der Tschechoslowakei besonders. Acta Universitatis Carolinae. Biologica 33: 1-42.

HINTON, W. F. 1970. The taxonomic status of *Physalis lanceolata* (Solanaceae) in the Carolina sandhills. Brittonia (22): 14-19.

HUNZIKER, A. T. 1979. South American Solanaceae: A synoptic survey. Pp. 49-85 in *The biology and taxonomy of the Solanaceae*, eds. J. G. Hawkes, R. N. Lester, and A. D. Skelding. London: Academic Press.

KENNELLY, E. J., C. GERHAEUSER, L. L. SONG, J. G. GRAHAM, C. W. W. BEECHER, J. M. PEZZUTO, and A. D. KINGHORN. 1997. Induction of quinone reductase by withanolides isolated from *Physalis philadelphica* (tomatillos). Journal of Agricultural and Food Chemistry 45 (10) 3771-3777.

MARTINEZ, M. 1993. Systematics of *Physalis* (Epeteiorhiza) Section *Epeteiorhiza*. Ph.D. dissertation, University of Texas at Austin.

MARTINEZ, M. 1998. Revision of *Physalis* Section *Epeteiorhiza* (Solanaceae). Anales del Instituto del Biologia Universidad Nacional Autonoma de Mexico, Serie Botanica 69(2): 71-117.

MARTINEZ, M. 1999. Infrageneric taxonomy of *Physalis*. Pp. 275-283 in *Solanaceae IV: Advances in biology and utilization*, eds. M. Nee, D. E. Symon, R. N. Lester and J. P. Jessop. Kew: Royal Botanic Gardens.

MATTHEWS, J. V. JR. and L. E. OVENDEN. 1990. Late Tertiary plant macrofossils from localities in arctic/subarctic North America: a reveiw of the data. Arctic 43(4): 364-392.

MAGALLON, S., P. R. CRANE, and P. S. HERENDEEN. 1999. Phylogenetic pattern, diversity, and diversification of eudicots. Annals of the Missouri Botanical Garden 86(2): 297-372.

MENZEL, M. Y. 1950. Cytotaxonomic observations on some genera of the Solanaceae: *Margaranthus*, *Saracha* and *Quincula*. American Journal of Botany 37: 25-30.

MENZEL, M. Y. 1951. The cytotaxonomy and genetics of *Physalis*. Proceedings of the American Philosophical Society 95(2): 132-183.

MILLER, R. E., M. D. RAUSHER, and P. S. MANOS. 1999. Phylogenetic systematics of *Ipomea* (Convolvulaceae) based on ITS and *waxy* sequences. Systematic Botany 24(2): 209-227.

MIONE, T., R. G. OLMSTEAD, R. K. JANSEN, and G. J. ANDERSON. 1994. Systematic implications of chloroplast DNA variation in *Jaltomata* and selected physaloid genera (Solanaceae). American Journal of Botany 81(7): 912-918.

OLMSTEAD, R. G. and J. D. PALMER. 1992. A chloroplast DNA phylogeny of the Solanaceae: subfamilial relationships and character evolution. Annals of the Missouri Botanical Garden 79: 346-360.

OLMSTEAD, R. G., J. A. SWEERE, R. E. SPANGLER, L. BOHS and J. D. PALMER. 1999. Phylogeny and provisional classification of the Solanaceae based on chloroplast data. Pp. 111-137 in *Solanaceae IV: Advances in biology and utilization*, eds. M. Nee, D. E. Symon, R. N. Lester and J. P. Jessop. Kew: Royal Botanic Gardens.

PIETRO, R. C. L. R., S. KASHIMA, D. N. SATO, A. H. JANUARIO, and S. C. FRANCA. 2000. In vitro antimycobacterial activities of *Physalis angulata* L. Phytomedicine Jena 7(4): 335-338.

RYDBERG, P. A. 1896. The North American species of *Physalis* and related genera. Memoirs of the Torrey Botanical Club 4(5): 297-372.

SALEHUZZAMAN, S. N. I. M., E. JACOBSEN, and R. G. F. VISSER. 1993. Isolation and characterization of a cDNA encoding granule-bound starch synthase in cassava (*Manihot esculenta* Crantz) and its antisense expression in potato. Plant Molecular Biology 23: 947-962.

SULLIVAN, J. R. 1984a. Systematic studies in *Physalis* (Solanaceae) Ph.D. dissertation, University of Oklahoma, Norman.

SULLIVAN, J. R. 1984b. Pollination biology of *Physalis viscosa* var. *cinerascens* (Solanaceae). American Journal of Botany 71(6): 815-820.

SWOFFORD, D. L. 1998. PAUP*: Phylogenetic Analysis Using Parsimony, version 4d64. Sinauer Associates, Sunderland, Massachusetts.

VAN DER LEIJ, F. R., R. G. F. VISSER, A. S. PONSTEIN, E. JACOBSEN, and W. J. FEENSTRA. 1991. Sequence of the structural gene for granule-bound starch synthase of potato (*Solanum tuberosum*) and evidence for a single point deletion in the amf allele. Molecular Gene Genetics 228: 240-248.

WATERFALL, U. T. 1958. *Physalis* in Mexico, Central America and the West Indes. Rhodora 69: 82-120 and 203-329.

WATERFALL, U. T. 1967. A taxonomic study of the genus *Physalis* in North America north of Mexico. Rhodora 60: 107-173.

WHITE, T. J., T. BRUNS, S. LEE, and J. TAYLOR. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pp. 315-322 in *PCR Protocols: A Guide to Methods and Applications*, eds. M. Innis, D. Gelfand, J. Sninsky, and T. White. San Diego: Academic Press.

YODER, A. D., J. A. IRWIN, and B. A. PAYSEUR. 2001. Faliure of the ILD to determine data compatibility for slow loris phylogeny. Systematic Biology 50(3): 408-424.

CHAPTER 2, WORKS CITED:

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

AVISE, J. C. 1994. *Molecular markers, natural history and evolution*. New York: Chapman and Hall.

BALDWIN, B. G., M. J. SANDERSON, J. M. PORTER, M. F. WOJCIECHOWSKI, C. S. CAMPBELL, and M. J. DONOGHUE. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Annals of the Missouri Botanical Garden 82: 247-277.

BUCKLER, E. S., A. IPPOLITO, and T. P. HOLTSFORD. 1997. The evolution of ribisomal DNA: divergent paralogues and phylogenetic implications. Genetics 145: 821-832.

CAMPBELL, C., M. F. WOJCIECHOWSKI, B. G. BALDWIN, L. A. ALICE, and M. J. DONOGHUE. 1997. Persistant nuclear ribosomal DNA sequence polymorphism in the *Amelanchier* agamic complex (Rosaceae). Molecular Biology and Evolution 14(1): 81-90.

DEBRY, R. W. and R. G. OLMSTEAD. 2000. A simulation study of reduced tree-search effort in bootstrap resampling analysis. Systematic Biology 49(1):171-179.

DENDUANGBORIPANT, J., and Q. C. B. CRONK. 2000. High intraindividual variation in internal transcribed spacer sequences in *Aeschynanthus* (Gesneriaceae): implications for phylogenetics. Proceedings of the Royal Society of London B. 267: 1407-1415.

DOYLE, J. J. and J. I. DAVIS. 1998. Homology in molecular phylogenetics: a parsimony perspective. Pp. 101-131 in *Molecular Systematics of Plants II*, eds. D. E. Soltis, P. S. Soltis, and J. J. Doyle. Boston: Kluwer Academic Press.

DOYLE, J. J. and J. L. DOYLE. 1987. A rapid DNA isolation for small quantities of fresh tissue. Phytochemical Bulletin 19: 11-15.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783-791.

FRANZKE, A. and K. MUMMENHOFF. 1999. Recent hybrid speciation in *Cardamine* (Brassicaceae) – conversion of nuclear ribosomal ITS sequences in statu nascendi. Theoretical Applied Genetics 98: 831-834.

HARTMANN, S., J. D. NASON, and D. BHATTACHARYA. 2001. Extensive ribosomal DNA variation in the columnar cactus Lophocereus. Journal of Molecular Evolution 53: 124-134.

HERSHKOVITZ, M. A. and E. A. ZIMMER. 1996. Conservation patterns in angiosperm rDNA ITS2 sequences. Nucleic Acids Research 24(15): 2857-2867.

HILLIS, D. M, C. MORTIZ, C. A. PORTER, and R. J. BAKER. 1990. Evidence for biased gene conversion in concerted evolution of ribosomal DNA. Science 251: 308-310.

HILLIS, D. M. and M. T. DIXON. 1991. Ribosomal DNA: molecular evolution and phylogenetic inference. Quarterly Review of Biology 66(4): 411-453.

HINTON, W. F. 1970. The taxonomic status of *Physalis lanceolata* (Solanaceae) in the Carolina sandhills. Brittonia (22): 14-19.

HINTON, W. F. 1975. Natural hybridization and extinction of a population of *Physalis virginiana* (Solanaceae). American Journal of Botany 62(2): 198-202.

HINTON, W. F. 1976. The systematics of *Physalis pumila* ssp. *hispida* (Solanaceae) Systematic Botany (1): 188-193.

JOBST, J., K. KING, and V. HEMLEBEN. 1998. Molecular evolution of the internal transcribed spacers (ITS1 and ITS2) and phylogenetic relationships among species of the family Cucurbitaceae. Molecular Phylogenetics and Evolution 9(2): 204-219.

KISS, T., M. KISS M, S. ABEL and F. SOLYMOSY. 1988. Nucleotide sequence of the 17S-25S spacer region from tomato rDNA. Nucleic Acids Research 16: 7179.

LALEV, A. I. and R. N. NAZAR. 1999. Structural equivalence in the transcribed spacers of pre-rRNA transcripts in *Schizosaccharomyces pombe*. Nucleic Acids Research 27(13): 3071-3078.

LIU, J. S. and C. L. SCHARDL. 1994. A conserved sequence in internal transcribed spacer 1 of plant nuclear rRNA genes. Plant Molecular Biology 26: 774-778.

MAI, J. C. and A. W. COLEMAN. 1996. The internal transcribed spacer 2 exhibits a common secondary structure in green algae and flowering plants. Molecular Evolution 44: 258-271.

MAYOL, M. and J. A. ROSELLO. 2001. Why nuclear ribosomal DNA spacers (ITS) tell different stories in *Quercus*. Molecular Phylogenetics and Evolution 19(2): 167-176.

MCDADE, L. 1992. Hybrids and phylogenetic systematics II: the impact of hybrids on cladistic analysis. Evolution 46(5): 1329-1346.

MENZEL, M. Y. 1951. The cytotaxonomy and genetics of *Physalis*. Proceedings of the American Philosophical Society 95(2): 132-183.

MUIR, G., C. C. FLEMING, and C. SCHLOTTERER. 2001. Three divergent rRNA clusters predate the species divergence in *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L.. Molecular Biology and Evolution 18(2): 112-119.

POPP, M. and B. OXELMAN. 2001. Inferring the history of the polyploid *Silene aegaea* (Caryophyllaceae) using plastid and homoelogous nuclear DNA sequences. Molecular Phylogenetics and Evolution. 20(3): 474-481.

Rieseberg, L. H., Ellstrand N. C. 1993. What can molecular and morphological markers tell us about plant hybridization? *Critical Reviews in Plant Sciences* 12(3): 213-241.

SANG, T., D. J. CRAWFORD, and T. F. STUESSY. 1995. Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. Proceedings of the National Academy of Science, USA 92(15): 6813-6817.

SULLIVAN, J. R. 1984a. Systematic studies in *Physalis* (Solanaceae) Ph.D. dissertation, University of Oklahoma, Norman.

SULLIVAN, J. R. 1984b. Pollination biology of *Physalis viscosa* var. *cinerascens* (Solanaceae). American Journal of Botany 71(6): 815-820.

SULLIVAN, J. R. 1985. Systematics of the *Physalis viscosa* complex (Solanaceae). Systematic Botany 10(4): 426-444.

Swofford, D. L. 1998. PAUP*: Phylogenetic Analysis Using Parsimony, version 4d64. Sinauer Associates, Sunderland, Massachusetts.

WHEELER, W. C. and R. L. HONEYCUTT. 1988. Paired sequence difference in ribosomal RNAs: evolutionary and phylogenetic implications. Molecular Biology and Evolution 5(1): 90-96.

WIDMER, A. and M. BALTISBERGER. 1999. Molecular evidence for allopolyploid speciation and a single origin of the narrow endemic *Draba ladina* (Brassicaceae). American Journal of Botany 86(9): 1282-1289.

ZIMMER, E. A., S. L. MARTIN, S. M. BEVERLEY, Y. W. KAN, and A. C. WILSON. 1980. Rapid duplication and loss of genes coding for the alpha chains of hemoglobin. Genetics 77(4): 2158-2162.

ZUCKER, M., D. H. MATHEWS, and D. H. TURNER. 1999. Algorithms and thermodynamics for RNA secondary structure prediction: a practical guide. In: *RNA biochemistry and biotechnology*, eds. J. Barciszewski and B. F. C. Clark. NATO ASI Series. Dordrecht: Kluwer Academic.

CHAPTER 3, WORKS CITED:

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

DATA DESCRIPTION INC. 1995. DataDesk 5.0. Ithaca, New York.

DOYLE, J. J. and J. L. DOYLE. 1987. A rapid DNA isolation for small quantities of fresh tissue. Phytochemical Bulletin 19: 11-15.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783-791.

FERSON, S., F. J. ROLF, and R. K. KOEHN. 1985. Measuring shape variation of two-dimensional outlines. Systematic Zoology 34(1): 59-68.

HILLIS, D. M., C. MORTIZ, C. A. PORTER, and R. J. BAKER. 1990. Evidence for biased gene conversion in concerted evolution of ribosomal DNA. Science 251: 308-310.

HINTON, W. F. 1970. The taxonomic status of *Physalis lanceolata* (Solanaceae) in the Carolina sandhills. Brittonia (22): 14-19.

HINTON, W. F. 1975. Natural hybridization and extinction of a population of *Physalis virginiana* (Solanaceae). American Journal of Botany 62(2): 198-202.

HINTON, W. F. 1976. The systematics of *Physalis pumila* ssp. *hispida* (Solanaceae) Systematic Botany (1): 188-193.

JAMES, C. W. 1961. Endemism in Florida. Brittonia 13: 225-244.

MENZEL, M. Y. 1951. The cytotaxonomy and genetics of *Physalis*. Proceedings of the American Philosophical Society 95(2): 132-183.

MITCHELL, R. S. 1963. Phytogeography and floristic survey of relic area in the Mariana Lowlands, Florida. American Midland Naturalist 69(2): 328-366.

NEI, M. 1972. Genetic distance between populations. American Naturalist 106: 283-292.

NEI, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89: 583-590.

PARKS, C. R., J. F. WENDEL, M. M. SEWELL, and Y. L. QIU. 1994. The significance of allozyme variation and introgression in the *Liriodendron tulipifera* complex (Magnoliaceae). American Journal of Botany 81(7): 878-889.

PLATT, W. J. and M. W. SCHWARTZ. 1990. Temperate hardwood forests. Pp. 194-229 in *Ecosystems of Florida*, eds. R. L. Meyers and J. J. Ewel. Orlando, Florida: University of Central Florida Press.

SALEHUZZAMAN, S. N. I. M., E. JACOBSEN, and R. G. F. VISSER. 1993. Isolation and characterization of a cDNA encoding granule-bound starch synthase in cassava (*Manihot esculenta* Crantz) and its antisense expression in potato. Plant Molecular Biology 23: 947-962.

SOLTIS, D. E. and P. R. SOLTIS, editors. 1989. Isozymes in plant biology. Portland, Oregon: Dioscorides Press.

SPSS INC. 1997. Systat 7.0. Chicago, Illinois.

SULLIVAN, J. R. 1984a. Systematic studies in *Physalis* (Solanaceae) Ph.D. dissertation, University of Oklahoma, Norman.

SULLIVAN, J. R. 1984b. Pollination biology of *Physalis viscosa* var. *cinerascens* (Solanaceae). American Journal of Botany 71(6): 815-820.

SWOFFORD, D. L. 1998. PAUP*: Phylogenetic Analysis Using Parsimony, version 4d64. Sinauer Associates, Sunderland, Massachusetts.

SWOFFORD, D. L. 1989. Manual for BIOSYS-1. Release 1.7. Illinois Natural History Survey, Champaign, Illinois.

UNIVERSAL IMAGING CORP. 1991. IMAGE-1, release 4.0. West Chester, Pennsylvania.

WATERFALL, U. T. 1967. A taxonomic study of the genus *Physalis* in North America north of Mexico. Rhodora 60: 107-173.

WOLFE, S. H., J. A. REIDNAUER, and D. B. MEANS. 1988. *An ecological characterization of the Florida panhandle*. Tallahassee, Florida: Florida State Dept. of Environmental Regulation.

WOLFRAM RESEARCH, INC. 1999. Mathematica 4.0. Champaign, Illinois.

CHAPTER 4, WORKS CITED:

HENDRYCH, R. 1989. *Physalis alkekengi*, in Europa und in der Tschechoslowakei besonders. Acta Universitatis Carolinae. Biologica 33: 1-42.

MENZEL, M. Y. 1951. The cytotaxonomy and genetics of *Physalis*. Proceedings of the American Philosophical Society 95(2): 132-183.

SULLIVAN, J. R. 1984. Systematic studies in *Physalis* (Solanaceae) Ph.D. dissertation, University of Oklahoma, Norman.

CHAPTER 5, WORKS CITED:

BRITTON, N. L. and A. BROWN. 1898. *An illustrated flora of the northern United States, Canada and the British possessions*. Vol. 3. New York: Charles Scribner's Sons. pp. 125-132.

BRITTON, N. L. and A. BROWN. 1913. *An illustrated flora of the northern United States*. 2nd edition. Vol.3. New York: Charles Scribner's Sons.

HITCHCOCK, A. S. and M. L. GREEN. 1929. Standard-species of Linnaean genera of Phanerogamae (1753-54). In: *International Botanical Congress*, Cambridge (England), 1930: Nomenclature. Proposals by British botanists. London: H.M.S.O. pp. 110-199.

GREUTER, W., J. MCNEILL, F. R. BARRIE, H. M. BURDET, V. DEMOULIN, T. S. FILGUEIRAS, D. H. NICHOLSON, P. C. SILVA, J. E. SKOG, P. TREHANE, N. J. TURLAND, and D. L. HAWKSWORTH. 2000. *International code of botanical nomenclature* (Saint Loius Code). Regnum vegetabile 138, Konigstein: Koeltz Scientific Books.

JARVIS, C. E., F. R. BARRIE, D. M. ALLAN, and J. L. REVEAL. 1993. A list of Linnaean generic names and their types. Regnum vegetabile 127, Konigstein: Koeltz Scientific Books.

LINNAEUS, C. 1753. *Species plantarum*. Stockholm. pp. 182-184.

MILLER, P. 1754. *The gardeners dictionary*. 4th ed. London: John and James Rivington.

MIONE, T., R. G. OLMSTEAD, R. K. JANSEN, and G. J. ANDERSON. 1994. Systematic implications of chloroplast DNA variation in *Jaltomata* and selected physaloid genera (Solanaceae). American Journal of Botany 81(7): 912-918.

OLMSTEAD, R. G., J. A. SWEERE, R. E. SPANGLER, L. BOHS and J. D. PALMER. 1999. Phylogeny and provisional classification of the Solanaceae based on chloroplast data. Pp. 111-137 in *Solanaceae IV: Advances in biology and utilization*, eds. M. Nee, D. E. Symon, R. N. Lester and J. P. Jessop. Kew: Royal Botanic Gardens.

RYDBERG, P. A. 1896. The North American species of *Physalis* and related genera. Memoirs of the Torrey Botanical Club 4(5): 297-372.

TREHANE, P. 1995. (1172) Proposal to conserve *Chrysanthemum* L. with a conserved type (Compositae). Taxon 44: 439-441.

WIJNANDS, D. O. and J. J. BOS. 1986. (856) Proposal to conserve 7436 *Petunia* with *P. nyctaginiflora* as typ. cons. (Solanaceae). Taxon 35: 74

BIOGRAPHY

Mary Kathryn ("Maggie") Whitson was born in Columbus, Mississippi on May 13, 1972. She received a B.S. in botany from the University of Florida, and graduated with highest honors in 1994.

Before applying to graduate school, she took two horticultural internships, the first at Selby Gardens in Sarasota, Florida, and the second at Longwood Gardens, in Kennett Square, Pennsylvania. In the summer of 1995, she collected plants for the Flora of La Selva Project, at the La Selva Biological Station in Costa Rica. She began graduate school in the Duke Botany Department in the fall of 1995.

Fellowships and awards:

A. W. Mellon Foundation and the Duke University plant systematics program: grant to study biosystematics of the genus *Physalis*.

Sigma Xi: Grant In Aid of Research, to study ITS variation within *Physalis* species.

Duke University Graduate School: Aleane Webb Dissertation Research Grant.

A. W. Mellon Foundation and the Duke University plant systematics program: fellowship, for completion of dissertation work, Spring 2000-Fall 2001.

American Society of Plant Taxonomists: 2001 Cooley Award for the best student/young Ph.D. talk in plant systematics.