Phylogenetic and phylogenomic studies of wild onions (Allium, Amaryllidaceae) at three taxonomic scales

---

A Dissertation

Presented to

The Faculty of the Graduate School

At the University of Missouri

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

By

ERICA JEAN WHEELER

J. Chris Pires, Dissertation Advisor

December 2011

The undersigned, appointed by the dean of the Graduate School,

have examined the Dissertation entitled

PHYLOGENETIC AND PHYLOGENOMIC STUDIES OF WILD ONIONS (ALLIUM, AMARYLLIDACEAE) AT THREE TAXONOMIC SCALES

Presented by Erica Jean Wheeler

A candidate for the degree of

Doctor of Philosophy

And hereby certify that, in their opinion, it is worthy of acceptance.

J. Chris Pires

Lori S. Eggert

Emmanuel Liscum

Stephen G. Pallardy

DEDICATION

I would like to dedicate this work to my Mum, Anne Wheeler, who supported me through some difficult times and who celebrated all of the small achievements along the way. Thanks Mum!

ACKNOWLEDGEMENTS

This research would not have been possible without the guidance and support of a number of people. I'd first like to thank Chris Pires for giving me the opportunity to pursue my interest in the genus *Allium*, and for supporting my ideas both intellectually and financially. Thanks also to the other members of my committee, Lori Eggert, Mannie Liscum and Steve Pallardy, who provided useful comments throughout the process. Dale McNeal (UOP), the North American *Allium* expert, inspired and encouraged me to continue research in the genus, was instrumental to the success of my field collections, and to the interpretation of the phylogeny of *Amerallium* species in North America. Saeideh Mashayekhi (RSABG) and I collaborated on many aspects of the *Amerallium* phylogeny project and I thank her for her hard work and collegiality.

During much of my time at MU, I was supported by a Life Sciences Fellowship, which made it possible for me to spend springs in the field for the first four years of my degree. During my last year and a half at MU, I was supported by a teaching assistantship in the Division of Biological Sciences. Much of the cost of my lab work was financed by an NSF Monocot Tree of Life Grant to Chris Pires. I also received student research grants from the American Philosophical Society, the American Society of Plant Taxonomists, the Missouri Native Plant Society, the Oregon Native Plant Society and the Washington Native Plant Society.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

My dissertation research centers on questions about evolutionary relationships among populations, species and subgeneric clades of wild onions (*Allium*), and on the inferences about evolutionary processes that can be drawn from phylogenies. I used standard molecular phylogenetic and emerging phylogenomic methods to address these questions. In chapter two, I conducted a molecular systematic study of subgenus *Amerallium* in North America. Eighty-four species of *Allium* are native to North America, 81 of which belong to the subgenus *Amerallium*. Based on morphological similarity, these species have been divided into eight informal alliances. To test the hypothesis that these morphological alliances reflect shared evolutionary history among species, I reconstructed the phylogeny of 74/81 North American *Amerallium* species based on DNA sequence variation in nuclear and chloroplast loci. Morphological alliances are largely congruent with the clades recovered in the molecular phylogenetic analysis. However, only three of eight alliances are strictly monophyletic. There is evidence of character reversal in bulb coat structure, a feature that has been important in delimiting morphological groups. In chapter three, I investigated multiple origins of pseudovivipary in the *Allium canadense* varietal complex. Pseudovivipary is a relatively rare reproductive strategy in angiosperms characterized by the replacement of flowers by asexual bulbils in the inflorescence. Whether this reproductive strategy evolves rarely or commonly within species has not been well-explored in a phylogenetic context. The *Allium canadense* varietal complex includes five floriferous varieties, and the pseudoviviparous variety *A. canadense* var. *canadense*. To test the hypothesis that pseudovivipary has evolved multiple times from sexual varieties in this species, I reconstructed the phylogeny of 119 populations of *A. canadense* using DNA sequence variation in the ETS region of the nuclear ribosomal RNA repeat. Results show that pseudovivipary has evolved at least six times in this species, and suggest that the transition to asexuality occurs commonly,

rather than rarely. Cytogenetic studies also showed that the transition to asexuality is accompanied by the transition to polyploidy. Although it is has been suspected for some time, this study is only the second to demonstrate that pseudovivipary is a labile trait in some angiosperm species. In chapter four, I conducted a chloroplast phylogenomic study of the genus *Allium* using genome survey sequencing methods. The genus *Allium* is currently divided into 15 subgenera and 72 sections based on nuclear ITS DNA sequences. Many relationships among subgenera and sections are resolved and well-supported. However, there are outstanding taxonomic questions in the genus that will require a significant increase in the number of phylogenetic characters. Next generation sequencing technologies are rapidly being applied to unresolved taxonomic problems in plants because they deliver megabases of data, including entire plastid genome sequences. I sequenced 18 *Allium* species representing eleven subgenera using genome survey sequencing (GSS) methods. My goals were to 1) assess the ability of GSS methods to recover entire plastid genome sequences in a taxon with a large nuclear genome, 2) provide the first intrageneric phylogenetic study in a non-grass monocot based on GSS methods, and 3) generate genomic resources that can be used in future phylogenetic and molecular evolution studies in the genus *Allium*. GSS methods recovered sufficient plastid sequences to assemble complete plastid genomes in all samples. Eighty protein coding genes were recovered, 79 of which were used to reconstruct the phylogeny of these taxa. Relationships among subgenera are largely congruent with previously published studies, with a couple of exceptions that may be the result of differences in taxon sampling density. All but three clades were supported by 100% bootstrap support. Although the aim of this study was not to resolve any particular taxonomic issues in the genus *Allium*, I hope its success will provide an impetus for *Allium* systematists in other parts of the world to embark on GSS projects aimed at resolving outstanding taxonomic questions. The data generated during this study can be used in future phylogenetic and molecular evolution studies within the genus *Allium*.

CHAPTER 1 INTRODUCTION

Molecular systematics of plants

The field of modern plant systematics is dedicated to 1) compiling an inventory of the world's plants, 2) producing a classification system that reflects plant evolution, 3) understanding the great diversity of the botanical world, and 4) communicating this information to the scientific community and to the general public (Woodward, 1997). The skills of plant systematists are of increasing importance in the 21st century, as human impacts on natural systems threaten to drive species to extinction, many of which remain unknown to science. World governments, through the Convention on Biodiversity (United Nations, 1992), have identified the lack of trained taxonomists (a subset of systematists) as a significant impediment to the sound management of biodiversity. In an effort to remedy this, the United Nations Environment Programme has instituted a Global Taxonomy Initiative, which provides taxonomic training to locals in developing countries, where the need is often the greatest.

Until relatively recently, the tools used to construct classifications have included descriptive and comparative studies of plant anatomy, morphology, palynology (pollen structure), cytogenetics (chromosome number and structure), micromolecules (flavonoids, terpenoids) and macromolecules (proteins). Over the past twenty years, the advent of automated DNA sequencing has revolutionized the field of plant systematics. Not only has the number of characters that can be used for comparison among taxa increased substantially, but since DNA sequence variation is historically ordered, it is now possible to reconstruct the history of groups of species over time. Today, analysis based on DNA sequence variation is the standard tool used to infer the phylogeny (evolutionary history) of groups of plants and modern classification systems often rely, at least in part, on results from molecular phylogenetic analysis.

The real power of DNA based phylogenies lies in comparing relationships recovered by these tools with classification schemes that have been hypothesized based on phenotypic characters. When these separate lines of evidence are congruent, the phenotypic characters used to construct a classification system truly reflect the genealogical relationships among taxa. However, when the DNA and phenotypic evidence are not congruent, this suggests that the species groupings suggested by phenotype may be artificial. Molecular phylogenetic studies in various angiosperm lineages have demonstrated that morphological classification systems do not always track genealogical relationships. For example, fruit morphology and seed embryo type had been used almost exclusively to infer relationships at all taxonomic levels in the mustard family, Brassicaceae; however, molecular phylogenetic analyses showed that these characters are highly homoplastic (have evolved repeatedly in unrelated taxa) and therefore classification systems based on these characters are artificial (reviewed in Al-Shehbaz, Beilstein, and Kellogg, 2006). This research resulted in a new classification system for the family and to the identification of other phenotypic characters, such as trichome morphology, that track the evolutionary history of the group better than fruit morphology and seed embryo type (Al-Shehbaz, Beilstein, and Kellogg, 2006).

Beyond informing classification systems, molecular phylogenetic methods can also be powerful tools in investigations of patterns and processes at lower taxonomic levels, within and among populations of the same species, for example. In angiosperms, intraspecific phylogenetic analysis has been key to demonstrating that variation in reproductive mode within a species (e.g., sexual seed production vs. asexual seed production) is sometimes the result of multiple transitions between these modes (Thompson and Whitton, 2006; Whitton et al., 2008). Similarly, the recurrent formation of polyploids within angiosperm species has also been documented using molecular phylogenetic methods (reviewed in Soltis and Soltis, 1993).

My interest in *Allium*

My interest in the genus *Allium* began when I was a Masters student at the University of Victoria, where I investigated the geographic distribution of genetic and chromosomal variation in the slim-leaf onion (*Allium amplectens*) (Wheeler, 2006). In the field, I was charmed by the search for populations of this species that led me through the foothills of the Sierra Nevada and the Coast Range of northern California. In the lab, I was thrilled to see my labor produce tangible and interesting results about the evolutionary history of this species. At this point in time, a classification system for North American *Allium* species had been established based on morphology (Saghir, Mann, and Ownbey, 1966), but no one had tested these hypotheses with molecular systematic methods. A well-resolved and well-supported species phylogeny would provide an evolutionary context upon which future studies could be based. A PhD student was born.

As is the case for many systematists, my research interests are driven by questions that arise from the close observation of a particular taxonomic group. What ties the three research projects in my dissertation together is my interest in *Allium* in general, and in particular, the evolution and systematics of the genus. Beyond contributing to a sound classification system for North American *Allium*, I am also interested in understanding the evolutionary processes that drive speciation in the genus, the evolution of asexual reproduction in some species, and the pathways to and consequences of polyploidy. If there is another unifying theme in my research, it is the approach I take to answering outstanding questions in the genus *Allium*: throughout this dissertation I use molecular phylogenetic and phylogenomic methods to investigate evolutionary relationships among populations, species, and intrageneric clades.

An introduction to the genus *Allium*

The genus *Allium* has had a place in human cuisine and folklore for millennia. Sculptural and painted representations from ancient Egypt suggest that onion (*A. cepa* L.), garlic (*A. sativum*

L.), and leek (*A. porrum* L.) were already cultivated at this time (Fritsch and Friesen, 2002), making these species some of the earliest to be domesticated as crops. Presently, at least 20 *Allium* species are cultivated as edible crops (Burba and Galmarini, 1994) and some edible species are collected from the wild including the European ramson (*A. ursinum* L.) and the North American ramp (*A. tricoccum* Solander). The humble onion has not only provided sustenance and flavor in the lives of humans, but has also been an inspiration for poetry. Among many others, two Nobel laureates in literature, Pablo Neruda (1971) and Wislawa Szymborska (1996), have turned their pens to writing about the onion (Neruda, 1994; Szymborska, 1998).

Plants in the genus *Allium* L. (Amaryllidaceae, Allioideae) are herbaceous geophytes, characterized by true bulbs, which are sometimes borne on rhizomes, and with a familiar onion or garlic odor and flavor (Figure 1.1). *Allium* is among the largest of the petaloid monocot genera with close 850 species currently recognized (N. Friesen, pers. comm., 2010). These species are distributed very broadly and almost exclusively in the northern hemisphere (Fritsch and Friesen, 2002). Many *Allium* species are xerophytic 'drought-avoiders' withstand drought conditions through dormancy. *Allium* is especially diverse in Mediterranean climates characterized by warm, dry summers and cool, wet winters. The center of diversity for the genus is in the Old World, where species richness is particularly high in central Asia and the Mediterranean Basin. A second less pronounced center of diversity occurs in the New World, in western North America (Figure 1.2). *Allium* is considered to be an ancient lineage which had already diversified in the Tertiary (Hanelt et al., 1992). This is based on the broad geographic distribution of centers of diversity among continents, and on unusually high genetic distances reported within the genus, i.e., distances that generally characterize angiosperm families or subfamilies rather than genera (Hanelt et al., 1992; Friesen, Fritsch, and Blattner, 2006).

Current taxonomy places *Allium* along with 14 other genera in the family Amaryllidaceae, subfamily Allioideae (Chase, Reveal, and Fay, 2009). Molecular phylogenies

demonstrate monophyly of *Allium* (Friesen, Fritsch, and Blattner, 2006; Nguyen, Driscoll, and Specht, 2008; Li et al., 2010). Within *Allium*, 15 subgenera and 72 sections are currently recognized (Friesen, Fritsch, and Blattner, 2006). Both morphological studies and molecular phylogenetic analyses support the division of the genus *Allium* into three evolutionary lines (Fritsch and Friesen, 2002; Friesen, Fritsch, and Blattner, 2006). The first evolutionary line includes the three subgenera *Nectaroscordum*, *Microscordum*, and *Amerallium*. The second evolutionary line includes the five subgenera *Caloscordum*, *Anguinum*, *Vvedenskya*, *Porphyroprason*, and *Melanocrommyum*. The third evolutionary line includes the seven subgenera *Butomissa*, *Cyathophora*, *Rhizirideum*, *Allium*, *Cepa*, *Reticulatobulbosa*, and *Polyprason*.

The scope of this dissertation

The three data chapters presented here stand alone as independent studies, each aimed at investigating patterns of genetic variation at a different taxonomic scale in the genus *Allium* and at answering different sets of questions. Chapter two is an interspecific molecular phylogenetic study of species in the subgenus *Amerallium* in North America, and aims to test the classification system based on morphological characters proposed by Marion Ownbey (Saghir, Mann, and Ownbey, 1966). Chapter three is an intraspecific molecular phylogeographic study of populations in the *Allium canadense* varietal complex of eastern and central North America. This study aims to test the hypotheses, also of Marion Ownbey, that a relatively rare form of asexual reproduction, pseudovivipary, has evolved numerous times in this species, and is correlated with the transition to polyploidy (Ownbey and Aase, 1955). Chapter four explores the use of Next Generation sequencing technology to recover entire plastid genome sequences from species in the genus *Allium*, and to reconstruct the phylogeny of the genus using 18 species, representing different subgenera and sections within subgenera from across the genus.

LITERATURE CITED

AL-SHEHBAZ, I. A., M. A. BEILSTEIN, and E. A. KELLOGG. 2006. Systematics and phylogeny of the Brassicaceae (Cruciferae): An overview. *Plant Systematics and Evolution* 259: 89-120.

BURBA, J. L., and C. R. GALMARINI. 1994. First international symposium on edible Alliaceae. International Society for Horticultural Science, Mendoza, Argentina.

CHASE, M. W., J. L. REVEAL, and M. F. FAY. 2009. A subfamilial classification for the expanded asparagalean families Amaryllidaceae, Asparagaceae and Xanthorrhoeaceae. *Botanical Journal of the Linnean Society* 161: 132-136.

FRIESEN, N., R. M. FRITSCH, and F. R. BLATTNER. 2006. Phylogeny and new intrageneric classification of *Allium* L. (Alliaceae) based on nuclear ribosomal DNA ITS sequences. *Aliso* 22: 372-395.

FRITSCH, R. M., and N. FRIESEN. 2002. Evolution, domestication and taxonomy. *In* H. D. Rabinovich AND L. Currah [eds.], *Allium* Crop Science – Recent Advances, vol. 1, 5-30. CABI Publishing.

HANELT, P., J. SCHULTZE-MOTEL, R. M. FRITSCH, J. KRUSE, H. I. MAAB, H. OHLE, and K. PISTRICK. 1992. Infrageneric grouping of *Allium* - the Gatersleben approach. The genus *Allium*: taxonomic problems and genetic resources, Gatersleben, Germany: 107-123.

LI, Q. Q., S. D. ZHOU, X. J. HE, Y. YU, Y. C. ZHANG, and X. Q. WEI. 2010. Phylogeny and biogeography of *Allium* (Amaryllidaceae: Allieae) based on nuclear ribosomal internal transcribed spacer and chloroplast rps16 sequences, focusing on the inclusion of species endemic to China. *Annals of Botany* 106: 709-733.

NATIONS, U. 1992. Convention on biological diversity [accessed November 17, 2011.

NERUDA, P. 1994. Odes to common things. Little, Brown and Company, Boston.

NGUYEN, N. H., H. E. DRISCOLL, and C. D. SPECHT. 2008. A molecular phylogeny of the wild onions (*Allium*; Alliaceae) with a focus on the western North American center of diversity. *Molecular Phylogenetics and Evolution* 47: 1157-1172.

OWNBEY, M., and H. C. AASE. 1955. Cytotaxonomic studies in *Allium* L.: the *Allium canadense* alliance, Research Studies of the State College of Washington, 1-106. State College of Washington.

SAGHIR, A. R. B., L. K. MANN, and M. OWNBEY. 1966. Composition of volatiles in relation to taxonomy of American *Alliums*. *American Journal of Botany* 53: 477-484.

SOLTIS, D. E., and P. S. SOLTIS. 1993. Molecular data and the dynamic nature of polyploidy. *Critical Reviews in Plant Sciences* 12: 243-273.

SZYMBORSKA, W. 1998. Poems New and Collected, 1957-1997. Harcourt Brace, New York.

THOMPSON, S. L., and J. WHITTON. 2006. Patterns of recurrent evolution and geographic parthenogenesis within apomictic polyploid Easter daises (*Townsendia hookeri*). *Molecular Ecology* 15: 3389-3400.

WHEELER, E. J. 2006. Autopolyploid evolution and phylogeography of the slim-leaf onion (*Allium amplectens*, Alliaceae)MSc, University of Victoria, Victoria.

WHITTON, J., C. J. SEARS, E. J. BAACK, and S. P. OTTO. 2008. The dynamic nature of apomixis in the angiosperms. *International Journal of Plant Sciences* 169: 169-182.

WOODWARD, D. W. 1997. Contemporary plant systematics. Second ed. Andrews University Press, Barrien Springs, Michigan.

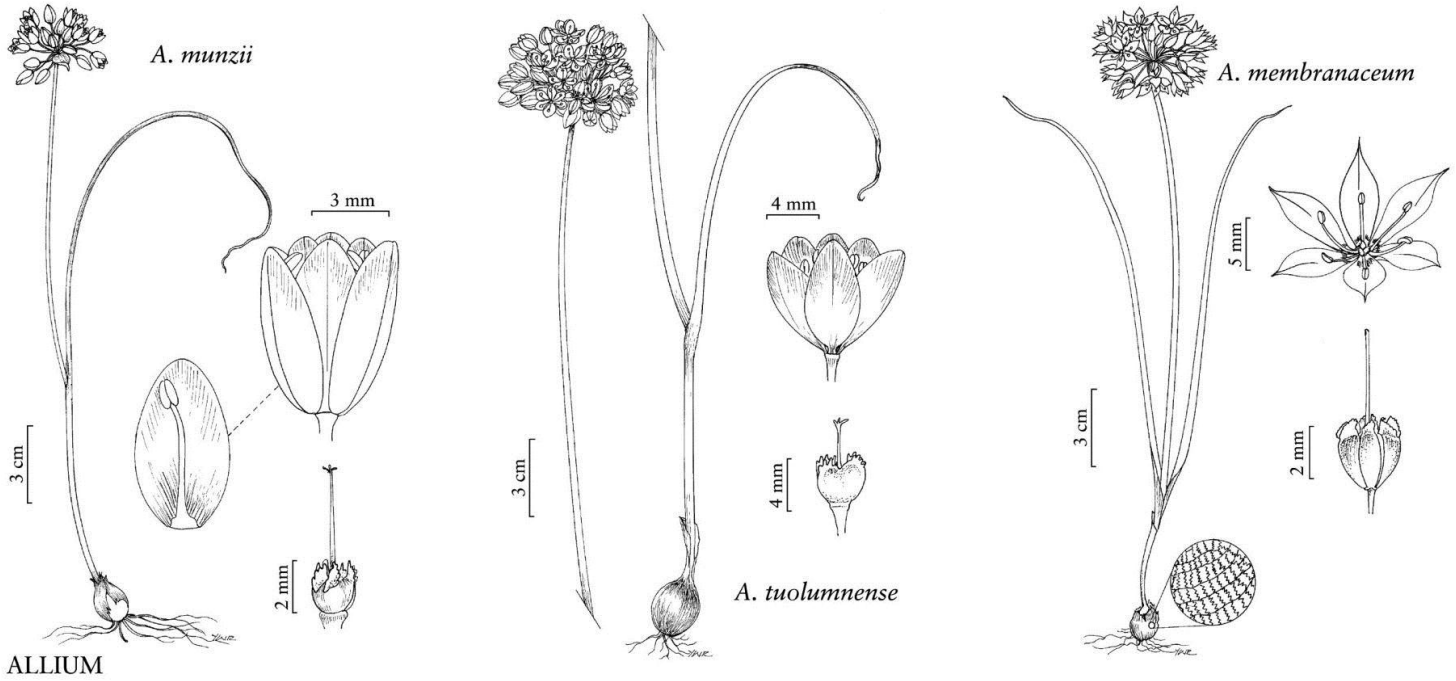

Figure 1.1 Illustrations of the morphology and general growth form of some *Allium* species from North America. From the Flora of North America, V 26 (2002).

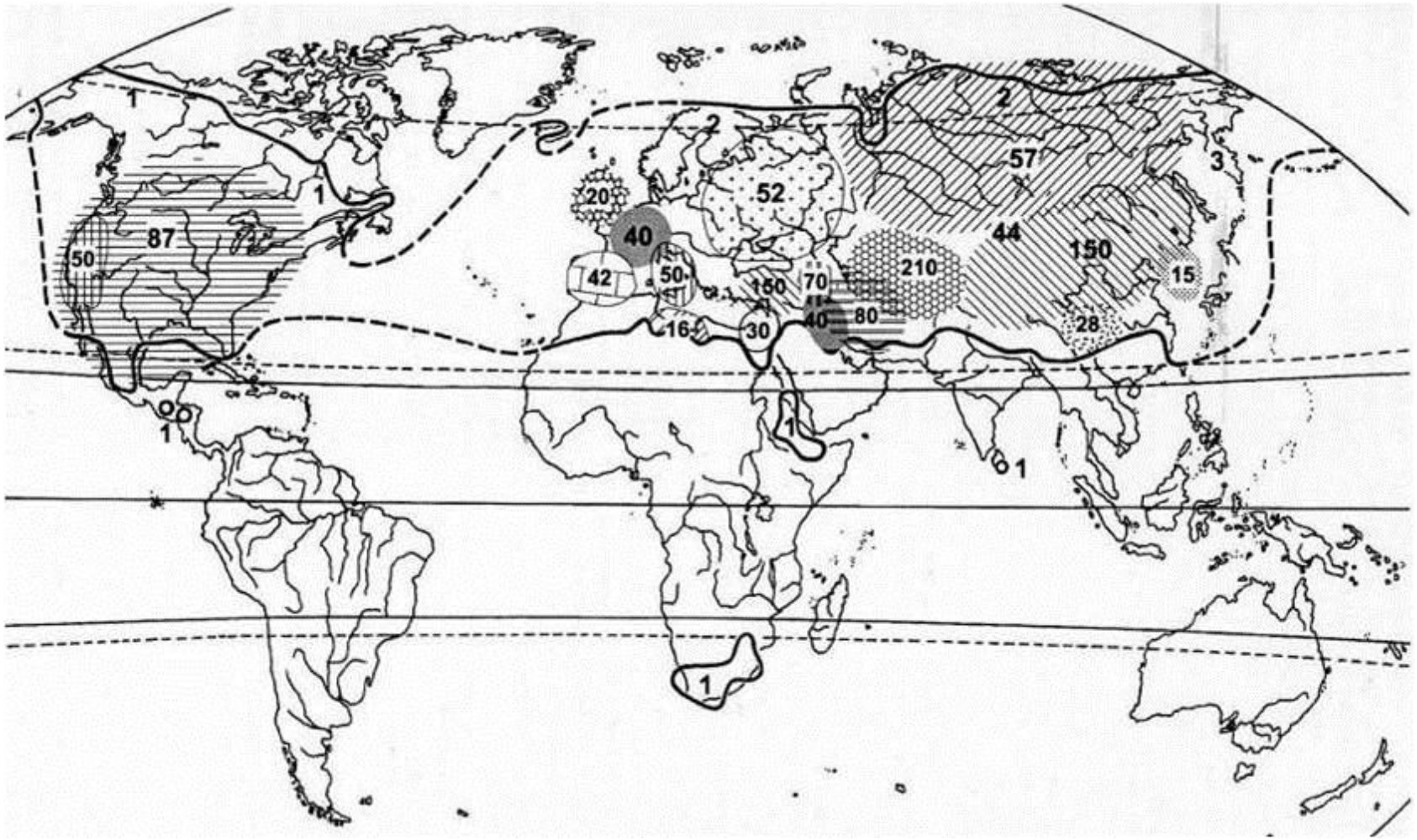

Figure 1.2 Global distribution of wild species of the genus *Allium*. The numbers on the map indicate the number of species found in each region. From Fritsch and Friesen, 2002.

CHAPTER 2

MOLECULAR SYSTEMATICS OF SUBGENUS *AMERALLIUM* IN NORTH AMERICA

ABSTRACT

In North America north of Mexico, wild onions (*Allium*, Amaryllidaceae) are represented by 84 native species, 81 of which belong to subgenus *Amerallium*. Based on morphology, these species have been divided into eight informal taxonomic 'alliances' hypothesized to represent shared evolutionary history among species. The main aim of this research was to test the monophyly of the alliances with molecular phylogenetic methods. We sampled 74 *Amerallium* species north of Mexico and two Mexican endemics and constructed a molecular phylogeny of subgenus *Amerallium* in North America based on predominantly non-coding sequences from two nuclear ribosomal RNA regions (ITS and ETS) and two chloroplast regions (*trnL–F* and *rpL32–trnL*). Most clades are well supported in analyses of nuclear data and when nuclear and chloroplast data are combined. However the chloroplast data alone did not produce a well-resolved or well-supported tree. The alliances are largely supported by the molecular phylogenetic data, although strict monophyly was observed in only three of eight alliances. Morphological characters used to delimit groups are generally well-founded although there is some evidence of character reversal in bulb-coat structure. There is considerable geographic structure to the distribution of the two main clades, which are largely separated by the Rocky Mountains.

INTRODUCTION

Subgenus *Amerallium*

The ~120 species in subgenus *Amerallium* are broadly distributed with centers of diversity in the Mediterranean Basin, eastern Asia and North America. These species are characterized by having a predominant base chromosome number of x=7, leaves with a single row of vascular bundles (Traub, 1968c) and subepidermal laticifers (Fritsch, 1988). Based on molecular phylogenetic analyses, *Amerallium* is monophyletic and along with subgenera *Nectaroscordum* and *Microscordum*, is sister to the rest of the genus and constitutes the first evolutionary line in *Allium*, (Dubouzet and Shinoda, 1999; Friesen, Fritsch, and Blattner, 2006). Recently, broader sampling of North American (Nguyen, Driscoll, and Specht, 2008) and Chinese species (Li et al., 2010) has confirmed that subgen. *Amerallium* is a monophyletic lineage. Six sections are recognized in the Old World and four sections in the New World (Traub, 1968c).

Amerallium is most diverse in North America with 81 species and 32 varieties currently recognized in the *Flora of North America, North of Mexico* (McNeal and Jacobsen, 2002). Examples of some endemic North American species are shown in Fig. 1.1. Approximately 26 species are reported for Mexico (Serna and Lopez Ferrari, 1992), 13 of which do not occur in the U.S. or Canada. New World *Amerallium* species form a clade that is sister to the Old World members of the subgenus (Dubouzet and Shinoda, 1999). This relationship is partly supported by cytogenetic evidence in that New World *Amerallium* species have a base chromosome number of x=7, whereas Old World species have a base chromosome number of x=7, 8, 9 or 10 (Friesen, Fritsch, and Blattner, 2006). Three North American *Allium* species belong to different subgenera that are otherwise centered in the Old World and have a base chromosome number of x=8: *A. tricoccum* Solander and *A. victorialis* L. (subgen. *Anguinum* (G. Don ex Koch) N. Friesen), and

the circumboreal *A. schoenoprasm* L. (subgen. *Cepa* (Mill.) Radic´) (Friesen, Fritsch, and Blattner, 2006).

Based on molecular phylogenetic evidence, New World *Amerallium* evolved from a common Old World ancestor (Hanelt et al., 1992; Dubouzet and Shinoda, 1999; Friesen, Fritsch, and Blattner, 2006; Nguyen, Driscoll, and Specht, 2008; Li et al., 2010) and diversified across the North American landscape to produce the current diversity and geographic distribution of species. Given the current geographic distribution of *Amerallium* species, it is most parsimonious to conjecture that New World species evolved from a common ancestor in the Mediterranean region or in eastern Asia. Hanelt et al. (1992) postulated an Asian origin and dispersal via the Bering land bridge. Recent biogeographic analysis also suggests that the common ancestor of North American species originated from high latitudes in eastern Asia and dispersed to the New World via the Bering land bridge (Li et al., 2010). However, a Mediterranean origin for this subgenus is also possible given the diversity of species in this region. Today *Amerallium* species cover most of the continent south of the 53rd parallel ranging through the oak hillsides of California and Oregon, deserts of Nevada and Texas, alpine meadows of Utah and Idaho, prairies of Nebraska and Manitoba, and forest glades of Missouri and Arkansas. Because of its broad distribution and diverse habitats, this group of plants is well suited to investigations of historical biogeography in the North American continent.

Classification of North American *Amerallium* species

Currently, there are two parallel classification systems used in reference to the North American *Amerallium* species. First, species north of Mexico and two Mexican endemics were divided into eight taxonomic 'alliances' based on morphological similarity: the *A. acuminatum*, *A. campanulatum*, *A. canadense*, *A. cernuum*, *A. falcifolium*, *A. kunthii*, *A. sanbornii* and *A. validum* alliances (Table 1.1). These alliances are not formally described taxa but are hypotheses of evolutionary relationship first proposed by Marion Ownbey in Saghir et al. (1966), hereafter

referred to as 'Ownbey's alliances'. Based largely on these alliances, Traub (1968c) formally arranged species in North America into four sections: *Amerallium* Traub, *Caulorhizideum* Traub, *Lophioprason* Traub and *Rhopetoprason* Traub. Within sect. *Lophioprason*, Traub recognized five of Ownbey's alliances as subsections (*Acuminata* Ownbey ex Traub, *Campanulata* Ownbey ex Traub, *Cernua* Ownbey ex Traub, *Falcifolia* Ownbey ex Traub, *Sanborniana* Owbey ex Traub,) and added subsect. *Bolanderiana* Traub. Within sect. *Amerallium*, Traub raised one alliance to subsectional status (*Canadensia* Ownbey ex Traub) and added subsect. *Mexicana* Traub (see Table 1.1 for a comparison of the two classification systems). This formal system has not been widely accepted by *Allium* taxonomists in North America due to lack of characters that unequivocally unite sections (McNeal, 1992). Throughout this paper, we use the alliance names of Ownbey to refer to groups of putatively closely related species. Based on the greatest sampling to date (46 North American *Amerallium* species), Nguyen et al. (2008) showed preliminary molecular phylogenetic evidence (ITS + ETS) that the *A. acuminatum, A. campanulatum, A. canadense* and *A. falcifolium* alliances are monophyletic but that the *A. cernuum, A. sanbornii* and *A. validum* alliances are not. Because only one species from the *A. kunthii* alliance has been sampled to date, no information about the monophyly of this alliance is available.

Morphological characters that have been used to circumscribe taxa within North American *Amerallium* include bulb morphology, in particular the structure of the outer bulb coat, which is often important for species identification (McNeal, 1973). Within most North American *Amerallium* species, the bulb coats are described as cellular-reticulate or fibrous-reticulate. These terms refer to whether the outer bulb coat is retained as a closed-cell, chartaceous 'skin' which forms an effective barrier to moisture, or an open-celled 'mesh' which likely has less moisture-retaining properties. Among species with cellular-reticulate bulb coats there is also considerable variation in the patterns made by cells on the inner surface of the outer bulb coat. These

characters are particularly important for identifying species in the *A. acuminatum* alliance. In addition to the cellular and fibrous-reticulate bulb coats, some species in the *A. validum* alliance are characterized by the presence of thick, *Iris*-like rhizomes. The presence, shape and orientation of ovarian crests are also important for species identification throughout the genus in North America, as is the number and shape of leaves.

Aims

Our molecular systematic study expands the number of species and the geographic range sampled in previously published research to include 74 of 81 species north of Mexico, and two Mexican *Amerallium* endemics. Additionally, we expand on past studies in this part of the genus by including both nuclear and chloroplast markers in our analyses. Our main objective is to test the taxonomic and evolutionary hypotheses suggested by Ownbey's morphological alliances.

MATERIALS AND METHODS

Taxon sampling

We sampled one individual from each of 74 *Amerallium* species north of Mexico, two endemic Mexican species and five outgroup taxa (Table 1.2). Outgroup species represent five subgenera as delimited by Friesen et al. (2006): Old World *Amerallium* (*A. roseum* L.), *Anguinum* (*A. tricoccum* Solander), *Butomissa* (*A. tuberosum* Rottler ex Sprengel), *Cepa* (*A. cepa* L.) and *Allium* (*A. flavum* L.). Leaf material was collected from plants in the wild, botanical-garden collections and herbarium specimens. Vouchers are deposited at CPH, RSA and UMO (Table 1.2).

DNA extraction, PCR amplification and DNA sequencing

DNA was extracted from dried leaves using a modified version of the CTAB DNA extraction protocol for plants (Doyle and Doyle, 1987). Two nuclear ribosomal RNA repeat regions, ITS and ETS, and two chloroplast regions, *trnL–F* and *rpL32–trnL* were used in this study. The internal transcribed spacer region (ITS1, 5.8S, ITS2), hereafter referred to simply as ITS, was amplified using plant-specific primers ITSA and ITSB (Blattner, 1999) or the universal primers ITS4 and ITS5 (White et al., 1990) under the following conditions: initial DNA denaturation at 95°C for two minutes, followed by 40 amplification cycles of 95°C for 20 seconds, 55°C for 30 seconds 70°C for one minute and a final extension at 70°C for seven minutes. The external transcribed spacer region (ETS) of the rRNA repeat was amplified using the *Allium*-specific forward primer *Allium* ETSF (Nguyen, Driscoll, and Specht, 2008) and the reverse primer 18S-ETS (Baldwin and Markos, 1998) under the following conditions: initial denaturation at 95°C for five minutes followed by 40 amplification cycles of 94°C for one minute, 50°C for one minute, 72°C for two minutes and a final extension at 72°C for ten minutes. The *trnL–F* region (*trnL* intron, 3' exon and *trnL–trnF* intergenic spacer) was amplified with the forward primer Tab-C and the reverse primer Tab-F, or alternatively in two sections with primer combinations Tab-C/D and Tab-E/F (Taberlet et al., 1991). The *rpL32–trnL* intergenic spacer was amplified with the forward primer *rpL32–F* and the reverse primer *trnL* (UAG) (Shaw et al., 2007). Both chloroplast regions were amplified under the following conditions: initial denaturation at 80°C for five minutes followed by 30 amplification cycles of 95°C for one minute, 50°C for one minute, 65°C for four minutes and a final extension at 65°C for five minutes. Amplification products were inspected for the presence of a single band by electrophoresis on 1% agarose gels in TBE buffer and products were purified by PEG precipitation (Johnson and Soltis, 1994) or QiaQuick PCR purification columns (Qiagen, Valencia California). PCR products were directly sequenced using Applied Biosystems Inc.

(Foster City California) BigDye dideoxy terminator sequencing kits v. 3.1. Sequences were analyzed at MU on an ABI 3730 automated sequencer or at RSABG on an ABI 3130XL automated sequencer. For all gene regions, amplification primers described above were also used in sequencing reactions. In some cases, sequencing of the *trnL–F* region required the replacement of Tab-E with the internal primer *trnL* INT3F primer (Columbus et al., 2007). Forward and reverse sequences were assembled in the programs SeqMan (DNAStar Inc., Madison Wisconsin) or Sequencher™ v. 4.1.2 (Gene Codes Corporation, Ann Arbor Michigan) and assessed for quality before exportation for alignment.

Sequence alignment and indel coding

Sequences from each region were aligned using the on-line alignment tool Muscle (Edgar, 2004), refined by eye and then concatenated into a single partitioned dataset in the program MacClade v. 4.06 (Maddison and Maddison, 1992). Regions in the alignment where homology could not be unambiguously assigned were excluded from the analysis. Parsimony informative insertions and deletions (indels) in the ingroup were treated as simple gaps and coded as presence or absence characters (Simmons and Ochoterena, 2000)

Phylogenetic analyses

To explore potential conflicts between the evolutionary history of the nuclear and chloroplast genomes in *Allium*, we analyzed the nuclear and chloroplast datasets independently. To assess data partition congruence (ITS vs. ETS, *trnL–F* vs. *rpL32–trnL* and nuclear vs. chloroplast), we employed incongruence length difference (ILD) tests (Farris et al., 1994) in PAUP* v. 4.0b10 (Swofford, 2002) with 100 replicates. In addition, we employed the approximately unbiased (AU) test (Shimodaira, 2002) in CONSEL (Shimodaira and Hasegawa, 2001)to evaluate whether trees obtained from the nuclear dataset (ITS + ETS) and the chloroplast dataset (*trnL–F* + *rpL32–trnL*) were significantly different. We also combined the nuclear and

chloroplast datasets to produce a total-evidence tree. Maximum parsimony (MP) heuristic tree searches were conducted in PAUP* with the TBR option in effect and majority-rule consensus trees constructed from a maximum of 1000 trees saved. Based on a hierarchical likelihood ratio test implemented in Modeltest v. 3.06 (Posada and Crandall, 1998), the best nucleotide substitution model for the sequence data across all partitions was GTR + G + I. Using the GTR + G + I model for sequence data and a restriction-site model for indel data, Bayesian analysis was implemented using MrBayes v.3.1 (Huelsenbeck and Ronquist, 2001) employing 2 000 000 generations in all analyses except the chloroplast dataset, for which 5 000 000 generations were run. Two independent runs were performed for each analysis and trees were sampled every 100 generations. All trees recorded prior to the point where the standard deviation of split frequencies reached 0.010 were discarded as burn-in and 50% majority rule consensus trees were calculated from the remaining trees. Measures of clade support included BI posterior probabilities and MP bootstrap percentages. MP bootstrap analysis was performed in PAUP* with 1000 pseudoreplicates that each used the TBR option. During each pseudoreplicate, a maximum of 1000 trees was saved.

Geographic ranges of main clades

Clade maps were constructed for two main clades by layering electronic species range maps published in the *FNA* (McNeal and Jacobsen, 2002) in the program Metamorph v. 6.1 (Universal Imaging Corporation, Sunnyvale California) to produce a composite multi-species range map including all species in a clade. The perimeter of each clade distribution was estimated by tracing composite maps and these range limits were superimposed on the *FNA* base map.

RESULTS

DNA sequence variation

The combined four-locus alignment was 3164 base pairs (bp) long and included 69 coded indel characters. The multi-region alignment can be found in TreeBase and individual sequences are deposited in GenBank. Among ingroup taxa, 2064 characters were invariable, 307 characters were parsimony uninformative, and 793 characters were parsimony informative. The ITS region was 786 aligned positions including 21 indels (317 informative characters); ETS was 583 bp including 19 indels (290 informative characters); *trnL–F* was 727 bp including 10 indels (55 informative characters); and *rpL32–trnL* was 1068 bp including 19 indels (131 informative characters). The ensemble consistency index (CI; Kluge and Farris, 1969) for ITS, ETS, *trnL–F* and *rpL32–trnL* were 0.475, 0.525, 0.821 and 0.744, respectively. A 210 bp hypervariable region was excluded from the middle of *rpL32–trnL* because this region was unalignable.

Congruence of nuclear and chloroplast datasets

Results from the ILD tests indicate that all data partitions are significantly incongruent ($P < 0.05$ for all tests). However, the AU test indicates that there is no significant difference between the tree topologies generated from the nuclear and chloroplast datasets ($p < 0.05$). The topology of the nuclear tree is well-supported, but there is weak support for many clades in the chloroplast tree. Topologies of the nuclear and chloroplast trees are similar in most respects, with a few exceptions. However, most cases of topological incongruence are weakly supported (< 70% bootstrap support), and considering also results from the AU test result, we focus on the tree based on the combined dataset. The single case of topologically well-supported incongruence is the placement of *A. validum*. In the nuclear gene tree, this species is placed as sister to the clade containing the *A. canadense* and *A. kunthii* alliances and other members of the *A. validum* alliance

(94% MP bootstrap / 1.0 Bayesian posterior probability) whereas in the chloroplast tree, it is grouped with the other members of the *A. validum* alliance (93 / 1.0).

Phylogeny based on the combined dataset

Based on the combined dataset (ITS, ETS, *trnL–F* and *rpL32–trnL*), the topologies of trees produced by the MP and BI analyses were similar and Fig. 2.2 shows the Bayesian consensus tree. The topology of the combined tree is very similar to the nuclear tree and the addition of the chloroplast data increased support values for some nodes. The 76 species of North American *Amerallium* (including two Mexican endemics) are supported as a monophyletic group in the MP and BI analyses (100/1.0). *Allium tricoccum*, a New World representative of subgen. *Anguinum*, is not a member of this clade, as has been shown in other analyses (Klaas and Friesen, 2002; Friesen, Fritsch, and Blattner, 2006; Nguyen, Driscoll, and Specht, 2008). Within North American *Amerallium*, two main lineages are well supported. The larger clade contains five of eight alliances (82/1.0) and the smaller clade contains three of the eight alliances (100/1.0). Alliances that are monophyletic as circumscribed by Ownbey include the *A. acuminatum*, *A. campanulatum*, and *A. falcifolium* alliances. Alliances that are not monophyletic include the *A. canadense*, *A. cernuum*, *A. kunthii*, *A. sanbornii* and *A. validum* alliances.

DISCUSSION

Phylogeny

Our phylogeny recovers many of the same clades and relationships among clades as Nguyen et al. (2008) and supports many of their findings. Our expanded sampling of species outside of the California Floristic Province (CFP) provides additional information about the relationships of species and clades in other parts of the North American continent. In addition,

increased sampling of species within the CFP allows us to see some patterns in the geographic distribution of species and clades.

Within *Amerallium* in North America, the two main lineages show considerable geographic structure. The larger lineage includes species from five alliances (*A. acuminatum, A. campanulatum, A. cernuum, A. falcifolium* and *A. sanbornii*) and is equivalent to Traub's (1968c) sect. *Lophioprason*. This clade, hereafter referred to as the western clade, occurs largely west of the Rocky Mountains, especially in California, Oregon and Washington (Fig. 2.3). The exceptions are *A. cernuum* and *A. stellatum*, which occur throughout the western and central plains and as far as the Pacific Northwest and the eastern seaboard in the case of *A. cernuum.* The species in this clade are characterized by a membranous to chartaceous outer bulb coat, with or without cellular reticulations or parallel fibers. The smaller lineage consists of three alliances (*A. canadense, A. kunthii* and *A. validum*), parallel to Traub's (1968c) sections *Amerallium, Rhophetoprason* and *Caulorhizideum* respectively, and includes the two endemic Mexican species sampled. This clade, hereafter referred to as the eastern clade, occurs east or south of the Rocky Mountains especially in Texas and some species range into Mexico (Fig. 2.3). The exceptions are *A. validum*, which occurs at high elevations in California, Oregon, Washington and British Columbia, and *A. geyeri*, whose range extends west of the continental divide. Most species in this group are characterized by a fibrous-reticulate bulb coat or an *Iris*-like rhizome. Our molecular phylogenies generally support the hypothesis that bulb morphology characters are useful for delimiting some natural groups. In particular, the membranous to chartaceous bulb coat of the western lineage is a synapomorphy that separates this group from other species in North America . The geographic ranges of the western and eastern clades overlap along their contact zone in the southwestern states, the Rocky Mountains, and in the Pacific Northwest. Despite this geographic overlap, there is no evidence of genetic exchange between plants

belonging to these two lineages as all species are placed reliably in one lineage or the other across all data partitions.

Ownbey's alliances

In general, the combined molecular phylogeny supports the hypotheses of species relationship first proposed by Ownbey (Saghir, Mann, and Ownbey, 1966) and later by Traub (1968c). Below, each alliance is treated separately with respect to monophyly and defining characters.

Allium falcifolium alliance – The *A. falcifolium* alliance is the largest alliance, with 27 species (24 sampled) and four infraspecific taxa currently recognized. This group is strongly supported as monophyletic (100/1.0). A morphological synapomorphy that defines the *A. falcifolium* alliance is more or less sickle-shaped, or falcate, leaves. In many species, the scape and the leaves break off at ground level at senescence. Members of this alliance are xerophytes, distributed throughout the CFP, the Pacific Northwest, the Great Basin Desert, and on the western slopes of the Rocky Mountains. Within this alliance, two clades with distinct geographic distributions are supported. North of the Sierra Nevada and east of the Cascade Range is a clade of 13 species (*A. aaseae*, *A. brandegeei*, *A. columbianum*, *A. douglasii*, *A. fibrillum*, *A. lemmonii*, *A. macrum*, *A. madidum*, *A. nevii*, *A. robinsonii*, *A. scilloides*, *A. simillimum* and *A. tolmiei*). Throughout the CFP and northward along the Coast Range to coastal British Columbia is a clade of 11 species (*A. anceps*, *A. burlewii*, *A. cratericola*, *A. crenulatum*, *A. falcifolium*, *A. hoffmanii*, *A. obtusum*, *A. parvum*, *A. platycaule*, *A. tribracteatum* and *A. yosemitense*). No obvious morphological synapomorphies define these two clades. Based on their geographic distribution, we predict that two of the species we did not sample, *A. constrictum* and *A. punctum* will be placed in the northern clade. The third unsampled species, *A. siskiyouense*, probably belongs in the southern clade based on geographic distribution. These predictions are consistent with the phylogenetic position of these three species in Nguyen et al. (2008).

Allium sanbornii alliance – The *A. sanbornii* alliance contains 17 species (15 sampled) and ten infraspecific taxa, all of which are xerophytes. This alliance is paraphyletic since a clade of three species in the *A. campanulatum* alliance is nested within this group. This larger group is well-supported in the combined analyses (84/1.0). A morphological synapomorphy that defines the *A. sanbornii* alliance is the single leaf (two in *A. bigelovii*) that is round in cross-section (terete) as opposed to flat or channeled as found in most other North American *Amerallium* species, including species in the *A. campanulatum* alliance. In addition, all species have six prominent ovarian crests. There is distinct geographic structure to the distribution of clades and grades in this alliance. A paraphyletic grade of five species occurs east of the Sierra Nevada in California, Arizona and Nevada and into the Transverse Ranges of southern California (*A. atrorubens*, *A. bigelovii*, *A. monticola*, *A. nevadense* and *A. shevockii*). The ovarian crests in these species tend to be entire or, at most, emarginate. In California west of the Sierra Nevada is a clade of ten species (*A. abramsii*, *A. diabolense*, *A. fimbriatum*, *A. howellii*, *A. jepsonii*, *A. munzii*, *A. parishii*, *A. parryi*, *A. sharsmithiae* and *A. tuolumnense*). The ovarian crests in these species tend to be more or less fimbriate. Based on their geographic distribution, we predict that the species we did not sample in this alliance, *A. denticulatum* and *A. sanbornii*, will be resolved in the second group and this is supported by the placement of these species in Nguyen et al. (2008).

Allium campanulatum alliance – The *A. campanulatum* alliance contains three species (all three sampled) and two infraspecific taxa. This group is supported as monophyletic (100/1.0) but is nested within the *A.sanbornii* alliance, as mentioned above. The species in this group are recognized by having mostly two leaves and six conspicuous ovarian crests. Members of this alliance are more or less mesophytic, occurring sometimes as understory plants or on shady slopes, suggesting an ecological transition from the closely related xerophytes in the *A. sanbornii* alliance. The presence of six prominent ovarian crests is shared by this alliance and the *A.*

sanbornii alliance although the shape of these crests differs in the two alliances. *Allium campanulatum* and *A. membranaceum* are found largely within the CFP whereas *A. bisceptrum* is found east of the Sierra Nevada and into the Rocky Mountains.

Allium cernuum alliance – The *A. cernuum* alliance contains three species (all three sampled) and is paraphyletic as circumscribed by Ownbey. *Allium cernuum* and *A. stellatum* form a clade (100/1.0) that does not include the other member of this alliance, *A. haematochiton*. This latter species is sister to the clade that contains the *A. falcifolium*, *A. campanulatum* and *A. sanbornii* alliances (89/1.0). Species in the *A. cernuum* alliance have several leaves, crested ovaries and elongate bulbs. The leaf blade of *A. haematochiton* is terete (round in cross-section) whereas the leaves of *A. cernuum* and *A. stellatum* are flat or channeled. *Allium cernuum* and *A. stellatum* are mesophytic whereas *A. haematochiton* is xerophytic. In the U.S., *A. haematochiton* is restricted to southern California, and it also ranges south into Mexico. The geographic range of the clade containing *A. cernuum* and *A. stellatum* is very large, covering much of the continent east and north of California, Oregon and Nevada. It is interesting that the common ancestor of *A. cernuum*, *A. stellatum* and *A. haematochiton* gave rise to only two extant species (*A. cernuum* and *A. stellatum*) to the east and into the Great Plains. In comparison, this same common ancestor gave rise to 58 extant species (*A. acuminatum*, *A. falcifolium*, and *A. sanbornii* alliances, plus *A. haematochiton*) in the mountainous west.

Allium acuminatum alliance – In the *A. acuminatum* alliance, 13 species (all 13 sampled) and six infraspecific taxa are recognized. This alliance is well-supported as a clade (100/1.0) and is strongly marked by the development of the inner epidermis of outer bulb coats, which becomes thickened with cellular markings that are often species specific. Members of the *A. acuminatum* alliance are xerophytes. This alliance is most diverse in the CFP but also ranges north to British Columbia and east of the Cascade Range in the Great Basin and the Blue Mountains of Washington and Oregon.

Allium canadense alliance – The *A. canadense* alliance includes 12 species (all 12 sampled) and ten infraspecific taxa and is marked by the presence of a fibrous-reticulate bulb coat. The clade containing most of the species in the *A. canadense* alliance is well-supported (100/1.0) but this alliance is not monophyletic as predicted by morphology. First, *A. elmendorfii*, a species considered to be part of the *A. kunthii* alliance because it lacks a fibrous-reticulum, is placed in the *A. canadense* alliance. Second, *A. plummerae*, which was considered part of the *A. canadense* alliance because of the presence of the fibrous-reticulum, is placed in the *A. validum alliance*, suggesting a reversal of this bulb-coat character. In contrast to the alliances of the western lineage, there is little geographic structure to the distribution of clades in this group in that ranges of clades are overlapping in most cases.

Allium kunthii alliance – The *A. kunthii* alliance contains four species (three sampled). The placement of *A. elmendorfii* in the *A. canadense* alliance indicates that the group as described by Ownbey is not monophyletic. The other species sampled, *A. glandulosum* and *A. kunthii*, form a well-supported clade (100/1.0). The species in this alliance are similar to those of the *A. canadense* alliance except that fibers are absent from the bulb coats. The *A. kunthii* alliance occurs in the southwestern U.S. in the states of Texas, New Mexico and Arizona and in Mexico.

Allium validum alliance – The *A. validum* alliance consists of four species (three sampled). As circumscribed by Ownbey, the *A. validum* alliance is not monophyletic. Molecular phylogenetic evidence suggests that, despite morphological differences, *A. plummerae* is placed in the *A. validum* alliance rather than the *A. canadense* alliance as predicted by Ownbey. The clade including *A. brevistylum*, *A. eurotophyllum* and *A. plummerae* is well supported (99/1.0). In addition, in the nuclear and combined (Figure 2.2) analyses, *A. validum* falls outside of the alliance, sister to the rest of the eastern clade. In Nguyen et al. (2008), based on analysis of ITS and ETS sequences, *A. validum* is sister to the clade containing the *A. canadense* alliance and the remaining species in the *A. validum* alliance. However, in our chloroplast phylogeny, *A. validum*

is resolved with other members of the *A. validum* alliance (a possible explanation for this conflict is discussed below). *Allium goodingii*, the only species we did not sample from the *A. validum* alliance, was shown by Dubouzet and Shinoda (1999) to be sister to the species *A. validum* based on an analysis of ITS sequences. As Ownbey described, the species included in this alliance have *Iris*-type rhizomes, several leaves and grow in swampy meadows and wet places. Species in this alliance are distributed throughout the mountainous west.

Phylogenetic incongruence and potential ancient hybrid species

Although hybridization is not hypothesized to have played a large role in the evolutionary history of *Allium*, there is evidence that some wild species are of hybrid origin including the Siberian allopolyploid species *A. altyncolicum* (Friesen et al., 1997). Fertile hybrids have also been made by artificial crosses between closely related and distantly related agricultural species (reviewed in Chuda and Adamus, 2009) as a means to introduce desirable traits into vegetable and horticultural crops. Whether hybridization has played a role in the evolutionary history of subgen. *Amerallium* in North America has not been assessed from a molecular phylogenetic perspective to date. Evidence from chromosome counts suggests that 5% (4/81) of these species occur only as polyploids, 12% (10/81) occur as diploids and polyploids and the remaining 83% (67/81) occur as diploids only (McNeal and Jacobsen, 2002). Similar patterns have been found in other parts of the world, including in the *Allium* flora of Greece (Tzanoudakis, 1992). Whether polyploid species and species that maintain polyploids are allopolyploids resulting from interspecific hybridization and subsequent genome doubling, or are autopolyploids, is an outstanding question.

We did not find extensive phylogenetic incongruence in North American *Amerallium*. However, the one case we did find calls into question the placement of *A. validum* and indicates that it may be of hybrid origin. In the nuclear and combined analyses, this species is sister to rest of the eastern lineage. However, the chloroplast data support the inclusion of this species with

other members of the *A. validum* alliance. *Allium validum* occurs as a tetraploid (2n=28) and as an octoploid (2n=56), an unusual situation in North American *Amerallium*, since most species that contain polyploids also include diploid populations (McNeal and Jacobsen, 2002). Since no diploids have been found in this species, *A. validum* may be an example of a polyploid hybrid species. To investigate this possibility further, future research could include cytological studies of this species aimed at detecting contributions from two divergent genomes. An alternate explanation for this pattern of incongruence is incomplete lineage sorting in the chloroplast genome relative to the pattern observed in the nuclear genome.

Taxonomy

The informal alliances proposed by Ownbey (Saghir, Mann, and Ownbey, 1966) were later formalized by Traub (1968c) into taxonomic ranks of sections and subsections. Since these infrageneric taxa are predominantly based on Ownbey's alliances, the two systems are largely congruent. However, two conspicuous differences exist between them. First, rather than including it in subsect. *Acuminata* (= *A. acuminatum* alliance), Traub placed *A. bolanderi* into a subsection of its own (*Bolanderiana*) based on the presence of rhizomes upon which renewal bulbs grow. Given that the phylogenetic position of *A. bolanderi* renders subsect. *Acuminata* paraphyletic, and that other species in the subsect. *Acuminata* also have rhizomes (e.g., *A. unifolium*), we advocate reducing subsect. *Bolanderiana* to synonymy and treating *A. bolanderi* in subsect. *Acuminata*. Second, Traub erected a subsection (*Mexicana*) into which he placed most of the Mexican species. Since we have only sampled two Mexican endemics, we cannot comment on the validity of subsect. *Mexicana* except to say that the two endemic Mexican species in our study do not form a clade; *A. eurotophyllum* is placed in the *A. validum* alliance and *A. glandulosum* is placed in the *A. kunthii* alliance. Future research should include comprehensive sampling of Mexican taxa to ascertain their evolutionary relationship to other *Amerallium* species in North America. Traub erected *Mexicana* as a subsection of sect.

Amerallium, which is parallel to Ownbey's *A. canadense* alliance. Based on this, we might predict that Mexican species will be resolved within the *A. canadense* alliance.

Allium sanbornii and *A. campanulatum* alliances – The *A. sanbornii* alliance is rendered paraphyletic because the *A. campanulatum* alliance is nested within it. Employing the criterion of monophyletic taxa (above the species level), one way to resolve this issue is to subsume the *A. campanulatum* alliance under the *A. sanbornii* alliance. The resulting expanded *A. sanbornii* alliance would then consist of 20 species in total, including the three species from the *A. campanulatum* alliance, *A. brevistylum*, *A. campanulatum* and *A. membranaceum*. Although this makes sense from a molecular phylogenetic perspective, there are no obvious morphological characters that unite this group of species as a whole.

Allium haematochiton – In our combined data analysis, *A. haematochiton* is well-supported (89/1.0) as sister to the clade containing the *A. campanulatum*, *A. falcifolium* and *A. sanbornii* alliances but is not placed in a clade with the other species in the *A. cernuum* alliance (Fig. 2). This result is consistent with that of Nguyen et al. (2008), who suggest that this species could be placed in its own subsection. Based on the chloroplast data, although not well supported (64/0.79), this species is sister to the *A. acuminatum* alliance. Regardless of the data partition being analyzed, *A. haematochiton* is not placed in an alliance, unless it is considered part of an expanded *A. acuminatum* alliance based on evidence from the chloroplast phylogeny.

Allium validum – The taxonomic placement of this species is uncertain. Analysis of our chloroplast data places the species *A. validum* with other members of the *A. validum* alliance. However, the nuclear and combined analysis places this species as sister to the rest of the species in the eastern lineage. We recommend that *A. validum* be considered *incertae sedis*, of uncertain placement, until further research provides more information about the placement of this species.

Future research

Our study includes the most complete phylogenetic sampling of North American *Amerallium* species to date. In combination with the results from Nguyen et al. (2008), this study provides a fairly robust assessment of whether the species in each of Ownbey's alliances reflect natural, monophyletic groups. However, additional North American species remain to be sampled including *A. constrictum* (*A. falcifolium* alliance), *A. goodingii* (*A. validum* alliance), *A. rhizomatum* (*A. kunthii* alliance) and as many as 13 *Amerallium* species described from Mexico that do not occur in the U.S. or Canada (Serna and Lopez Ferrari, 1992). The validity and phylogenetic placement of these species needs to be assessed before we can have a complete picture of the evolutionary history of the genus *Allium* in North America.

ACKNOWLEDGEMENTS

This project was a collaborative effort between Erica J. Wheeler and Saeideh Mashayekhi, a PhD student at Rancho Santa Ana Botanical Garden in Claremont California. We were guided through this project by Dale W. McNeal, the taxonomic expert in North American *Allium*. Thank you to J. Travis Columbus for multiple rounds of excellent advice and manuscript editing. Thank you also to T. Arias, K. Hertweck, L. McDade, J. M. Porter and P. R. Steele for helpful comments on earlier versions of this manuscript. Thanks to K. Hertweck for conducting the AU test. The following people helped with field collections: B. Carr, J. Frazee, E. Keith, D. Knoke, T. Patrick, M. Powell and J. Smith. The following herbaria generously allowed us to remove leaf material from their specimens for DNA analysis: CPH, LL, NAU, RSA, SRSC and WS. We thank the National Parks Service and biologists at the Petrified Forest National Park and Craters of the Moon National Monument for permission to collect. In addition, we thank the California State Parks and biologists at Point Cabrillo Light Station State Historic Park for permission to collect. This research was supported by a National Science Foundation Grant (DEB 0829849), a University of Missouri Life Science Fellowship, Rancho Santa Ana Botanical Garden research awards and summer fellowships, and grants from the American Philosophical Society, the American Society of Plant Taxonomists, the Native Plant Society of Oregon and the Washington Native Plant Society.

LITERATURE CITED

BALDWIN, B. G., and S. MARKOS. 1998. Phylogenetic utility of the external transcribed spacer (ETS) of 18S-26S rDNA: congruence of ETS and ITS trees of *Calycadenia* (Compositae). *Molecular Phylogenetics and Evolution* 10: 449-463.

BLATTNER, F. 1999. Direct amplification of the entire ITS region from poorly preserved plant material using recombinant PCR. *Biotechniques* 27: 1180-1186.

CHUDA, A., and A. ADAMUS. 2009. Aspects of interspecific hybridization within edible Alliaceae. *Acta Physiologiae Plantarum* 31: 223-227.

COLUMBUS, J. T., R. CERROS-TLATILPA, M. S. KINNEY, M. E. SIQUEIROS-DELGADO, H. L. BELL, M. P. GRIFFITH, and N. F. REFULIO-RODRIGUEZ. 2007. Phylogenetics of Chloridoideae (Gramineae): A preliminary study based on nuclear ribosomal transcribed spacer and chloroplast trnL-F sequences. *Aliso* 23: 565-579.

DOYLE, J. J., and J. D. DOYLE. 1987. A rapid DNA isolation procedure from small quantities of fresh leaf tissues. *Phytochemical Bulletin* 19: 11-15.

DUBOUZET, J. G., and K. SHINODA. 1999. Relationships among Old and New World *Alliums* according to ITS DNA sequence analysis. *TAG Theoretical and Applied Genetics* 98: 422-433.

EDGAR, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl. Acids Res.* 32: 1792-1797.

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

FRIESEN, N., R. M. FRITSCH, and F. R. BLATTNER. 2006. Phylogeny and new intrageneric classification of *Allium* L. (Alliaceae) based on nuclear ribosomal DNA ITS sequences. *Aliso* 22: 372-395.

FRIESEN, N., N. BORISJUK, T. H. M. MES, M. KLAAS, and P. HANELT. 1997. Allotetraploid origin of *Allium altyncolicum* (Alliaceae, *Allium* sect. *Schoenoprasm*) as investigated by karyological and molecular markers. *Plant Systematics and Evolution* 206: 317-335.

FRITSCH, R. M. 1988. Anatomische Untersuchungen an der Blattspreite bei *Allium* L. (Alliaceae) I. Arten mit einer einfachen Leitbundelreihe. *Flora* 181: 83-100.

HANELT, P., J. SCHULTZE-MOTEL, R. M. FRITSCH, J. KRUSE, H. I. MAAB, H. OHLE, and K. PISTRICK. 1992. Infrageneric grouping of *Allium* - the Gatersleben approach. The genus *Allium*: taxonomic problems and genetic resources, Gatersleben, Germany: 107-123.

HUELSENBECK, J. P., and F. R. RONQUIST. 2001. MrBayes: Bayesian inference of phylogeny. *Biometrics* 17: 754-755.

JOHNSON, L. A., and D. E. SOLTIS. 1994. *matK* DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. *Systematic Botany* 19: 143-156.

KLAAS, M., and N. FRIESEN. 2002. Molecular markers in *Allium*. *In* H. D. Rabinowitch, L. Currah [ed.], *Allium* crop science: recent advances, 159-185. CABI Publishing, New York.

LI, Q. Q., S. D. ZHOU, X. J. HE, Y. YU, Y. C. ZHANG, and X. Q. WEI. 2010. Phylogeny and biogeography of *Allium* (Amaryllidaceae: Allieae) based on nuclear ribosomal internal transcribed spacer and chloroplast rps16 sequences, focusing on the inclusion of species endemic to China. *Annals of Botany* 106: 709-733.

MADDISON, W. P., and D. R. MADDISON. 1992. MacClade: analysis of phylogeny and character evolution.

MCNEAL, D. W. 1973. Bulb morphology in some western North America species of *Allium*. *Madrono* 22: 10-24.

MCNEAL, D. W. 1992. Taxonomy of North American species of *Allium*. The Genus *Allium* - Taxonomic Problems and Genetic Resources, Institute of Genetics and Crop Plant Research, Gatersleben, Germany: 195-204.

MCNEAL, D. W., and T. D. JACOBSEN. 2002. *Allium*, Flora of North America North of Mexico, vol. 26, 224-276. Oxford University Press, New York.

NGUYEN, N. H., H. E. DRISCOLL, and C. D. SPECHT. 2008. A molecular phylogeny of the wild onions (*Allium*; Alliaceae) with a focus on the western North American center of diversity. *Molecular Phylogenetics and Evolution* 47: 1157-1172.

POSADA, D., and K. A. CRANDALL. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14: 817-818.

SAGHIR, A. R. B., L. K. MANN, and M. OWNBEY. 1966. Composition of volatiles in relation to taxonomy of American *Alliums*. *American Journal of Botany* 53: 477-484.

SERNA, A. E., and A. R. LOPEZ FERRARI. 1992. Las monocotiledoneas Mexicanas: una sinopsis floristica, Parte I: Agavaceae, Alismataceae, Alliaceae, Alstromeriaceae y Amaryllidaceae.

SHAW, J., E. B. LICKEY, E. E. SCHILLING, and R. L. SMALL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *Am. J. Bot.* 94: 275-288.

SHIMODAIRA, H. 2002. An approximately unbiased test of phylogenetic tree selection. *Systematic Biology* 51: 492-508.

SHIMODAIRA, H., and M. HASEGAWA. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17: 1246-1247.

SIMMONS, M. P., and H. OCHOTERENA. 2000. Gaps as characters in sequence-based phylogenetic analysis. *Systematic Biology* 49: 369-381.

SWOFFORD, D. L. 2002. PAUP*: Phylogenetic analysis using parsimony (*and other methods), version 4.0b10.

TABERLET, P., L. GIELLY, G. PAUTOU, and J. BOUVET. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology (Historical Archive)* 17: 1105-1109.

TRAUB, H. P. 1968c. The subgenera, sections and subsections of *Allium* L. *Plant Life* 24: 147-163.

TZANOUDAKIS, D. 1992. Karyotype variation and evolution in the Greek *Allium*. The genus *Allium* - taxonomic problems and genetic resources, Gatersleben, Germany: 305-320.

WHITE, T. J., T. BRUNS, S. LEE, and J. TAYLOR. 1990. Amplication and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* M. Innis, D. Gelfant, J. Sinnsky, AND T. White [eds.], PCR protocols: A guide to methods and applications, 315-322. Academic Press, San Diego.

Table 2.1 Comparison of the classification systems of Ownbey (Saghir et al., 1966) and Traub (1968c).

OWNBEY	TRAUB	
Alliances	**Section/Type**	**Subsection**
A. acuminatum	*Lophioprason/A. sanbornii*	*Acuminata*
–		*Bolanderiana*
A. campanulatum		*Campanulata*
A. cernuum		*Cernua*
A. falcifolium		*Falcifolia*
A. sanbornii		*Sanborniana*
A. canadense	*Amerallium/A. canadense*	*Canadensia*
–		*Mexicana*
A. validum	*Caulorhizideum/A. validum*	–
A. kunthii	*Rhophetoprason/A. glandulosum*	–

Table 2.2 Taxa and collections sampled for DNA sequencing (species not sampled are also noted). Collection numbers are those of E. Wheeler unless indicated otherwise. Most determinations were verified by D.W. McNeal. Vouchers are deposited at UMO unless indicated otherwise. Taxa with an asterisk (*) were sampled from herbarium sheets, herbarium and accession number noted. Taxa with a double asterisk (**) were sampled from botanical gardens, garden and accession number noted.

Taxon	Voucher	Source
***A. acuminatum* alliance**		
A. acuminatum Hooker	07–59	USA: CA
A. amplectens Torrey	s.n.	USA: WA
A. bolanderi S. Watson	830924** (UCBG)	USA: CA
A. crispum Greene	07–36	USA: CA
A. dichlamydeum Greene	07–50	USA: CA
A. dictuon H. St. John	J Frazee s.n.	USA: WA
A. hickmanii Eastwood	07–45	USA: CA
A. hyalinum Curran	07–43	USA: CA
A. lacunosum S. Watson	S Mashayekhi 011(RSA)	USA: CA
A. peninsulare Lemmon ex Greene	07–33	USA: CA
A. praecox Brandegee	S Mashayekhi 009 (RSA)	USA: CA
A. serra McNeal & Ownbey	291683* (WS)	USA: CA
A. unifolium Kellogg	07–49	USA: CA
***A. campanulatum* alliance**		
A. bisceptrum S. Watson	08–28	USA: UT
A. membranaceum Ownbey ex Traub	07–27	USA: CA
A. campanulatum S. Watson	07–78	USA: CA
***A. canadense* alliance**		
A. canadense Linnaeus	07–04	USA: OK
A. coryi M. E. Jones	00318718* (LL)	USA: TX
A. cuthbertii Small	T Patrick s.n. (none)	USA: GA
A. drummondii Regel	07–04	USA: OK
A. geyeri S. Watson	S Mashayekhi 005 (RSA)	USA: NM
A. macropetalum Rydberg	07–12	USA: AZ
A. passeyi N. H. Holmgren & A. H. Holmgren	08–27	USA: UT
A. perdulce S. V. Fraser	07–10	USA: NM
A. plummerae S. Watson	S Mashayekhi 007 (RSA)	USA: AZ
A. runyonii Ownbey	E Keith s.n.	USA: TX
A. speculae Ownbey & Aase	T Patrick s.n. (none)	USA: GA
A. textile A. Nelson & J. F. Macbride	08–31	USA: ID

Taxon	Voucher	Source
***A cernuum* alliance**		
A. cernuum Roth	S Mashayekhi 002 (RSA)	USA: NM
A. haematochiton S. Watson	07–15	USA: CA
A. stellatum Ker Gawler	2002-1694** (MO)	USA
***A. falcifolium* alliance**		
A. aaseae Ownbey	JF Smith 2739 (SRP)	USA: ID
A. anceps Kellogg	08–19	USA: CA
A. brandegeei S. Watson	08–29	USA: UT
A. burlewii Davidson	07–23	USA: CA
A. columbianum (Ownbey & Mingrone) P. M. Peterson, Annable & Rieseberg	08–45	USA: WA
A. constrictum (Ownbey & Mingrone P. M. Peterson, Annable & Rieseberg	Not sampled	
A. cratericola Eastwood	D McNeal 4661(CPH)	USA: CA
A. crenulatum Wiegand	s.n. (none)	Canada: BC
A. douglasii Hooker	08–44	USA: WA
A. falcifolium Hooker & Arnott	07–93	USA: CA
A. fibrillum M. E. Jones ex Abrams	08–42	USA: OR
A. hoffmanii Ownbey ex Traub	07–95	USA: CA
A. lemmonii S. Watson	08–21	USA: CA
A. macrum S. Watson	D Knoke, s.n.	USA: OR
A. madidum S. Watson	07–124	USA: OR
A. nevii S. Watson	s.n.	USA: WA
A. obtusum Lemmon	07–42	USA: CA
A. parvum Kellogg	06–04 (none)	USA: CA
A. platycaule S. Watson	07–80	USA: CA
A. punctum L. F. Henderson	Not sampled	
A. robinsonii L. F. Henderson	D Knoke 1740	USA: WA
A. scilloides Douglas ex S. Watson	D Knoke 1798	USA: WA
A. siskiyouense Ownbey ex Traub	Not sampled	
A. simillimum L. F. Henderson	08–38	USA: ID
A. tolmiei Baker	07–121	USA: OR
A. tribracteatum Torrey	291693* (WS)	USA: CA
A. yosemitense Eastwood	16017* (CPH)	USA: CA
***A. kunthii* alliance**		
A. elmendorfii M. E. Jones ex Ownbey	B Carr 26649	USA: TX
A. glandulosum Link & Otto	00112207* (LL)	Mexico: Tamaulipas
A. rhizomatum Wooton & Standley	Not sampled	

Taxon	Voucher	Source
A. kunthii G. Don	S Mashayekhi 001 (RSA)	USA: TX
***A. sanbornii* alliance**		
A. abramsii (Ownbey & Aase ex Traub) McNeal	07–39	USA: CA
A. atrorubens S. Watson	08–13	USA: NV
A. bigelovii S. Watson	34620* (NAU)	USA: AZ
A. denticulatum (Ownbey & Aase ex Traub) McNeal	Not sampled	
A. diabolense (Ownbey & Aase ex Traub) McNeal	07–21	USA: CA
A. fimbriatum S. Watson	07–28	USA: CA
A. howellii Eastwood	S Mashayekhi 016 (RSA)	USA: CA
A. jepsonii (Ownbey & Aase ex Traub) S. S. Denison & McNeal	07–66	USA: CA
A. monticola Davidson	21766 (RSA))	USA: CA
A. munzii (Ownbey & Aase ex Traub) McNeal	Terri Parr.I. 21761 (RSA)	USA: CA
A. nevadense S. Watson	255919* (WS)	USA: NV
A. parishii S. Watson	S Mashayekhi 015 (RSA)	USA: CA
A. parryi S. Watson	14576* (CPH)	USA: CA
A. sanbornii Alph. Wood	Not sampled	
A. sharsmithiae (Ownbey & Aase ex Traub) McNeal	07–37	USA: CA
A. shevockii McNeal	D McNeal 3155(CPH)	USA: CA
A. tuolumnense (Ownbey & Aase ex Traub S. S. Denison & McNeal	D McNeal 4663 (CPH)	USA: CA
***A. validum* alliance**		
A. brevistylum S. Watson	08–34	USA: ID
A. eurotophyllum Wiggs	Thorne 60887 (RSA)	Mexico: Baja CA
A. gooddingii Ownbey	Not sampled	
A. validum S. Watson	07–125	USA: Oregon
Outgroups		
A. cepa Linnaeus	none	Cultivated
A. flavum Linnaeus	81.0952** (UCBG)	France
A. roseum Linnaeus	20060378** (UCBG)	USA: MO
A. tricoccum Solander	09-65	Cultivated
A. tuberosum Rottler ex Sprengel	11-02	France


A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R

Figure 2.1 Photographs of some *Amerallium* species from North America. A. *A. atrorubens* var. *atrorubens*, B. *A. unifolium*, C. *A. sanbornii* var. *sanbornii*, D. *A. jepsonii*, E. *A. falcifolium*, F. *A. madidum* bulblets, G. *A. hoffmanii*, H. A. *canadense* var. *canadense*, I. *A. dichlamydeum*, J. *A. cuthbertii*, K. *A. platycaule*, L. *A. nevadense*, M. *A. drummondii* bouquet, N. *A. speculae*, O. *A. burlewii*, P. *A. passeyi* bulbs, Q. *A. simillimum*, R. *A. canadense* var. *hyacinthoides*. Photos by Erica Wheeler except for *A. cuthbertii* and *A. speculae*, which are courtesy of Jim Allison.

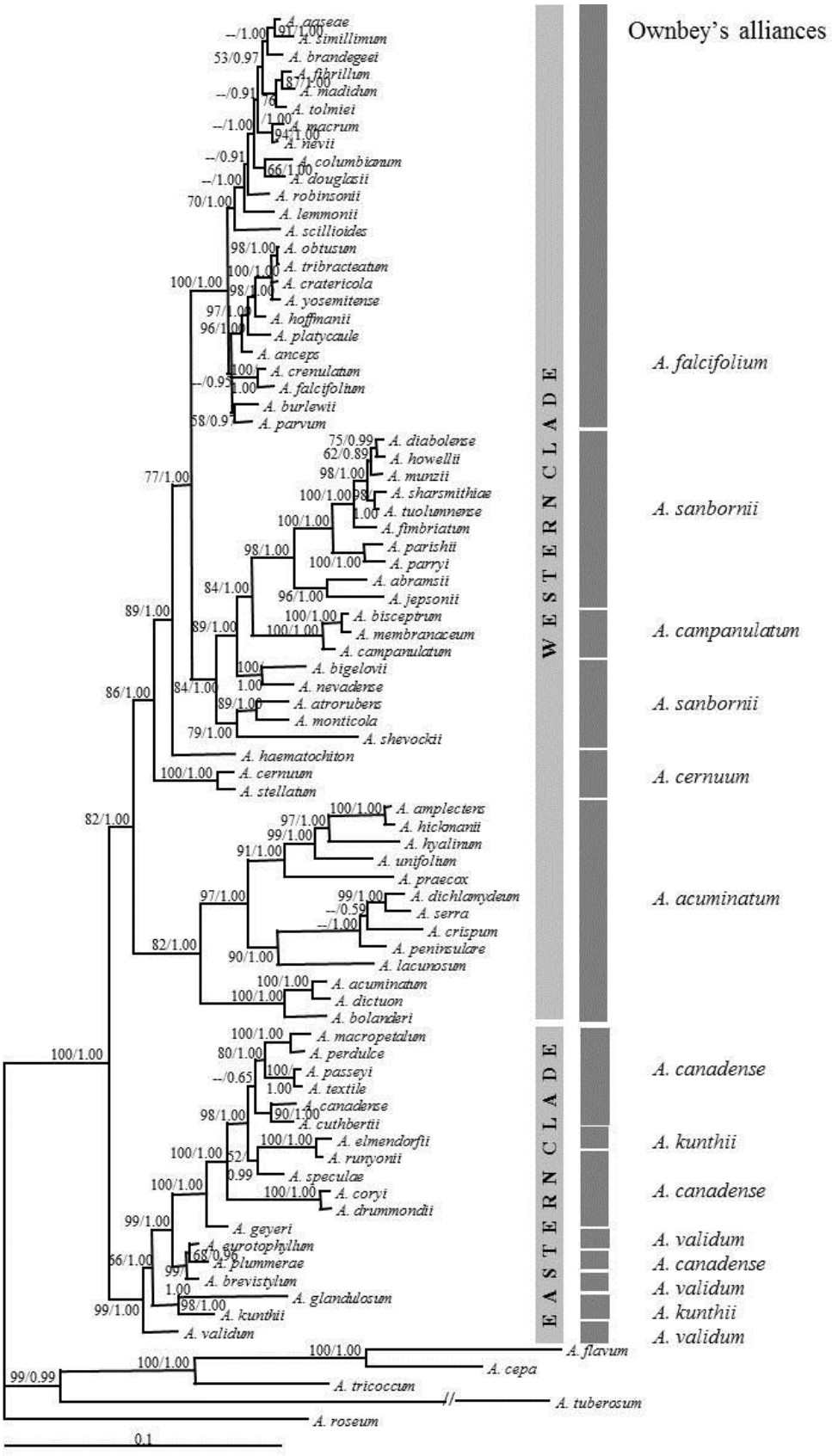
Ownbey's alliances
WESTERN CLADE
EASTERN CLADE
A. falcifolium
A. sanbornii
A. campanulatum
A. sanbornii
A. cernuum
A. acuminatum
A. canadense
A. kunthii
A. canadense
A. validum
A. canadense
A. validum
A. kunthii
A. validum
A. aqueae
A. simillimum
A. brandegeei
A. fibrillum
A. madidum
A. tolmiei
A. macrum
A. nevii
A. columbianum
A. douglasii
A. robinsonii
A. lemmonii
A. scillioides
A. obtusum
A. tribracteatum
A. cratericola
A. yosemitense
A. hoffmanii
A. platycaule
A. anceps
A. crenulatum
A. falcifolium
A. burlewii
A. parvum
A. diabolense
A. howellii
A. munzii
A. sharsmithiae
A. tuolumnense
A. fimbriatum
A. parishii
A. parryi
A. abramsii
A. jepsonii
A. bisceptrum
A. membranaceum
A. campanulatum
A. bigelovii
A. nevadense
A. atrorubens
A. monticola
A. shevockii
A. haematochiton
A. cernuum
A. stellatum
A. amplectens
A. hickmanii
A. hyalinum
A. unifolium
A. praecox
A. dichlamydeum
A. serra
A. crispum
A. peninsulare
A. lacunosum
A. acuminatum
A. dictuon
A. bolanderi
A. macropetalum
A. perdulce
A. passeyi
A. textile
A. canadense
A. cuthbertii
A. elmendorfii
A. runyonii
A. speculae
A. coryi
A. drummondii
A. geyeri
A. gyrotophyllum
A. plummerae
A. brevistylum
A. glandulosum
A. kunthii
A. validum
A. flavum
A. cepa
A. tricoccum
A. tuberosum
A. roseum
0.1

Figure 2.2. Bayesian consensus tree showing the evolutionary relationships among *Amerallium* species in North America based on the combined dataset (ITS + ETS + *trnL-F* + *rpL32-trnL*). Maximum parsimony bootstrap and Bayesian posterior probabilities are shown at nodes, respectively. Scale bar units = substitutions/site.

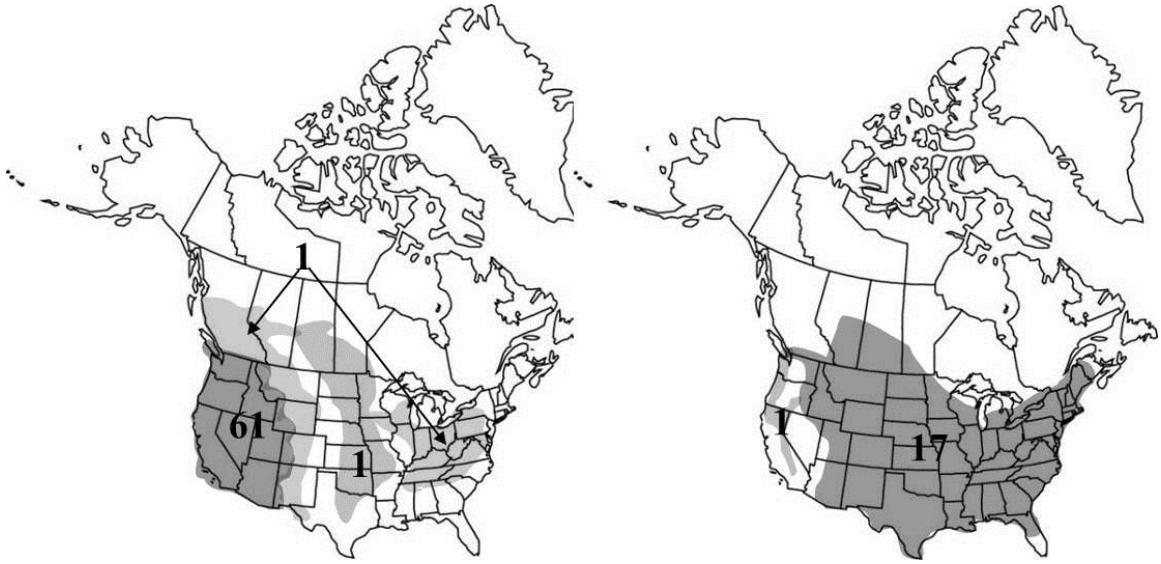

Figure 2.3. Geographic distribution of the western (left) and eastern (right) clades. Numbers of species occupying each shaded area is shown and includes all 81 *Amerallium* species recognized in the Flora of North America, north of Mexico. Base map courtesy of the Flora of North America Association.

CHAPTER 3

MULTIPLE ORIGINS OF PSEUDOVIVIPARY AND POLYPLOIDY IN THE WILD ONION, *ALLIUM CANADENSE*

ABSTRACT

Pseudovivipary is a relatively rare form of asexual reproduction in plants, characterized by the development of asexual propagules in place of sexual reproductive structures in the inflorescence. Whether this reproductive strategy arises rarely or frequently from sexual relatives has not been well studied to date. *Allium canadense* is a polyploid agamic complex that is widespread throughout much of the eastern US. One of six varieties in this complex (var. *A. c. canadense*) is pseudoviviparous. In their cytotaxonomic study of this species, Ownbey and Aase (1955) suggested that the var. *A. c. canadense* had arisen multiple times from the sexual varieties. To test this hypothesis, we conducted a molecular phylogenetic study of 119 individuals including representatives of all varieties using the external transcribed spacer region (ETS) of the nuclear ribosomal DNA repeat. Our results show that there have been at least six transitions to pseudovivipary among the samples we analyzed. We also confirm the result of Ownbey and Aase (1955) that pseudovivipary is correlated with polyploidy in this species by making somatic chromosome counts. Patterns of additive molecular variation in the ETS sequences suggest that there is historical and/or ongoing gene flow among ribotypes. The geographic distribution of ribotypes shows that the northern part of the range of this species was colonized by one ribotype

INTRODUCTION

Although sexual reproduction is overwhelmingly the dominant mode of reproduction in angiosperms, many species maintain the ability to reproduce asexually. Despite the theoretical drawbacks to clonality, such as the loss of genetic diversity due to a lack of recombination and an increased mutational load (Lynch et al., 1993), asexual taxa can be widespread and more successful than their sexually reproducing relatives in some environments (Vandel, 1928; Horandl, 2006). There is a diversity of non-exclusive explanations for this phenomenon. First, asexual plants may have a colonizing advantage over their out-crossing sexual relatives because a new population can be established with the migration of just one individual (selfing taxa are expected to have this same advantage). Second, most plants that reproduce asexually by seed (apomixis) are also polyploids (Asker and Jerling, 1992), and polyploidy itself may confer an advantage over diploidy in terms of the tolerance of ecological extremes (Levin, 1983; Otto and Whitton, 2000). Third, some asexual taxa are of hybrid origin and may benefit from a fixed state of heterosis (Richards, 2003). Fourth, asexual taxa may have evolved 'general-purpose genotypes' that enable them to successfully occupy a wide range of habitats (Baker, 1965).

Another factor that may contribute to the success of asexual plants is the apparent lability of reproductive systems in some species. In the case of apomixis, polyploid asexuals have been shown to evolve multiple times from different diploid sexual populations of the same species or species complex (Thompson and Whitton, 2006; Cosendai, Rodewald, and Horandl, 2011). Under these conditions, a diversity of asexual clones may form a patchwork of divergent genotypes across the landscape, each locally adapted to their environment. Hence, multiple origins of asexuality may lead to increased genetic diversity in taxa that might otherwise be considered genetically homogeneous.

Pseudovivipary

Pseudovivipary is defined as asexual reproduction in which plants produce asexual propagules or plantlets in the inflorescence in place of sexual reproductive structures (Elmqvist and Cox, 1996). Propagules grow on and gain resources from the parent plant before dispersing and becoming independent ramets. This habit should not be confused with true vivipary, in which sexually produced seed germinates and grows while still attached to the parent plant. In some habitats, pseudoviviparous plants may have ecological and evolutionary advantages that are specific to this reproductive strategy. In particular, resources obtained from the parental plant can outweigh those gained by seed, and the lack of dormancy enables plantlets to establish in the same season in which they are produced (Lee and Harmer, 1980; Ronsheim, 1996). This could enhance the ability of pseudoviviparous plants to colonize open habitat more quickly than seed, which may need to go through a period of stratification before germination will occur. Pseudovivipary is a relatively rare phenomenon, and has been documented in about 50 angiosperm species, including *Allium* (Elmqvist and Cox, 1996). Some well-known examples include *Polygonum viviparum* (Polygonaceae), *Poa alpinum* (Poaceae) and *Festuca vivipara* (Poaceae).

In some grasses, pseudovivipary is environmentally influenced by factors such as excess water about the roots, excess shade, high humidity, submergence, abrupt changes in moisture, day length, or temperature and insufficient vernalization (Beetle, 1980). However, the stable nature of the pseudoviviparous habit after transplantation to a common environment indicates that there is a heritable component to the expression of this reproductive strategy in some species (Flovik, 1938; Heide, 1994; Chiurugwi et al., 2010). Based on correlations between polyploidy and pseudovivipary, it has been suggested that polyploidization may play a role (Flovik, 1938; Beetle, 1980). These correlations are often not perfect and contrary examples, where diploid taxa are pseudoviviparous, can be found (Levan, 1933). An alternative (but not exclusive) hypothesis

suggests that hybridization plays a role (Flovik, 1938). Both polyploidy and hybridization can cause a disruption of the meiotic process due to improper chromosome pairing and may lead to selection for asexuality as an escape from sterility (Asker and Jerling, 1992). Since many polyploids are allopolyploids and are therefore of hybrid origin, it is difficult to disentangle the individual effects of each of these processes on the evolution of the pseudoviviparous habit. The co-occurrence of polyploidy and hybridization in apomicitic lineages similarly confounds efforts to assign causality to either of these processes in the evolution of apomixis.

Little attention has been paid to whether pseudovivipary has evolved multiple times in species and complexes that exhibit this reproductive trait and in which this trait is heritable. A recent exception to this is a molecular phylogeographic study of the pseudoviviparous *Festuca vivipara* (Poaceae) and its seminiferous (seed producing) relative *F. ovina* which demonstrates that *F. vivipara* has evolved multiple times from *F. ovina* populations (Chiurugwi et al., 2010). Considering the fact that this is the only study of its kind to date, more case studies are needed in order to better understand the causes and the evolutionary implications of pseudovivipary, especially on levels of genetic diversity maintained by asexuals.

***Allium canadense* L.**

Our study organism, *Allium canadense*, is a highly variable and widespread polyploid agamic complex that ranges throughout much of central and eastern North America (Fig. 3.1). Six varieties are recognized in the Flora of North America: *A. c. fraseri* Ownbey, *A. c. mobilense* Regel (Ownbey), *A. c. lavendulare* (J. M. Bates) Ownbey & Aase, *A. c. hyacinthoides* (Bush) Ownbey, *A. c. ecristatum* Ownbey, and *A. c. canadense* (McNeal and Jacobsen, 2002). The first five varieties are floriferous and reproduce by seed, presumably by sexual reproduction. The sixth variety, *A. c. canadense*, is pseudoviviparous and can be easily distinguished from the other five varieties by the replacement, at least in part, of flowers by asexual bulbils in the inflorescence (Fig. 3.2). Seminiferous and pseudoviviparous plants maintain their respective

phenotypes in a common garden (Ownbey and Aase, 1955), indicating that there may be genetic control of the pseudoviviparous phenotype. In the field, pseudoviviparous plants are morphologically diverse and produce a variety of forms. Some inflorescences produce only bulbils, whereas some produce bulbils and a few flowers. Still other inflorescences produce bulbils, flowers and secondary inflorescences which in turn produce bulbils, flowers and tertiary inflorescences (E. Wheeler, personal observation). While most of the flowers produced by pseudoviviparous plants are sterile, seed is occasionally set. Normal meiosis has been documented in some of these plants (Ownbey and Aase, 1955), suggesting that sexual reproduction may occasionally occur.

In their cytotaxonomic study of the six varieties of *Allium canadense*, Ownbey and Aase (1955) hypothesized that both polyploidy and hybridization have played a role in shaping the morphological and reproductive variation in this complex. Based on morphology and geographic distribution, they suggest that the diploid (2n=14) sexual vars. *A. c. fraseri* and *A. c. mobilense* hybridized at three different locations along their zone of contact (Fig. 3.1) to form the other sexual vars. *A. c. ecristatum* (2n=28), *A. c. hyacinthoides* (2n=14) and *A. c. lavendulare* (2n=14, 28) (Fig. 3.3). Each sexual variety is distinguishable morphologically at the center of its range, but in areas where the ranges of these come into contact, plants of intermediate morphology occur, suggesting that there is continued gene flow among varieties (Ownbey and Aase, 1955).

Based on morphological characters shared between local sexual and asexual forms (e.g. tepal color), Ownbey and Aase (1955) hypothesized that the pseudoviviparous var. *A. c. canadense* has evolved multiple times and is derived from the seminiferous varieties *A. c. mobilense*, *A. c. fraseri* and *A. c. lavendulare* (Fig. 3.3). Despite the possibility that the var. *canadense* might be a polyphyletic taxon, these authors argue that the asexual races are knit into 'such a tight agamic complex morphologically that all must be treated as one', taxonomically. Hence, all pseudoviviparous plants in this species are currently treated as the var. *A. c. canadense*,

regardless of their evolutionary origin. Ownbey and Aase (1955) also demonstrated a correlation between the transition to pseudovivipary and polyploidy; pseudoviviparous plants are polyploid (x=7; 2n=28, one instance of 2n=21) whereas most seminiferous plants are diploid (2n=14). An exception to this is that the seminiferous var. *A. c. lavendulare* occurs as a diploid and as a tetraploid (2n=28). Whether polyploids are the result of genome doubling within a taxon (autopolyploidy), or hybridization between divergent taxa followed by genome doubling (allopolyploidy) has not been thoroughly investigated. However, Ownbey and Aase (1955) observed that pollen meiosis occurs normally in polyploid pseudoviviparous plants, without the disturbances often associated with tetravalent formation in autopolyploids. This preliminary evidence suggests that polyploids may be allopolyploids and are therefore hybrids, potentially between closely related taxa.

The geographic distribution of pseudoviviparous plants in the *A. canadense* complex is also of interest. Although sexual and asexual forms are sympatric across much of the species range, at least half of the range of *A. canadense* is occupied only by the asexual var. *A. c. canadense*, especially in the northern and eastern parts of the range (Fig. 3.1). The geographic separation of sexual and asexual forms of closely related taxa has been termed 'geographical parthenogenesis' (Vandel, 1928), and is a well-described phenomenon in both animals (reviewed in Lynch, 1984) and in plants (reviewed in Bierzychudek, 1985). In plants, some general trends are that asexual forms have larger geographic ranges, occur at higher latitudes, at higher elevations and in previously glaciated regions (Bierzychudek, 1985). In studies where genetic variation has been examined in asexual plant populations, it is generally the case that multiple clones rather than one widespread clone have colonized the asexual part of the range (Ellstrand and Roose, 1987). In their study of *A. canadense*, Ownbey and Aase (1955) hypothesized that the northern part of the asexual part of the range is occupied by one genetic clone that is derived from the tetraploid form of the sexual var. *A. c. lavendulare*. They suggest that this clone extends from

'southern Missouri, north and eastward to the limits of the species'. At the southern end of the solely asexual part of the range they suggest that the asexual clones(s) are derived from the sexual var. *A. c. mobilense*. The asexual clones derived from the var. *A. c. fraseri*, are largely restricted to Texas and do not contribute to the var. *A. c. canadense* in the asexual part of the range.

Nuclear ribosomal DNA as molecular markers

Sequence variation in nuclear ribosomal DNA is widely used in phylogenetic studies in angiosperms, especially the internal transcribed spacer region (ITS) (Álvarez and Wendel, 2003). One advantage of rDNA is that it is biparentally inherited and intraindividual polymorphisms can sometimes be detected with direct sequencing methods (Whittall et al., 2000). When this molecular variation is additive, i.e. each of the nucleotides found in a polymorphic individual can also be found separately in other individuals, these additive polymorphic sites (APS) can be used to make inferences about hybridization among ribotypes and about the reticulate history of species (Sang, Crawford, and Stuessy, 1995; Whittall et al., 2000; Fuertes Aguilar and Nieto Feliner, 2003; Nieto Feliner, Gutierrez Larena, and Fuertes Aguilar, 2004). The external transcribed spacer region (ETS) of the nuclear ribosomal DNA repeat is even more variable than the ITS region (Baldwin and Markos, 1998), and has been shown to be phylogenetically informative among North American species of *Allium* (Nguyen, Driscoll, and Specht, 2008). However, ETS has not been applied to intraspecific studies in the genus *Allium* to date. The inclusion of putative hybrid individuals in traditional phylogenetic analyses can be problematic because divergence and reticulation have opposite effects on the topology of phylogenetic trees. However, there is evidence from morphology that hybrids are unlikely to break down cladistic structure unless they are between distantly related parents (McDade, 1992). Several studies based on nuclear rDNA have included putative hybrid individuals in traditional phylogenetic analyses (Whittall et al., 2000; Fuertes Aguilar and Nieto Feliner, 2003) and found that phylogenetic resolution is somewhat decreased with the addition of these taxa.

Aims

The aim of this research was to use molecular phylogenetic methods to test some of the hypotheses advanced by Ownbey and Aase (1955) regarding the *Allium canadense* complex, especially the origin(s) of the widespread asexual var. *canadense*. Our primary goal was to determine if the var. *A. c. canadense* is a monophyletic, or a polyphyletic taxon i.e. whether pseudovivipary has evolved once or more than once from sexual relatives. Our second goal was to add to the cytogenetic data gathered by Ownbey and Aase (1955) to establish whether polyploidy is correlated with the pseudoviviparous habit. Our third goal was to determine whether pseudoviviparous plants are of hybrid origin by assessing patterns of additive polymorphic variation in the ETS region. Our final goal was to examine the geographic distribution of genetic variation in the *A. canadense* complex, and to determine the diversity and taxonomic origins of plants in the solely asexual part of the range

MATERIALS AND METHODS

Plant material

We sampled 119 individuals of *A. canadense*, including multiple representatives of each of the six varieties for the molecular phylogeny (Table 3.1). Our study included 21 samples of herbarium specimens from the original cytotaxonomic study by Ownbey and Aase (1955), on which chromosome counts were recorded for 14 specimens. Thirty samples came from herbarium specimens of the var. *A. c. canadense* in the solely asexual part of the range. Sixty-eight samples came from field collections made during this study. Vouchers are deposited at the University of Missouri, Dunn-Palmer Herbarium (UMO). In order to sample a large number of populations and thus increase the likelihood of detecting transitions between reproductive modes, we included only one individual from each population, or two if both sexual and asexual

phenotypes occurred in the same population. We concentrated our field collections in Texas in an effort to detect the genetic boundary between the vars. *A. c. mobilense* and *A. c. A. c. fraseri*, putative parents of the putative hybrid vars. *A. c. ecristatum*, *A. c. hyacinthoides* and *A. c. lavendulare*.

Common garden

To confirm that the pseudoviviparous and floriferous habits are maintained in a common environment, we repeated Ownbey and Aase's experiment and grew bulbs collected from the wild in the common environment of a greenhouse. After collection in the wild, bulbs were allowed to dry down at room temperature for a month and then were kept in 4°C cold storage for 12-16 weeks to induce flowering. Bulbs were then planted in four-inch square pots in Sunshine Professional Growing Mix (Sun Gro Horticulture, Seba Beach Alberta) and fertilized with Miracle-Gro Water Soluble Bloom Booster Flower Food (Scotts Miracle-Gro Products Inc., Marysville Ohio) as per manufacturer's instructions. When and if plants reproduced, we noted whether they produced floriferous or pseudoviviparous inflorescences and compared this to the kind of inflorescence they had produced in the wild.

DNA extraction, PCR and sequencing

DNA was extracted from dried leaves using a modified version of the CTAB protocol (Doyle and Doyle, 1987). The external transcribed spacer region (ETS) of the ribosomal RNA repeat was amplified using the forward primer *Allium* ETS (Nguyen, Driscoll, and Specht, 2008) and the reverse primer 18SETS (Baldwin and Markos, 1998) under the following conditions: initial DNA denaturation at 95°C for five minutes followed by 40 amplification cycles of 95°C for one minute, 50°C for one minute, 72°C for two minutes and a final extension at 72°C for ten minutes. PCR products were stained with ethidium bromide and examined for quality and concentration on a 1% agarose gel run in TBE buffer and then purified using QiaQuick PCR

purification columns (Qiagen, Valencia California). PCR products were directly sequenced using Applied Biosystems Inc (Foster City California) BigDye dideoxy terminator sequencing kits v. 3.1. Amplification primers described above were also used in sequencing reactions. Sequences were analyzed on an ABI 3730 automated sequencer at the University of Missouri DNA Core Facility. Forward and reverse sequences from each sample were assembled in the program SeqMan (DNAStar Inc. Madison, Wisconsin) and assessed for quality before exportation for alignment. Sequences were aligned in Clustal X v. 1.83 (Thompson et al., 1997) and exported for phylogenetic analysis.

Phylogenetic analysis

A heuristic search for trees was performed with maximum parsimony as the optimality criterion in the program PAUP v.4.0b10 (Swofford, 2002) using default settings of the program. A maximum of 1000 most parsimonious trees was saved and a strict consensus tree was calculated. To assess support for clades, maximum parsimony bootstrap analysis was performed in PAUP with 500 replicate searches using the TBR option. During each replicate, a maximum of 100 trees was saved. Ambiguous nucleotide characters caused by double-peaks in the sequence traces were left in the data matrix as IUPAC nucleotide ambiguity codes which PAUP treats as missing data when DATATYPE=DNA is specified. We interpret these polymorphisms as evidence that these individuals may be of hybrid origin. Four closely related species of *Allium* from the *A. canadense* alliance (Saghir, Mann, and Ownbey, 1966) were used as outgroups in the phylogenetic analysis: *A. geyeri* S. Watson, *A. passeyi*, N.H. Holmgren & A.H. Holmgren, *A. perdulce* S.V. Fraser and *A. macropetalum* Rydberg. ETS sequences for individual samples are deposited in GenBank and the alignment can be found in TreeBASE.

Determination of ploidy level

Somatic chromosomes were counted following a modified version of the method described by Allen (1984). Mitotically active root tips were excised from plants, placed in 0.03% colchicine solution and stored at 4°C overnight to arrest the cell cycle and accumulate metaphase cells. Root tips were rinsed with distilled water, blotted dry and fixed in 3:1 ethanol:glacial acetic acid for a minimum of 48 hours. Root tips were then hydrolyzed in 1M HCl at 60°C for 15 minutes, rinsed in distilled water and stained in leuco-basic fuchsin for 1-2 hours. Highly stained root meristem tissue was dissected from the root tips and squashed by hand on a microscope slide in 45% acetic acid. Chromosomes were visualized at 100X magnification with an Olympus BX61 microscope (Olympus America Inc., Center Valley Pennsylvania).

RESULTS

Common garden

Without exception, plants that produced a pseudoviviparous inflorescence in the wild produced a pseudoviviparous inflorescence in a common garden, and plants that produced a floriferous inflorescence in the wild produced a floriferous inflorescence in a common garden. The production of secondary and tertiary inflorescences seen in pseudoviviparous plants in the wild was not seen in the controlled environment of the greenhouse, suggesting that this phenotypic variation is the result of environmental variation during the development of the inflorescence in the wild.

Phylogeny

Of the 551 base pair alignment, 458 characters were constant, 31 were variable but uninformative and 62 were parsimony informative (A=0.20, C=0.20, G=0.26, T=0.34). The length of the most parsimonious trees was 130 and the consistency index was 0.8. The strict

consensus tree shows that the six varieties of *A. canadense* form a well-supported clade, confirming that this complex is monophyletic (97 % MP bootstrap support) (Fig. 3.4). Two main clades are supported corresponding to the two putative parental varieties in the complex, *A. c. mobilense* and *A. c. fraseri.* Clade I contains individuals of the vars. *A. c. mobilense*, *A. c. lavendulare* and *A. c. canadense* (68%). Clade II contains individuals of the vars. *A. c. fraseri*, *A. c. hyacinthoides*, *A. c. ecristatum* and *A. c. canadense* (94%). None of the varieties except *A. c. ecristatum* is monophyletic.

The pseudoviviparous variety *A. c. canadense* occurs in all main clades in the ETS strict consensus tree, and is phylogenetically related to different seminiferous varieties. This result indicates that the var. *A. c. canadense* is not a monophyletic taxon. By examining the phylogenetic relationships among seminiferous and pseudoviviparous samples, we conservatively infer that pseudovivipary has evolved six times in the *A. canadense* complex based on our sampling (Fig. 3.4). Because there is a lack of fine resolution in the phylogenetic tree, in some cases it is not possible to determine which particular seminiferous variety is most closely related to samples of the var. *A. c. canadense*. For example, in Clade IA2, it is not possible to say whether the samples of the var. *A. c. canadense* are most closely related to samples of the var. *A. c. mobilense* or the var. *A. c. lavendulare*. However, it is clear that there have been four transitions to asexual reproduction from sexual reproduction in the *A. c. mobilense* clade (Clade I) and two transitions in the *A. c. fraseri* clade (Clade II).

Polyploidy

Chromosome counts were made of 13 individuals during this study. All seven plants of the var. *A. c. canadense* that were examined were tetraploid (2n=28) (Fig. 3.5). The one plant of the var. *A. c. fraseri* that was examined was diploid (2n=14). The two plants of the var. *A. c. mobilense* that were examined were diploid. Two plants of the var. *A. c. lavendulare* were examined; one was diploid + one supernumerary chromosome, and the other was tetraploid. The

one plant of the var. *A. c. hyacinthoides* that was examined was diploid. No live plant material of the var. *A. c. ecristatum* was collected for cytogenetic analysis. Based on the 13 chromosome counts made in this study and the 14 made by Ownbey and Aase (1955), we infer that polyploidy has arisen at least four times from diploid relatives in this species (Fig. 3.4).

Additive polymorphic sites

At 55 nucleotide positions in the alignment (10%), double-peaks appeared in the electropherograms and resulted in IUPAC ambiguous nucleotide characters being called in some samples. In 32 cases this molecular variation was additive i.e. when there was an ambiguous character at a particular site in a sample, both of the nucleotides that define that ambiguous character were present in other samples at that position in the alignment. For example, in the case of a double peak resulting in an 'R (A/G)' in one sample, some samples in the alignment had an 'A' at that position and other samples had a 'G'. We interpret these additive polymorphic sites (APS) as allelic polymorphisms in heterozygote individuals and as evidence of past and/or ongoing gene flow among ribotypes. At the other sites that showed double peaks, the variation was not completely additive i.e. only one of the nucleotides that define the ambiguous base was present in other samples at that site. There are a few ways to explain this non-additive variation. First, we may not have sampled individuals with the missing ribotype. Second, concerted evolution may have acted to homogenize the two copies in one direction or the other in most individuals. Third, the double-peaks at these positions could be the result of background signal in the electropherograms. To avoid the problem of mis-interpreting background signal as something of biological significance, we only consider the 32 APSs as evidence of gene flow among ribotypes.

APSs occurred in samples of the vars. *A. c. mobilense* (3/20), *A. c. fraseri* (4/15), *A. c. lavendulare* (1/7), *A. c. ecristatum* (1/2) and *A. c. canadense* (29/66). No APSs were detected in the var. *A. c. hyacinthoides* (0/6) or in the unassigned samples *A. c. mobilense/fraseri* (0/3). None

of the nucleotide sites that define the separation of the *mobilense* and the *fraseri* clades exhibited APSs, suggesting that there is little gene flow between these two main clades. Instead, APSs that were detected were the result of gene flow within each of the main clades, among more closely related subclades.

Geographic distribution of clades and colonization of the asexual part of the range

There is structure to the geographic distribution of genetic variation in the samples included in the ETS phylogeny (Fig. 3.6). Clade I includes samples of the vars. *A. c. mobilense*, *A. c. lavendulare* and *A. c. canadense* and extends north and east from south central Texas to the limits of the species range (colored circles in Fig. 5). Clade II includes the vars. *A. c. fraseri*, *A. c. hyacinthoides*, *A. c. ecristatum* and *A. c. canadense* and extends north from south central Texas into Kansas (black circles in Fig. 3.6). The geographic boundary between the *A. c. mobilense* clade to the east (Clade I) and the *A. c. fraseri* clade to the west (Clade II) is especially evident in Texas, where our field sampling was most concentrated. The location of this boundary coincides with the boundary between the vars. *A. c. fraseri* and *A. c. mobilense* in the south, and the boundary between the vars. *A. c. fraseri* and *A. c. lavendulare* in the north, as delimited by Ownbey and Aase (1955) (Fig. 3.1). However, there is some geographic overlap of these clades around their zone of contact.

Based on the position of clades in the phylogeny and on their geographic distributions, there is evidence of range expansions of clades out of Texas, both northward and eastward into the asexual part of the range. This is most evident in the *A. c. mobilense* clade. Clade IC (orange circles, Fig. 3.6) is the earliest diverging lineage in the *A. c. mobilense* clade and is also the most southerly. It extends from eastern Texas to southern Mississippi and into South Carolina. This clade is associated with the sexual var. *A. c. mobilense*, and is a source of the var. *A. c. canadense* in the southern part of the asexual part of the range. In north Texas and Oklahoma, Clade IC gives way to the sister clades IB and IA. Clade IB (blue circles, Fig. 5) is associated with the

sexual var. *A. c. lavendulare*. It extends north into Nebraska, but does not extend eastward into the asexual part of the range. Clade IA (yellow, red and green circles, Fig. 3.6), is associated with the sexual vars. *A. c. mobilense* and *A. c. lavendulare*, and is the most widespread clade. It extends from Oklahoma and Nebraska in the west to Georgia and South Carolina in the east and New Brunswick in the north. Clade IA1 (green circles, Fig. 5) is the source of the var. *A. c. canadense* throughout most of the asexual part of the range, especially in the north. All individuals in this clade have the same ETS ribotype and based on this marker, can be seen as one widespread clone.

DISCUSSION

Despite having been documented in over 50 angiosperm species (Elmqvist and Cox, 1996), little is known about the evolvability of the asexual pseudoviviparous habit. In particular, whether this reproductive strategy arises rarely or often within species in which it appears to be heritable, has largely gone untested (but see Chiurugwi et al., 2010). Using DNA sequence variation in the ETS region of nuclear rDNA, we investigated genealogical relationships among seminiferous and pseudoviviparous populations of the *A. canadense* polyploid agamic complex in North America to gain insight into the evolutionary history of the pseudoviviparous var. *A. c. canadense*. In addition, we combined previously obtained chromosome counts (Ownbey and Aase, 1955) with those newly obtained in this study to determine if polyploidy is correlated with the pseudoviviparous habit.

Multiple origins

The ETS phylogeny shows that the pseudoviviparous var. *A. c. canadense* is not a monophyletic taxon, but has evolved at least six times among the populations we sampled. This result supports the hypothesis of Ownbey and Aase (1955) that var. *A. c. canadense* has arisen *de novo* from different populations of the sexual varieties. Our results are also consistent with other

studies of clonal plants and animals which show that asexual taxa are often polyphyletic in origin (reviewed in Vrijenhoek and Parker, 2009). The various evolutionary and taxonomic origins of the var. *A. c. canadense* helps to explain the breadth of morphological variation found in this taxon; asexual clones resemble their sexual progenitor(s) in characters such as tepal color, leaf width etc (Ownbey and Aase, 1955). Given the number of independent origins of pseudovivipary detected among our limited set of specimens, it is likely that further sampling in regions we were unable to include in our study would uncover evidence of additional transitions between sexual and asexual reproduction in *A. canadense*.

The polyphyly of var. *A. c. canadense* does not negate its standing as an infraspecific taxon of the species *A. canadense*. Indeed, there is precedence for the maintenance of polyphyletic plant taxa even at the specific rank. In their molecular phylogeny of populations of the pseudoviviparous *Festuca vivipara* (Poaceae) and the seminiferous *F. ovina* in Great Britain, Chiurugwi et al.(2010) found that *F. vivipara* was polyphyletic and had arisen multiple times from local populations of *F. ovina*. Despite this, these authors argue that *F. vivipara* is a morphologically and ecologically, if not genetically distinct species. In a well-known example that does not involve pseudovivipary, the allopolyploid species *Tragopogon mirus* and *T. miscellus* have arisen multiple times by hybridization between the diploid species *T. dubius*, *T. porrifolius*, and *T. pratensis* (Symonds, Soltis, and Soltis, 2010). The same can be argued in the case of the var. *A. c. canadense* for the maintenance of its infraspecific status. Despite a lack of genetic coherence, the var. *A. c. canadense* is held together by striking morphological and reproductive characters. It would not be practical (or even possible) to segregate the pseudoviviparous races of each variety taxonomically. However, it does make sense to include information about the multiple origins of the var. *A. c. canadense* in floristic treatments to avoid an oversimplification of evolutionary relationships in this complex.

Polyploidy

All pseudoviviparous plants examined to date are also polyploid, indicating a strong correlation between polyploidization and this reproductive strategy in *A. canadense*. However, the converse is not true; all polyploids are not pseudoviviparous as demonstrated by the tetraploid form of the seminiferous var. *A. c. lavendulare*. This pattern suggests that polyploidy is necessary, but may not always be sufficient to induce the transition to pseudovivipary in *A. canadense*. Among other *Allium* species that exhibit pseudovivipary, the correlation with polyploidy varies. In the North American species *A. geyeri*, Ownbey and Aase (1955) also demonstrated that seminiferous plants (var. *geyeri*) are largely diploid whereas pseudoviviparous plants (var. *tenerum*) are always polyploid. Levan (1933) found that pseudovivipary was correlated with polyploidy in the European species *A. carinatum* and *A. oleraceum* but also showed that the pseudoviviparous species *A. scorodoprasum* and *A. cepa* var. *proliferum* were diploid. In addition, although species in the genus *Allium* clearly have a capacity for pseudovivipary, polyploidization does not lead to this phenotype in many species. Of the 84 *Allium* species native to North America, at least 12 species maintain both diploids and polyploids in their populations (McNeal and Jacobsen, 2002). However, only in *A. canadense* and *A. geyeri* are polyploids pseudoviviparous; in all other species in North America, polyploids remain seminiferous.

Hybridization

It has also been suggested that hybridization is correlated with pseudovivipary, especially in some arctic and alpine grasses (Flovik, 1938). The presence of additive polymorphic sites (APS) in the ETS region of nuclear rDNA indicates that there has been a history of gene flow among different ribotypes in the *A. canadense* complex. We detected APSs in 29/66 samples (44%) of the var. *A. c. canadense* included in our study, suggesting that at least some populations

of this taxon are of hybrid origin. However, we also detected APSs in 9/53 samples (17%) of the sexual varieties, indicating that some sexual populations may also be of hybrid origin.

Pseudoviviparous plants are capable of normal meiosis and sometimes produce seed from apparently normal flowers (Ownbey and Aase, 1955), so it is possible that sexual reproduction occasionally occurs in the var. *A. c. canadense*. Whereas polyploid pseudoviviparous plants are likely to be reproductively isolated from their diploid sexual relatives as a result of ploidy differences, sexual reproduction between polyploid pseudoviviparous plants is more likely. If genetically divergent clones of the var. *A. c. canadense* come into contact with each other as a result of range expansions from their local origins, occasional sexual reproduction may also result in pseudoviviparous plants that are of hybrid origin. What is not known is whether seed produced by pseudoviviparous plants is viable, and more importantly, whether the seed develops into the pseudoviviparous phenotype. This aspect of the reproductive biology of *A. canadense* requires further investigation.

Geographical parthenogenesis

Despite its asexual habit, the var. *canadense* is not genetically homogeneous. Based on the ETS phylogeny, multiple clones representing independent evolutionary and taxonomic origins blanket the species range, especially in areas where sexual and asexual forms are sympatric. Competition for resources among clones and between sexual and asexual forms is expected to result in the evolution of either generalist or specialist asexual genotypes, according to two competing theories that have been advanced to explain the phenomenon of geographical parthenogenesis (Vrijenhoek and Parker, 2009). The general purpose genotype (GPG) hypothesis (Baker, 1965; Parker et al., 1977) argues that among the asexual clones produced by a sexual population, there will be variation for niche breadth i.e. some clones will be able to withstand a wider range of environmental conditions than others. Over several generations, natural selection will fix those clonal genotypes with wider than average tolerances (Vrijenhoek and Parker, 2009),

leading to GPGs. Proponents of this model suggest that the colonizing success of asexuals in cases of geographical parthenogenesis is the result of the range expansion of GPGs that can be successful in a variety of environments. The frozen niche variation (FNV) hypothesis (Vrijenhoek, 1979; Vrijenhoek, 1984) argues that interclonal competition and competition between asexuals and sexuals leads to the fixation of asexual genotypes that can specialize on resources that are underutilized by their competitors. Proponents of this model suggest that the colonizing success of asexuals in cases of geographical parthenogenesis is the result of the range expansion of specialist genotypes into environments to which they are particularly well adapted.

One line of evidence in support of the GPG hypothesis is the occurrence of very widespread clones across a range of habitats (Vrijenhoek and Parker, 2009). In the case of *A. canadense*, all individuals in Clade IA1 have the same ribotype and can be seen as one widespread clone. Based on our sampling, this clone occupies much of the asexual part of the range north of South Carolina to the northeastern limits of the species range in New Brunswick. The broad distribution of this clone supports the GPG hypothesis in the northern part of the range. It also supports the hypothesis of Ownbey and Aase (1955) that the northern part of the asexual part of the range was colonized by a single clone that was derived from the sexual var. *A. c. lavendulare*.

Putative hybrid varieties

Based on morphology and geography, Ownbey and Aase (1955) hypothesized that the sexual vars. *A. c. mobilense* and *A. c. fraseri* hybridized at three different locations to give rise to the other sexual vars. *A. c. ecristatum*, *A. c. hyacinthoides* and *A. c. lavendulare*. If hybridization had occurred between the *A. c. mobilense* and *A. c. fraseri* clades to form these other taxa, we might expect to see additive molecular variation in the putative hybrids at the nucleotide sites that define each of these two clades (Sang, Crawford, and Stuessy, 1995; Whittall et al., 2000; Fuertes Aguilar and Nieto Feliner, 2003). However, no heterozygotes were detected at these sites with

direct sequencing in the putative hybrid taxa, or in any other samples. If hybridization did in fact occur, there are a few ways to explain the fact that we did not find additive molecular variation using direct sequencing. First, extensive backcrossing to one parent or the other may have resulted in the homogenization of the rDNA repeat to represent just one parent. Second, concerted evolution may have acted to homogenize the rDNA repeat, a process that can occur very rapidly (Kovarik et al., 2005). Third, the PCR may have randomly amplified only one copy of the rDNA repeat in an individual. Alternatively, this evidence suggests that the vars. *A. c. ecristatum, A. c. hyacinthoides* and *A. c. lavendulare* are not of hybrid origin but instead are segregates of the vars. *A. c. mobilense* and *A. c. fraseri*. In particular the var. *A. c. lavendulare* may be a segregate of the var. *A. c. mobilense* and the vars. *A. c. hyacinthoides* and *A. c. ecristatum* may be segregates of the var. *A. c. fraseri*. Either way, our data do not lend support to the hypothesis that the vars. *A. c. ecristatum, A. c. hyacinthoides* and *A. c. lavendulare* formed as a result of hybridization between the vars. *A. c. mobilense* and *A. c. fraseri*.

Future directions

In garlic (*A. sativum*), the differential expression of two alternatively spliced *LEAFY* homolog variants (*gaLFY*) is implicated in the developmental differences between seminiferous and pseudoviviparous plants (Rotem et al., 2007; Rotem et al., 2011). In particular, a long variant of *gaLFY* is expressed in seminiferous plants during reproductive development, but is not expressed in pseudoviviparous plants. A short variant of *gaLFY* is expressed in both phenotypes throughout development. Unlike the situation in *A. canadense*, both sexual and asexual forms of *A. sativum* are diploid (2n=16), so polyploidy is not considered to be a factor in the transition to pseudovivipary. In *Arabidopsis*, the *LEAFY* gene encodes a transcription factor that controls the transition from the inflorescence meristem to a floral meristem and regulates the expression of floral identity genes (Weigel et al., 1992). *LEAFY* mutants produce leafy shoots rather than flowers in the inflorescence. Could disruption of the expression of the homolog of *gaLFY* in *A.*

canadense hybrid polyploids be linked to the pseudoviviparous phenotype in this species? Considering the evidence from *A. sativum*, *gaLFY* would be an excellent candidate gene for experiments aimed at finding the genetic underpinnings of this reproductive strategy in this species.

ACKNOWLEDGEMENTS

I would like to thank curators from the following herbaria for permission to remove leaf material from their collections of *A. canadense*: CAN, F, ISC, KY, MICH, MIN, MISS, NEB, NY, PH, TENN, USCH, VT, WS, and WVU. Thanks to April Diebold and Victoria Docktor for their help in the lab, Carol Clark for her help in the field, and to Melody Kroll for her help with preparing figures. This research was supported by a University of Missouri Life Science Fellowship and grants from the National Science Foundation (DEB 0829849), the American Society of Plant Taxonomists, and the Missouri Native Plant Society.

LITERATURE CITED

ALLEN, G. A. 1984. Morphological and cytological variation in the western North American *Aster occidentalis* complex (Asteraceae). *Systematic Botany* 9: 175-191.

ÁLVAREZ, I., and J. F. WENDEL. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* 29: 417-434.

ASKER, S. E., and L. JERLING. 1992. Apomixis in plants. CRC-Press, Boca Raton, FL.

BAKER, H. G. 1965. Characteristics and modes of origins of weeds. *In* H. G. Baker AND G. L. Stebbins [eds.], Genetics of colonizing species, 147-172. Academic Press, New York.

BALDWIN, B. G., and S. MARKOS. 1998. Phylogenetic utility of the external transcribed spacer (ETS) of 18S-26S rDNA: congruence of ETS and ITS trees of *Calycadenia* (Compositae). *Molecular Phylogenetics and Evolution* 10: 449-463.

BEETLE, A. A. 1980. Vivipary, proliferation and phyllody in grasses. *Journal of Range Management* 33: 256-261.

BIERZYCHUDEK, P. 1985. Patterns in plant parthenogenesis. *Cellular and Molecular Life Sciences* 41: 1255-1264.

CHIURUGWI, T., M. A. BEAUMONT, M. J. WILKINSON, and N. H. BATTEY. 2010. Adaptive divergence and speciation among sexual and pseudoviviparous populations of *Festuca*. *Heredity* 106: 854-861

COSENDAI, A.-C., J. RODEWALD, and E. HORANDL. 2011. Origin and distribution of autopolyploids via apomixis in the alpine species *Ranunculus kuepferi* (Ranunculaceae). *Taxon* 60: 355-364.

DOYLE, J. J., and J. D. DOYLE. 1987. A rapid DNA isolation procedure from small quantities of fresh leaf tissues. *Phytochemical Bulletin* 19: 11-15.

ELLSTRAND, N. C., and M. L. ROOSE. 1987. Patterns of genotypic diversity in clonal plant species. *American Journal of Botany* 74: 123-131.

ELMQVIST, T., and P. A. COX. 1996. The evolution of vivipary in flowering plants. *Oikos* 77: 3-9.

FLOVIK, K. 1938. Cytological studies of arctic grasses. *Hereditas* 24: 265-376.

FUERTES AGUILAR, J., and G. NIETO FELINER. 2003. Additive polymorphisms and reticulation in an ITS phylogeny of thrifts (*Armeria*, Plumbaginaceae). *Molecular Phylogenetics and Evolution* 28: 430-447.

HEIDE, O. M. 1994. Control of flowering and reproduction in temperate grasses. *New Phytologist* 128: 347-362.

HORANDL, E. 2006. The complex causality of geographical parthenogenesis. *New Phytologist* 171: 525-538.

KOVARIK, A., J. C. PIRES, A. R. LEITCH, K. Y. LIM, A. M. SHERWOOD, R. MATYASEK, J. ROCCA, et al. 2005. Rapid concerted evolution of nuclear ribosomal DNA in two *Tragopogon* allopolyploids of recent and recurrent origin. *Genetics* 169: 931-944.

LEE, J. A., and R. HARMER. 1980. Vivipary, a reproductive strategy in response to environmental stress? *Oikos* 35: 254-265.

LEVAN, A. 1933. Cytological studies in *Allium* III. *Allium carinatum* and *Allium oleraceum*. *Hereditas* 18: 101-114.

LEVIN, D. A. 1983. Polyploidy and novelty in plants. *The American Naturalist* 122: 1-25.

LYNCH, M. 1984. Destabilizing hybridization, general-purpose genotypes and geographic parthenogenesis. *The Quarterly Review of Biology* 59: 257-290.

LYNCH, M., R. BURGER, D. BUTCHER, and W. GABRIEL. 1993. The mutational meltdown in asexual populations. *Journal of Heredity* 84: 339-344.

MCDADE, L. A. 1992. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. *Evolution* 46: 1329-1346.

MCNEAL, D. W., and T. D. JACOBSEN. 2002. *Allium*, Flora of North America North of Mexico, vol. 26, 224-276. Oxford University Press, New York.

NGUYEN, N. H., H. E. DRISCOLL, and C. D. SPECHT. 2008. A molecular phylogeny of the wild onions (*Allium*; Alliaceae) with a focus on the western North American center of diversity. *Molecular Phylogenetics and Evolution* 47: 1157-1172.

NIETO FELINER, G., B. GUTIERREZ LARENA, and J. FUERTES AGUILAR. 2004. Fine-scale geographical structure, intra-individual polymorphism and recombination in nuclear ribosomal internal transcribed spacers in *Armeria* (Plumbaginaceae). *Ann Bot* 93: 189-200.

OTTO, S. P., and J. WHITTON. 2000. Polyploid incidence and evolution. *Annual Review of Genetics* 34: 401-437.

OWNBEY, M., and H. C. AASE. 1955. Cytotaxonomic studies in *Allium* L.: the *Allium canadense* alliance, Research Studies of the State College of Washington, 1-106. State College of Washington.

PARKER, E. D., JR., R. K. SELANDER, R. O. HUDSON, and L. J. LESTER. 1977. Genetic diversity in colonizing parthenogenetic cockroaches. *Evolution* 31: 836-842.

RICHARDS, A. J. 2003. Apomixis in flowering plants: an overview. *Philosophical Transactions: Biological Sciences* 358: 1085-1093.

RONSHEIM, M. L. 1996. Evidence against a frequency-dependent advantage for sexual reproduction in *Allium vineale*. *The American Naturalist* 147: 718-734.

ROTEM, N., R. DAVID-SCHWARTZ, Y. PERETZ, I. SELA, H. RABINOWITCH, M. FLAISHMAN, and R. KAMENETSKY. 2011. Flower development in garlic: the ups and downs of *gaLFY* expression. *Planta* 233: 1063-1072.

ROTEM, N., E. SHEMESH, Y. PERETZ, F. AKAD, O. EDELBAUM, H. D. RABINOWITCH, I. SELA, et al. 2007. Reproductive development and phenotypic differences in garlic are associated with expression and splicing of LEAFY homologue gaLFY. *J. Exp. Bot.* 58: 1133-1141.

SAGHIR, A. R. B., L. K. MANN, and M. OWNBEY. 1966. Composition of volatiles in relation to taxonomy of American *Alliums*. *American Journal of Botany* 53: 477-484.

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 Sciences of the United States of America* 92: 6813-6817.

SWOFFORD, D. L. 2002. PAUP*: Phylogenetic analysis using parsimony (*and other methods), version 4.0b10.

SYMONDS, V. V., P. S. SOLTIS, and D. E. SOLTIS. 2010. Dynamics of polyploid formation in *Tragopogon* (Asteraceae): recurrent formation, gene flow, and population structure. *Evolution; international journal of organic evolution* 64: 1984-2003.

THOMPSON, J. D., T. J. GIBSON, F. PLEWNIAK, F. JEANMOUGIN, and D. G. HIGGINS. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 24: 4876-4882.

THOMPSON, S. L., and J. WHITTON. 2006. Patterns of recurrent evolution and geographic parthenogenesis within apomictic polyploid Easter daises (*Townsendia hookeri*). *Molecular Ecology* 15: 3389-3400.

VANDEL, A. 1928. La parthenogenese geo- graphique. Contribution 'a l'etude biologique et cytologique de la parthenogenese naturelle. *Bull. Biol. France Belg* 62: 164-281.

VRIJENHOEK, R. 1984. Ecological differentiation among clones: the frozen niche variation model. *In* K. Wöhrmann AND V. Loeschcke [eds.], Population biology and evolution. Springer-Verlag, Heidelberg.

VRIJENHOEK, R. C. 1979. Factors affecting clonal diversity and coexistence. *American Zoologist* 19: 787-797.

VRIJENHOEK, R. C., and E. D. J. PARKER. 2009. Geographical parthenogenesis: general purpose genotypes and frozen niche variation. *In* I. Schön, K. Martens, AND P. v. Dijk [eds.], Lost sex: the evolutionary biology of parthenogenesis, vol. 1, 99-131. Springer, Heidelberg.

WEIGEL, D., J. ALVAREZ, D. R. SMYTH, M. F. YANOFSKY, and E. M. MEYEROWITZ. 1992. *LEAFY* controls floral meristem identity in Arabidopsis. *Cell* 69: 843-859.

WHITTALL, J., A. LISTON, S. GISLER, and R. J. MEINKE. 2000. Detecting nucleotide additivity from direct sequences is a SNAP: an example from *Sidalcea* (Malvaceae). *Plant Biology* 2: 211-217.

Table 3.1. Samples of *A. canadense* used in the molecular phylogenetic study. Sample numbers are indicated on the strict consensus tree in Fig. 4. Herbarium collections are preceded by the Index Herbariorum code; those preceded by WS were included in Ownbey and Aase's study (1955). Field collections made by the first author are preceded by EW. In three cases, the morphology of plants collected in the field were intermediate to the vars. *A. c. mobilense* and *A. c. fraseri*, so could not be assigned to one taxon or the other. These samples are assigned '*mobilese/fraseri*' indicating that their taxonomic status is unknown.

#	Variety	Collection	State/Province	County
1	*canadense*	EW-2011-001	Missouri	Boone
2	*canadense*	CAN 472311	Quebec	Maddington
3	*canadense*	CAN 484173	Ontario	Huron
4	*canadense*	CAN 489922	New Brunsw.	York
5	*canadense*	ISC 420620	Iowa	Story
6	*canadense*	ISC 438857	Iowa	Hamilton
7	*canadense*	MICH M.J. Oldham 21544	Ontario	Peterborough
8	*canadense*	MICH A.A. Reznicek 8523	Ontario	Huron
9	*canadense*	MICH E.G. Voss 16643	Michigan	Mackinac
10	*canadense*	MIN 919332	Minnesota	Pine
11	*canadense*	MIN 781480	Minnesota	Jackson
12	*canadense*	MIN 508432	Minnesota	Wabasha
13	*canadense*	NEB 291513	Nebraska	Johnson
14	*canadense*	Nils Holmgren 3172	Missouri	
15	*canadense*	TENN 1049 V.E. McNeilus 92-217	Tennessee	Jackson
16	*canadense*	TENN 1049 Dwayne Estes 535	Tennessee	Maury
17	*canadense*	TENN 1049 Dwayne Estes s.n.	Tennessee	Giles
18	*canadense*	TENN 1049 Dwayne Estes 4673	Tennessee	Maury
19	*canadense*	USCH 63664	South Carolina	York
20	*canadense*	VT Peter F. Zika 8004	Vermont	Chittenden
21	*canadense*	VT Peter F. Zika 8985	Vermont	Rutland
22	*canadense*	VT Pamela Weatherbee 4082	Vermont	Bennington
23	*canadense*	WVU 163309	West Virginia	Summers
24	*lavendulare*	WS 281411	Missouri	Montgomery
25	*mobilense*	WS 259552	Arkansas	Independence
26	*canadense*	WS 281594	Missouri	Jefferson
27	*canadense*	WS 6483	Oklahoma	Rogers
28	*lavendulare*	WS 281416	Missouri	Johnson
29	*mobilense*	WS 281484	Missouri	Taney
30	*canadense*	EW-2009-047	Missouri	Dade
31	*lavendulare*	EW-2009-06	Missouri	Dade
32	*mobilense*	EW-2009-042	Oklahoma	Pushmataha
33	*canadense*	EW-2009-044	Oklahoma	Cherokee

#	Variety	Collection	State/Province	County
34	*canadense*	EW-2009-048	Missouri	Jasper
35	*canadense*	EW-2009-049	Missouri	Jasper
36	*canadense*	WS281587	Georgia	Oglethorpe
37	*canadense*	EW-2009-050	Missouri	Newton
38	*canadense*	NEB 307084	Nebraska	Stanton
39	*canadense*	USCH 61382	South Carolina	Cherokee
40	*canadense*	MISS 72149	Mississippi	Yalobusha
41	*mobilense/fraseri*	EW-2009-025	Texas	Bosque
42	*canadense*	EW-2010-031	Texas	Hardin
43	*lavendulare*	WS281413	Nebraska	Clay
44	*canadense*	NEB 313856	Nebraska	Platte
45	*canadense*	EW-2009-026	Texas	Hill
46	*canadense*	EW-2009-027	Texas	McLennan
47	*canadense*	EW-2009-040	Oklahoma	Love
48	*canadense*	EW-2009-041	Oklahoma	Bryan
49	*lavendulare*	EW-2009-058	Kansas	Little Walnut
50	*canadense*	EW-2010-005	Texas	Cooke
51	*canadense*	NEB 309014	Nebraska	Colfax
52	*canadense*	EW-2010-006	Texas	Collin
53	*canadense*	EW-2009-033	Texas	Houston
54	*canadense*	EW-2009-034	Texas	Houston
55	*canadense*	EW-2009-052	Missouri	Newton
56	*canadense*	NEB 269181	Nebraska	Hall
57	*canadense*	EW-2009-053	Missouri	Newton
58	*canadense*	EW-2010-008	Texas	Denton
59	*lavendulare*	EW-2009-053	Missouri	Newton
60	*lavendulare*	EW-2009-055	Kansas	Montgomery
61	*canadense*	NEB 309105	Nebraska	Colfax
62	*canadense*	EW-2010-011	Texas	Johnson
63	*canadense*	Jim Varnum s.n.	Texas	Dallas
64	*mobilense*	EW-2010-035	Texas	Sabine
65	*mobilense*	EW-2010-032	Texas	Jasper
66	*mobilense*	EW-2009-035	Texas	Houston
67	*canadense*	MISS 67037	Mississippi	Simpson
68	*mobilense*	WS 188684	Texas	Lamar
69	*mobilense*	EW-2010-031	Texas	Hardin
70	*canadense*	USCH 106130	South Carolina	Richland
71	*mobilense*	WS188690	Texas	Angelina
72	*mobilense*	EW-2009-037	Texas	Anderson
73	*mobilense*	EW-209-038	Texas	Anderson

#	Variety	Collection	State/Province	County
74	*canadense*	EW-2010-036	Louisiana	
75	*mobilense*	EW-2009-003	Texas	Montgomery
76	*mobilense/fraseri*	EW-2009-029	Texas	Limestone
77	*canadense*	EW-2009-003	Texas	Montgomery
78	*mobilense/fraseri*	EW-2009-008	Texas	Gonzales
79	*mobilense*	EW-2009-015	Texas	Bastrop
80	*mobilense*	EW-2009-013	Texas	Caldwell
81	*mobilense*	EW-2009-017	Texas	Grimes
82	*mobilense*	EW-2010-020	Texas	Goliad
83	*mobilense*	EW-2010-26	Texas	Refugio
84	*mobilense*	EW-2010-27	Texas	Victoria
85	*mobilense*	EW-2010-028	Texas	Jackson
86	*fraseri*	EW-2010-012	Texas	Bosque
87	*fraseri*	EW-2020-013	Texas	San Saba
88	*fraseri*	EW-2010-018	Texas	Bandera
89	*hyacinthoides*	EW-2010-007	Texas	Collin
90	*hyacinthoides*	EW-2010-010	Texas	Dallas
91	*fraseri*	WS 281376	Texas	Gonzales
92	*fraseri*	EW-2009-060	Kansas	Reno
93	*fraseri*	EW-2009-062	Kansas	Reno
94	*fraseri*	WS 281372	Texas	Comal
95	*hyacinthoides*	WS 168539	Texas	Dallas
96	*fraseri*	WS 286367	Kansas	Cloud
97	*canadense*	EW-2009-028	Texas	Limestone
98	*hyacinthoides*	WS188709	Texas	Bosque
99	*hyacinthoides*	WS281442	Texas	Dallas
100	*canadense*	EW-2010-017	Texas	Kerr
101	*fraseri*	EW-2010-015	Texas	Mason
102	*hyacinthoides*	EW-2010-006	Texas	Collin
103	*fraseri*	EW-2010-001	Texas	Cooke
104	*canadense*	EW-2009-019	Texas	Washington
105	*fraseri*	WS40119	Texas	Tarrant
106	*fraseri*	EW-2010-022	Texas	Gonzales
107	*canadense*	EW-2010-029	Texas	Chambers
108	*canadense*	WS181945	Texas	Travis
109	*canadense*	WS281357	Texas	Travis
110	*fraseri*	EW-2009-023	Texas	Burnett
111	*fraseri*	EW-2009-024	Texas	Burnett
112	*fraseri*	EW-2010-016	Texas	Llano
113	*canadense*	EW-2010-020	Texas	Goliad

#	Variety	Collection	State/Province	County
114	*canadense*	EW-2010-037	Louisiana	Webster
115	*ecristatum*	WS188704	Texas	Wilson
116	*ecristatum*	WS188714	Texas	Wilson
117	*canadense*	EW-2009-020	Texas	Travis
118	*canadense*	EW-2009-018	Texas	Washington
119	*canadense*	EW-2010-028	Texas	Jackson

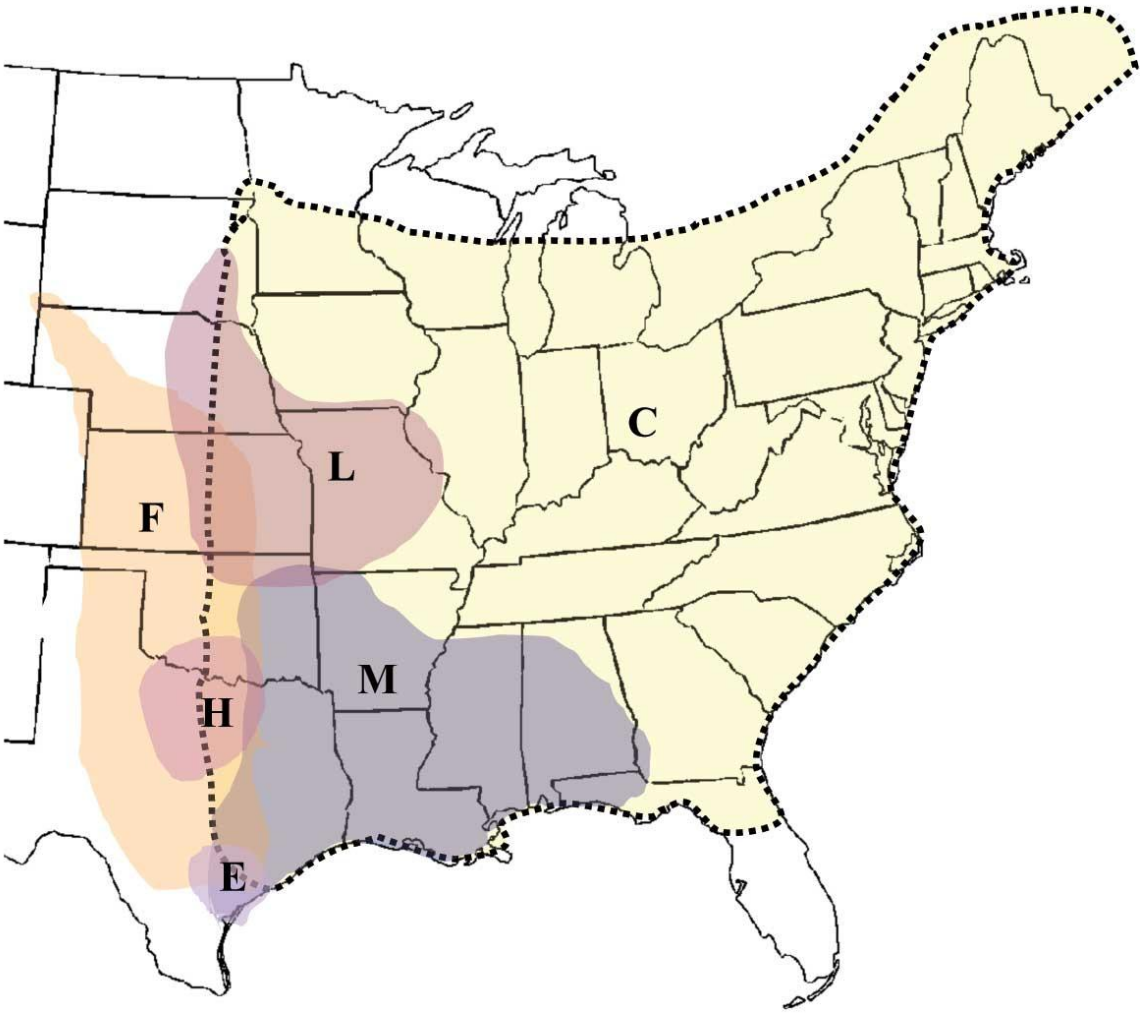

Figure 3.1 Geographic ranges of the six varieties of *A. canadense*. C=var. *A. c. canadense*, E=var. *A. c. ecristatum*, F=var. *A. c. fraseri*, H=var. *A. c. hyacinthoides*, L=var. *A. c. lavendulare* and M=var. *A. c. mobilense*. The dotted line highlights the range of the pseudoviviparous var. *A. c. canadense*. Redrawn from Ownbey and Aase (1955).

Figure 3.2 Photographs of the sexual (above) and asexual (below) forms of *A. canadense*. In this case, the pseudoviviparous plant bears some apparently normal flowers in addition to asexual bulbils.

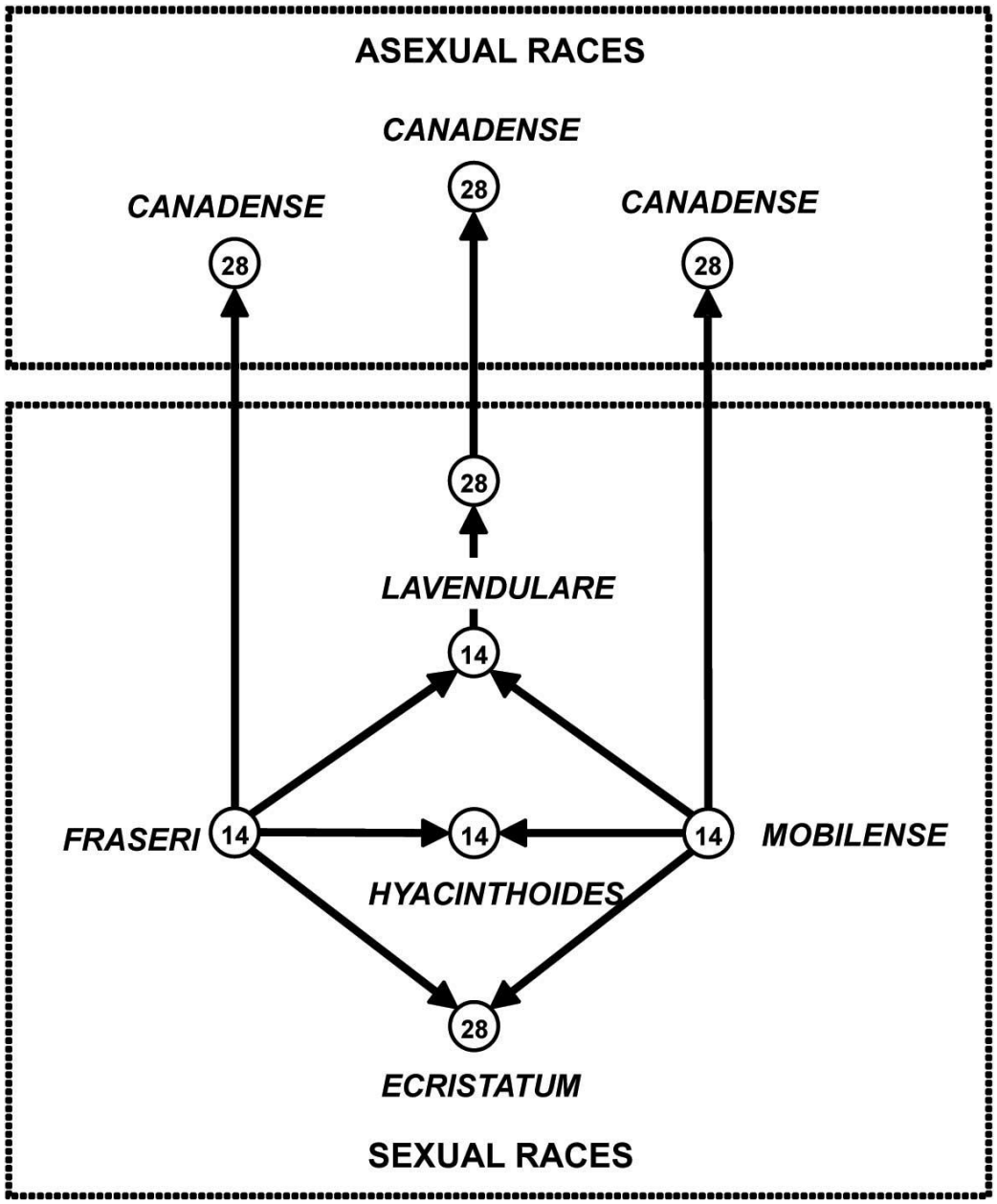

Figure 3.3 Proposed relationships among the six varieties and ploidy levels of *A. canadense*. Numbers in circles are somatic chromosome numbers. Redrawn from Ownbey and Aase, 1955.

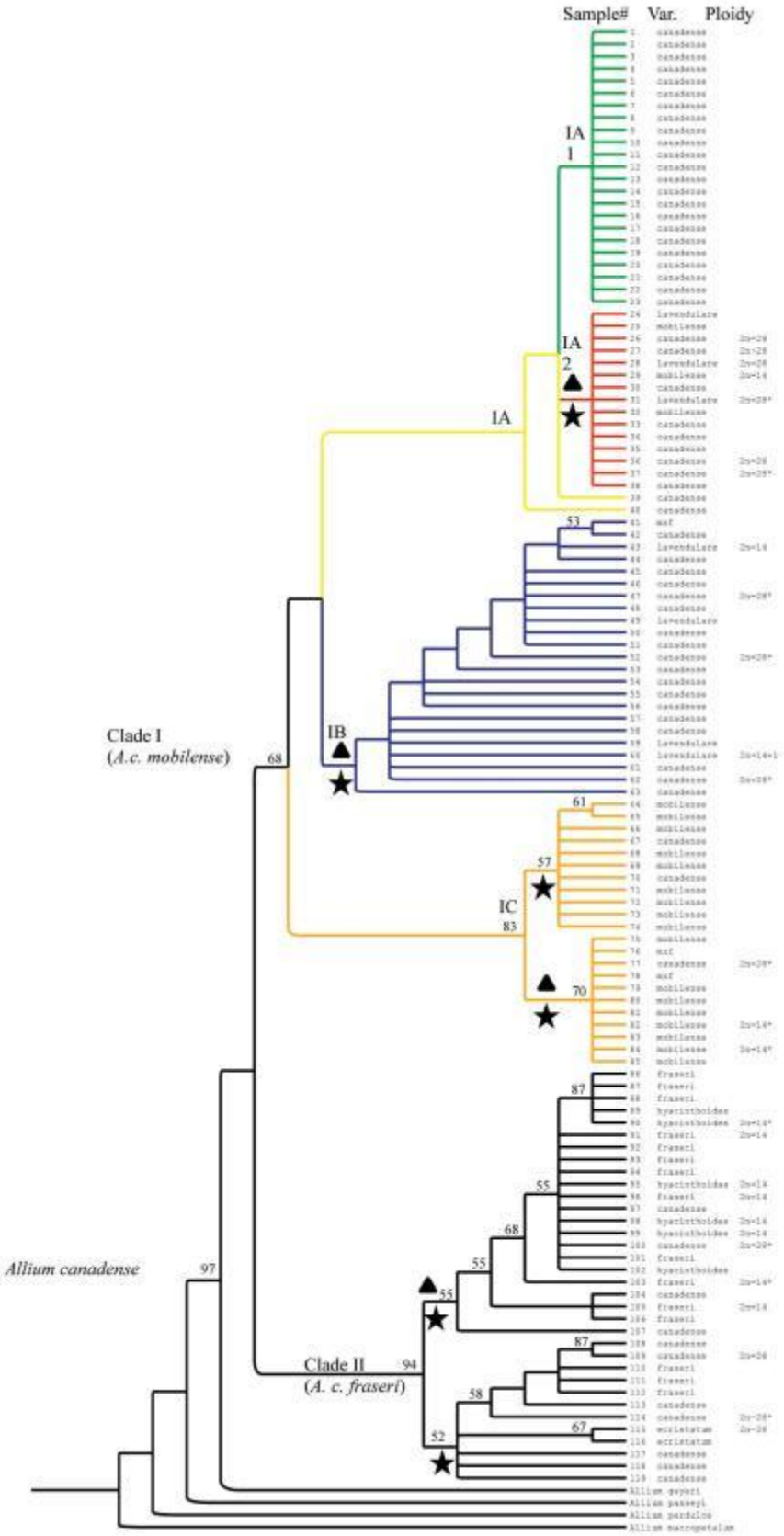
Sample#
Var.
Ploidy
IA
1
IA
2
IA
IB
IC
Clade I
(A.c. mobilense)
Clade II
(A. c. fraseri)
Allium canadense
53
68
61
57
83
70
87
55
68
55
55
97
94
87
58
67
52

Figure 3.4 Strict consensus tree of 1000 most parsimonious trees showing the phylogenetic relationships among 119 samples of *A. canadense*. MP bootstrap support values > 50% are shown at nodes. Sample numbers refer to those listed in Table 3.1. Chromosome numbers are shown to the right of taxon names. Chromosome counts determined during this study are marked with an asterisk (*); other chromosome counts were determined by Ownbey and Aase (1955). Approximate chromosome counts are denoted by (~). Stars indicate clades in which a transition between seminifery and pseudovivipary is inferred. Triangles indicate clades in which a transition between diploidy and polyploidy is inferred. Clade colors correspond to color coded samples on the map (Fig.3. 6).

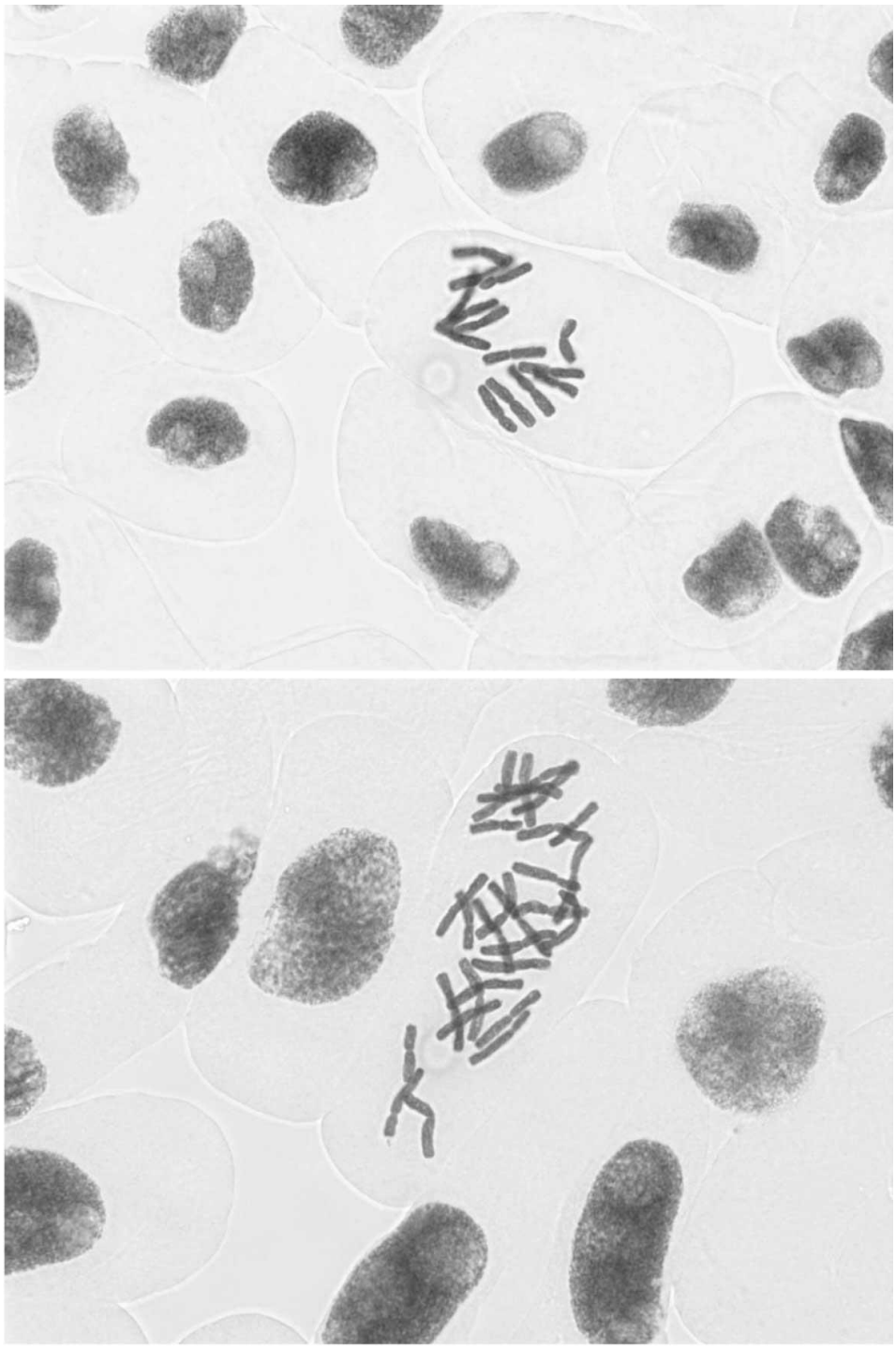

Figure 3.5 Photographs of *Allium canadense* root tip cells at metaphase. Above is the var. *A. c. fraseri* (EW-2009-063) showing the diploid (2n=14) chromosome complement. Below is the var. *A. c. canadense* (EW-2009-003) showing the tetraploid chromosome (2n=28) complement.

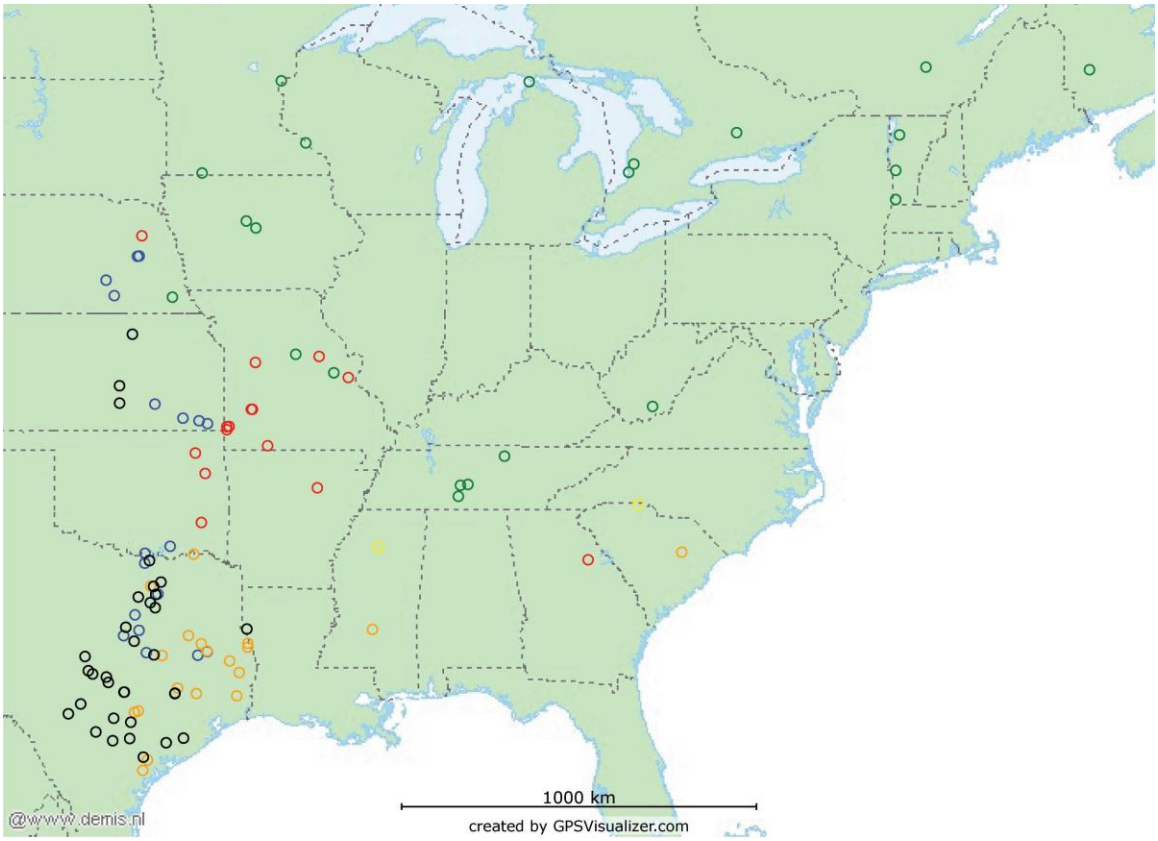

Figure 3.6 Geographic distribution of samples and of clades. Colors of circles correspond to colors of clades in the strict consensus tree (Fig. 3.4). For herbarium specimens that did not include latitude and longitude data, the latitude and longitude of the closest town reported on the herbarium label was used as an approximation of collection location. The map was created using the on-line mapping tool GPSVisualizer.com. The basemap was provided by DEMIS (www.demis.nl).

CHAPTER 4

CHLOROPLAST PHYLOGENOMICS OF 18 *ALLIUM* SPECIES USING GENOME SURVEY SEQUENCING

ABSTRACT

The approximately 850 species in genus *Allium* are currently divided among 15 subgenera and 72 sections based on morphology and on molecular phylogenetic analysis of the nuclear ITS region. Although many of the evolutionary relationships among these groups are resolved and well-supported, questions remain about relationships among some subgenera, and among sections within subgenera. Resolution of these issues in *Allium* will likely require a substantial increase in the number of phylogenetic characters available for analysis. As the cost of next generation sequencing methods decrease, these technologies are rapidly being applied to unresolved taxonomic questions in plants because they deliver megabases of data, including entire plastid genome sequences. We sequenced 18 *Allium* species representing eleven subgenera from across the genus in a multiplex arrangement of six samples per lane on an Illumina GAIIx Genome Analyzer. The goals of this research were to 1) assess the capacity of genome survey sequencing (GSS) methods to capture entire plastome sequences in a genus of flowering plants characterized by very large nuclear genome size, 2) provide the first intrageneric phylogenetic study of a non-grass monocot genus based on GSS methods, and 3) generate genomic resources for *Allium* that can be used for future investigations of phylogenetic relationships and molecular evolution with the genus. Despite the large nuclear genome size in *Allium*, chloroplast sequences recovered by GSS represented between 0.3% and 4.38% of total reads, and from these data we obtained DNA sequence for 80 plastid protein coding genes. Maximum likelihood phylogenetic analysis of 79 of these genes shows that the evolutionary relationships across the genus are largely congruent with previously published phylogenies, with a couple of exceptions. Bootstrap

support for clades was 100% in all cases but three, providing a marginal improvement to those reported to date. Because of restricted taxon sampling, this study does not provide resolution of any of the outstanding taxonomic questions in *Allium*. However, the success of this small-scale project should encourage *Allium* systematists to embark on GSS projects with deeper taxon sampling aimed at resolving these outstanding questions. Sequence data obtained during this study can also be used in future investigations of evolutionary relationships and molecular evolution within the genus *Allium*.

INTRODUCTION

Intrageneric relationships in genus *Allium*

The genus *Allium* has been the subject of numerous molecular phylogenetic studies over the last decade and a half (Dubouzet and Shinoda, 1999; Friesen, Fritsch, and Blattner, 2006; Nguyen, Driscoll, and Specht, 2008; Gurushidze, Fritsch, and Blattner, 2010; Li et al., 2010; Wheeler et al., Submitted). Of particular importance to a genus-wide classification is the work by Friesen et al. (2006), who proposed a new intrageneric classification based on evidence from DNA sequence variation in the internal transcribed spacer (ITS) region of the nuclear ribosomal repeat. As a result of this research, more subgenera than had previously been recognized were named and some sections were shown to be artificial. Currently, 15 subgenera and 72 sections are recognized (Friesen, Fritsch, and Blattner, 2006).

Many of the relationships among subgenera and sections of *Allium* have been resolved and are well-supported by previously published phylogenies. However, relationships among four subgenera in the third evolutionary line, *Allium, Cepa, Polyprason,* and *Reticulatobulbusa* have not been resolved (Friesen, Fritsch, and Blattner, 2006; Li et al., 2010). Li et al. (2010) suggest that this is due to rapid speciation within this lineage. In addition, there are outstanding questions about relationships among sections within other subgenera. Of particular relevance to North American *Allium* is the question of the geographic origin and center of dispersal of subgenus *Amerallium,* which has representatives in the New World, in the Mediterranean Basin and the Middle East, and in eastern Asia (Fritsch and Friesen, 2002; Li et al., 2010). Currently, the relationship between two Old World clades that correspond to the Mediterranean and the east Asian centers of diversity is not well-supported, making it difficult to draw inferences about historical distribution and migration of taxonomic groups within and among continents.

Resolution of these outstanding questions will require a substantial increase in the number of characters provided by the few loci examined to date. DNA sequence variation found in rapidly evolving non-coding, intergenic spacers in the plastid genome is often higher than in coding regions (Shaw et al., 2007), and has been successfully applied to intrageneric phylogenetic studies in many taxa, including some clades within *Allium* (Gurushidze, Fritsch, and Blattner, 2010; Li et al., 2010; Wheeler et al., Submitted). However, the application of rapidly evolving plastid intergenic spacers to cross-genus *Allium* phylogenetics is often hampered by numerous insertions and deletions that make sequence alignment and homology assessment difficult, and sometimes impossible. This may be due to the fact that the genus *Allium* is considered to be an ancient angiosperm lineage that was already well-differentiated in the early Tertiary (Hanelt et al., 1992; Dubouzet and Shinoda, 1999; Li et al., 2010), allowing for the accumulation of many mutations in non-coding regions over evolutionary time. One approach to this challenge in cross-genus phylogenetic studies in *Allium* is to rely on protein coding sequences in the plastid genome, which are functionally constrained and therefore evolve more slowly than intergenic spacers. However, the number of informative characters/gene is expected to be relatively low and therefore a large number of loci would be needed to provide the number of characters necessary to resolve evolutionary relationships within the genus.

Next Generation sequencing and molecular phylogenetics

In the past decade, there have been significant developments in DNA sequencing methods and technology that have improved our capacity to investigate patterns of genetic variation across a wide range of biological fields. In particular, traditional Sanger sequencing methods (Sanger, Nicklen, and Coulson, 1977), which currently produce single DNA sequence reads of up to 800-1000 nucleotides long from PCR amplified loci, are rapidly being replaced by a suite of closely related methods commonly referred to as 'Next Generation' (NG) or 'Massively Parallel' (MP) sequencing (reviewed in: Ansorge, 2009; Lister, Gregory, and Ecker, 2009;

Pettersson, Lundeberg, and Ahmadian, 2009; Metzker, 2010). The crucial technical advance common to these methods is that they employ parallel sequencing reactions to synthesize millions of highly redundant short DNA sequence reads. These short sequence reads (~100bp) are then assembled into long contiguous sequences (contigs) by means of 'reference-guided' (e.g. YASRA, Ratan, 2009) or 'de novo' assembly algorithms (e.g. NextGene; Softgenetics, State College, Pennsylvania).

The amount of sequence data resulting from these methods is orders of magnitude more than that obtained by Sanger sequencing, at a fraction of the cost per nucleotide (Steele and Pires, 2011) and has the capacity to increase the number of characters available for phylogenetic analyses. The effect of data matrix length (number of nucleotide characters) on the percentage of well-supported clades (>95%) has been tested using matrices generated by Sanger sequencing methods; in a meta-analysis of intrageneric phylogenetic studies based on chloroplast data matrices with an average length of 2574 nucleotides, Parks et al. (2009) found that data matrix length was directly and significantly correlated with clade support. With this in mind, the increase in data matrix length provided by MP sequencing methods may improve support values for phylogenetic relationships that have remained poorly-supported by single or even multi-locus sequence analysis.

Molecular phylogenetic studies published to date using MP sequencing methods to sequence entire chloroplast genomes do indeed demonstrate a notable increase in highly supported clades (>95% bootstrap) compared to previous studies based on fewer characters generated by Sanger sequencing of one or even multiple chloroplast loci (Parks, Cronn, and Liston, 2009; Zhang, Ma, and Li, 2011). However, MP sequencing methods have not been widely applied to molecular phylogenetic studies that aim to reconstruct the evolutionary history of multiple related organisms. This may be due to the fact that, until recently, the per taxon cost of sequencing was prohibitive for most systematics labs. In addition, extensive manipulation of

these very large datasets is required before phylogenetic analysis can be conducted, and becoming familiar with these bioinformatics methods represents a steep learning curve. Finally, these methods are relatively new and it is only recently that many research institutions have acquired access to the technology required to conduct phylogenetic studies using these sequencing methods.

Genome Survey Sequencing

Genome survey sequencing (GSS) presents a methodologically simplified and cost-effective alternative for phylogenomic projects that seek to include numerous taxa (Nock et al., 2011; Steele and Pires, 2011). Rather than relying on enrichment procedures to increase the concentration of target DNA (i.e. PCR, chloroplast isolations), GSS methods take a shotgun approach by randomly sequencing fragments from total DNA extractions. Genome regions that are highly represented in total DNA (i.e. chloroplast genome, mitochondrial genome, nuclear ribosomal RNA, repeat regions) are preferentially sequenced based on their relative abundance. DNA fragments from each taxon can be tagged with unique adapter sequences and multiple samples can be sequenced at once in a multiplex arrangement, lowering the sequencing cost per taxon substantially. Although the data that result from this multiplex sequencing arrangement provide lower genomic coverage and redundancy per taxon than other MP methods, GSS has been successfully used to recover entire chloroplast genome sequences across a wide range of monocot taxa and to provide increased support for phylogenetic relationships at the ordinal and familial levels (Givnish et al., 2010; Nock et al., 2011; Steele et al., Submitted). A concern in species with large nuclear genomes is that GSS may not recover plastid sequences because they constitute a smaller proportion of total genomic DNA extractions than species with smaller nuclear genomes. However, based on a comparison of 50 species in the order Asparagales, the percentage of plastid sequences recovered using GSS methods is not negatively correlated with nuclear genome size (Steele et al., Submitted).

Aims

The aim of this study was to use GSS methods to obtain complete plastid genome sequences from multiple *Allium* species, representing all three evolutionary lineages in the genus. Using these data, our goal was to reconstruct the evolutionary relationships among *Allium* subgenera and sections based on plastid protein coding gene sequences and to compare these results with previously published phylogenetic analyses. Since the taxonomic position and number of species we sampled in this study was restricted, we do not anticipate that our results will resolve any of the outstanding taxonomic questions in *Allium* mentioned above. However, data generated by this small-scale study 1) provided the first intrageneric phylogenetic study of a non-grass monocot genus based on GSS methods, 2) assessed the capacity of GSS methods to capture entire plastome sequences in a genus of flowering plants characterized by very large nuclear genome size, and 3) generated genomic resources for *Allium* that can be used for future investigations of phylogenetic relationships and molecular evolution with the genus.

MATERIALS AND METHODS

Taxon sampling

To assess the congruence of a whole chloroplast genome phylogeny with previously published phylogenies, we included one species from each of following eleven subgenera: *Allium*, *Amerallium*, *Anguinum*, *Caloscordum*, *Cepa*, *Cyathophora*, *Melanocrommyum*, *Microscordum*, *Nectaroscordum*, *Reticulatobulbosa*, and *Rhizirideum* (Table 4.1) as delimited by Friesen et al. (2006). We are interested in identifying variable chloroplast regions for future studies aimed at resolving evolutionary relationships and outstanding biogeographic questions in subgen. *Amerallium*. To this end, we included seven species from five sections of *Amerallium* (*Amerallium*, *Arctoprason*, *Bromatorrhiza*, *Lophioprason*, and *Narkissoprason*) with

representation from the Eurasian and North American centers of diversity. To investigate chloroplast sequence variation between closely related species, we included two species from subgen. *Cepa*, section *Cepa*, subgen. *Amerallium* section *Lophioprason* and subgen. *Amerallium* section *Amerallium*. Outgroups included four other members of the family Amaryllidaceae: *Gillesia graminea*, *Tulbaghia violacea* (subfamily Alliodeae), *Crinum asiaticum* (subfamily Amarylidoideae), and *Agapanthus africanus* (subfamily Agapanthoideae). For six species (*Allium cepa*, *Allium fistulosum*, *Agapanthus africanus*, *Crinum asiaticum*, *Gillesia graminea* and *Tulbaghia violacea*), library preparation, sequence assembly and annotation was done by P. Roxanne Steele and these data have been previously published in Steele et al. (Submitted).

Illumina DNA library preparation and sequencing

Sequencing libraries were made using reagents from an NEB Prep kit E6000L (New England Biolabs, Inc., Ipswich Massachusetts) and the following protocol. Total genomic DNA was extracted from ~20 mg silica dried leaves from each sample using a DNeasy Plant DNA Extraction Kit (Qiagen, Germantown Maryland). Approximately five μg of DNA was diluted in double distilled water (ddH_20) to a final concentration of 6.25 ng/μL, in a total volume of 800 μL. The DNA was sheared into ~300 base pair fragments using a Bioruptor® sonication device (Diagenode, Denville New Jersey), run for 24 min. Sheared DNA was purified and concentrated using a QIAquick PCR Purification kit (Qiagen, Germantown Maryland). To assess the approximate size of sheared DNA fragments, 200 ng DNA was run against a 100 base pair size standard on a 2% low-melt agarose gel at 120 V for one hour, stained with ethidium bromide and then visualized by UV light.

End-repair of the ~300 bp DNA fragments, A-tailing and ligation of adapter oligos were carried out as per the instructions in the NEB Prep kit E6000L. In order to sequence six individual samples in the same lane on the Illumina Genome Analyzer, each of the six samples destined for the same lane was assigned a different adapter tag. Ligation products were purified

and eluted in a volume of 20 μL. This entire volume was run against a 100 base pair standard on a 2% low-melt agarose gel. The portion of the gel containing ~300 base pair DNA fragments was excised using an x-tracta Disposable Gel Extraction Tool (USA Scientific, Ocala Florida) and DNA was purified from this 200 μL gel plug with a Gel Extraction kit (Qiagen, Germantown Maryland). The ~300 base pair DNA fragments were enriched by PCR amplification in a 50 μL volume containing 3 μL ligation product, 20 μL ddH_2O, 25 μL master mix (NEB kit), and 1 μL each of 25 μM forward and reverse primers (forward 5'- AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC* T -3'; reverse 5'- CAA GCA GAA GAC GGC ATA CGA GAT CGG TCT CGG CAT TCC TGC TGA ACC GCT CTT CCG ATC* -3'). The thermal cycling conditions were as follows: 98°C for 30 sec, followed by 15 cycles of 98°C for 10 sec, 65°C for 30 sec, and 72°C for 30 sec, with a final extension step of 72°C for 5 min. The amplification product was run on a 2% low-melt agarose gel and the ~300 base pair target product was excised and purified with the tools described above. Sample libraries were quantified and then sequenced on a GAIIx Genome Analyzer (Illumina, San Diego California), six samples per lane, by the staff at the DNA Core Facility at the University of Missouri-Columbia.

Sequence assembly, alignment, and annotation

Sequencing reads from each lane were parsed into six bins representing the six samples per lane using a perl script that recognizes the unique adapter tag assigned to each sample. This perl script also removed the adapter sequence from each read, all reads with more than five ambiguous nucleotide positions (N), the sequence quality information, and constructed a Fasta text file of the reads for each sample. For each sample, contiguous sequences (contigs) were assembled with the reference-based assembly program YASRA (Ratan, 2009), using the *Allium cepa* chloroplast genome sequence as a reference (Steele et al., Submitted), with the similarity threshold between sample reads and the reference sequence set at medium (85%). For each

sample, the multiple contigs constructed in YASRA were exported to the program Geneious v.5.4 (Drummond et al., 2011) and reassembled to the *A. cepa* reference. At positions in the assembly where there were gaps between sample contigs relative to the reference sequence, a place-holder 'N' was assigned to indicate missing data. A consensus sequence of each sample assembly was constructed.

The initial assemblies described above were constructed using reads that had not been quality-trimmed based on the Phred scores reported for each nucleotide in the original Fastq file. To correct this, we mapped quality-trimmed data back to each consensus sequence and visualized the depth of coverage across the entire chloroplast genome in NextGene (Softgenetics, State College, Pennsylvania). In areas where there were less than two quality-trimmed reads, we changed the consensus sequence to 'N' in those positions to indicate missing data if this region was a coding sequence. The consensus sequence of each sample including these amendments was constructed and exported for alignment.

An alignment of consensus sequences from the 22 samples was made in Geneious using the MAFFT (Katoh, Asimenos, and Toh, 2009) plug-in. The limits of protein coding genes were identified in each sample by comparison with the previously annotated *A. cepa* chloroplast genome sequence (Steele et al., Submitted) and annotated accordingly. In cases where stop codons (UAG, UAA, or UGA) appeared prematurely relative to the reference, the 3' end of the gene in question was annotated at this truncation point in that particular sample.

Phylogenetic Analysis

The DNA sequence of 79 genes (no intron, intergenic spacer, rRNA or tRNA sequences were used) was concatenated into one multi-locus sequence and exported for phylogenetic analysis. A search for the best tree was conducted with maximum likelihood (ML) as the optimality criterion using the on-line tool RAxML Black Box (Stamatakis, Hoover, and Rougemont, 2008). The model of nucleotide evolution implemented by this program is GTR +

CAT, which approximates the GTR + Γ model (General Time Reversible with a gamma distribution of rate heterogeneity). Support for clades was assessed with 100 replicate bootstrap searches, as implemented in RAxML.

RESULTS

Sequence data

The number of sequence reads obtained from each sample ranged from 2, 170, 547 in *A. geyeri*, to 11, 255, 997 in *A. flavidum* (Table 4.1). The percentage of total reads that assembled to the chloroplast reference genome and is therefore considered to be of plastid origin, ranged from 0.30% in *A. neriniflorum* to 4.38% in *A. cepa.* The average sequence length was uniform within individual sequencing runs, but varied from 77 bp to 117 bp among runs. The average depth of sequence coverage across all regions of the reference genome ranged from 14.7 X in *A. bulgaricum* to 214.2 X in *A. sativum*. The number of contigs assembled in YASRA from each sample varied from two in *A. canadense*, *A. sativum* and *A. wallichi* to 32 in *A. bulgaricum*.

When we assembled the quality-trimmed reads from each sample to the consensus sequence for that sample made from reads that were not quality-trimmed, we found numerous places where the trimmed sequences no longer provided at least 2X coverage in coding regions. The number of instances within a sample ranged from one in *A. tolmiei* to 52 in *A. neriniflorum.* As noted in the methods, nucleotides at these positions were replaced with Ns to indicate missing data.

In total, 80 protein coding genes were recovered and annotated for all 22 taxa including: *accD, atpA, atpB, atpE, atpF, atpH, atpI, ccsA, cemA, clpP, infA, matK, ndhA, ndhB, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK, petA, petB, petD, petG, petL, petN, psaA, psaB, psaC, psaI, psaJ, psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL,*

psbM, psbN, psbT, psbZ, rbcL, rpl14, rpl16, rpl2, rpl20, rpl22, rpl23, rpl32, rpl33, rpl36, rpoA, rpoB, rpoC1, rpoC2, rps11, rps12, rps14, rps15, rps16, rps18, rps19, rps2, rps3, rps4, rps7, rps8, ycf1, ycf2, ycf3, ycf4, and *ycf68.* The *infA* gene was excluded from phylogenetic analysis because this gene was attenuated in all samples by stop codons, leaving a total of 79 genes for phylogenetic analysis.

Phylogeny

The aligned length of the concatenated dataset for 22 samples was 70, 260 nucleotides. The most likely tree recovered by maximum likelihood analysis (-ln = 181817.939680) is shown in Figures 4.1 and 4.2. Genus *Allium* is supported as monophyletic with 100% ML bootstrap support as is each of the three evolutionary lines. The seven species sampled from subgenus *Amerallium* form a monophyletic group with 100% bootstrap support. The phylogenetic relationships among the species of *Allium* sampled are also supported by 100% bootstrap values, except in three cases: the sister relationship of *A. sativum* and *A. tuvinicum* is supported with 89%, the sister relationship of *A. cristophii* and *A. neriniflorum* is supported with 85%, and the sister relationship of *A. wallichii* and (*A. insubricum* + *A. ursinum*) is supported with 95%.

DISCUSSION

Intrageneric relationships in *Allium*

The phylogenetic relationships among subgenera and sections of *Allium* recovered using 79 protein coding chloroplast genes are largely congruent with previously published studies, with two notable exceptions. In our phylogeny, subgenera *Allium* and *Rhizirideum* are sisters whereas the subgenus *Rhizirideum* is sister to the clade containing the subgenera (*Allium* + *Cepa* + *Polyprason* + *Reticulatobulbosa*) in Friesen et al. based on the nuclear ITS region (2006); this relationship was also recovered by Li et al. (2010) based on a dataset that combined DNA

sequence from ITS and the chloroplast *rps16* intron. In addition, the sister relationships among the subgenera *Anguinum*, *Caloscordum* and *Melanocrommyum* differ in our phylogeny when compared to that of Friesen et al. (2006). In particular, subgenus *Anguinum* is sister to the clade containing subgenera (*Caloscordum* + *Melanocrommyum*) in our phylogeny whereas the subgenus *Caloscordum* is sister to the clade containing subgenera (*Anguinum* + *Melanocrommyum*) in Friesen et al. (2006).

These topological differences may be explained in a couple of ways. First, our taxon sampling was sparse compared to Friesen et al. (2006) and differences in taxon density among analyses have been shown to shift topological relationships among taxa (Heath, Hedtke, and Hillis, 2008). Future studies based on GSS data should endeavor to include more species than we were able to include here in order to avoid this problem. Second, since our analysis is based on DNA sequence variation in plastid genes, this topology reflects the evolutionary history of the plastid genome which exhibits uniparental inheritance in most angiosperms (Birky Jr, 1995) and may not track the evolutionary history of the biparentally inherited nuclear ITS locus used in previous studies (Wendel and Doyle, 1998).

Except for the topological differences noted above, our study reiterates and supports previous findings about evolutionary relationships in the genus *Allium*. Friesen et al. (2006) reported maximum parsimony bootstrap values of 100% for all but one of the clades we recovered. For the sister relationship between section *Bromatorrhiza* and sections (*Narkissoprason* + *Arctoprason*) subgenus *Amerallium*, bootstrap support increased from 86% MP bootstrap support to 95% ML bootstrap support. Because our taxon sampling was limited, we were unable to address any of the unresolved relationships mentioned above with GSS methods. However, we hope that the success of this small-scale project will provide incentive to *Allium* systematics researchers to embark on GSS projects with increased taxonomic sampling directed at resolving outstanding taxonomic questions.

Research avenues in genus *Allium*

The sequence data obtained during this project can be used in a variety of ways in future research. First, the plastid sequence data assembled and analyzed here may be used to identify variable coding and non-coding regions that can be used as phylogenetic markers in *Allium*. This information would be valuable to *Allium* systematists in different parts of the world working at various taxonomic levels in the genus. Second, we may be able to recover other highly represented genomic regions from these data, including mitochondrial genes or genomes, nuclear ribosomal repeat regions, and other repetitive elements. Of particular interest is the evolutionary history of a group II intron in the mitochondrial gene *cox2*, which is putatively *trans*-spliced in *A. cepa*, but *cis*-spliced in *A. sativum* and *A. ramosum* (Kim and Yoon, 2010). In plants, most mitochondrial group II introns are *cis*-spliced and this is only the second time *trans*-splicing has been identified and it may be of recent evolutionary origin. Third, the proliferation of repetitive elements in the *Allium* genome, is considered to be responsible for the large nuclear genome size in the genus and for the wide variation of nuclear genome size among even closely related species (Evans, James, and Barnes, 1983). By comparing the DNA sequence of repeats present in *Allium* with other members of the subfamily Allioideae (*Gillesia*, *Gethyum*, *Ipheon*, *Miersia*, *Speea*, *Tristagma*, and *Tulbaghia*) for which we will soon have GSS sequence data, we may be able to learn something about the evolution of repetitive elements in these taxa.

GSS methods and intrageneric phylogenetic relationships

This is the first study to use GSS methods to investigate intrageneric phylogenetic relationships in a non-grass monocot. Other studies have demonstrated the phylogenetic power of these methods in the grass genera *Oryza* (Poaceae) (Nock et al., 2011) and *Bambusa* (Poaceae) (Zhang, Ma, and Li, 2011) and in the conifer genus *Pinus* (Pinaceae) (Parks, Cronn, and Liston, 2009). GSS studies in monocots have otherwise focused on ordinal and familial relationships to date (Givnish et al., 2010; Steele et al., Submitted).

Despite the large nuclear genome size characteristic of the genus *Allium* (Ohri, Fritsch, and Hanelt, 1998; Pires et al., 2006), GSS methods recovered sufficient plastid DNA sequence to assemble all plastid protein coding genes for all of the species sampled. This result supports the assertion of Steele et al. (Submitted), that large nuclear genomes do not restrict the ability of GSS methods to recover plastid sequences at a level that is appropriate for phylogenetic analysis. In addition to recovering the entire plastid genome, an advantage of GSS methods is the recovery of DNA sequence from all genome partitions within plants including the mitochondrial and nuclear genomes. Phylogenetic comparisons across components of these three genomic partitions using GSS has already been demonstrated (Steele et al., Submitted) and the recovery of these genomic components promises to improve with increased depth of sequencing.

The cost per taxon of GSS is rapidly decreasing and the depth of sequencing is rapidly increasing due to continued advances in the application of MP sequencing. It is not unreasonable to suggest that these methods will eventually replace the isolation and amplification of individual loci by PCR for many phylogenetic applications. The laboratory protocol for GSS employs standard molecular biology methods that are easily carried out in most molecular systematics labs. However, the lack of accessible bioinformatics tools for the manipulation and analysis of these data still restricts the use of these methods to a small subset of systematics researchers. This impediment will be lessened as new tools are developed and the systematics community starts to adopt these bioinformatics methods.

ACKNOWLEDGEMENTS

Thank you to N. Friesen (Botanical Garden of the University of Osnabrück, Germany) for guidance regarding taxonomic sampling and leaf material for *Allium* species not native to North America (except *A. cepa* and *A. fistulosum*). Thank you to P. R. Steele for contributing annotated plastome sequences for six species and for instruction on Illumina library preparation and data analysis. Thanks to T. Arias, P. Edger and D. Mayfield for their advice and assistance. This research was supported by an NSF grant (DEB 0829849), a DFG grant (DFG: FR 1431/4-1), and by a University of Missouri Life Sciences Fellowship.

LITERATURE CITED

ANSORGE, W. J. 2009. Next-generation DNA sequencing techniques. *New Biotechnology* 25: 195-203.

BIRKY JR, C. W. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: Mechanisms and evolution. *Proceedings of the National Academy of Sciences of the United States of America* 92: 11331-11338.

DRUMMOND, A. J., B. ASHTON, S. BUXTON, M. CHEUNG, A. COOPER, C. DURAN, M. FIELD, et al. 2011. Geneious, version 5.4.

DUBOUZET, J. G., and K. SHINODA. 1999. Relationships among Old and New World *Alliums* according to ITS DNA sequence analysis. *TAG Theoretical and Applied Genetics* 98: 422-433.

EVANS, I. J., A. M. JAMES, and S. R. BARNES. 1983. Organization and evolution of repeated DNA sequences in closely related plant genomes. *Journal of Molecular Biology* 170: 803-826.

FRIESEN, N., R. M. FRITSCH, and F. R. BLATTNER. 2006. Phylogeny and new intrageneric classification of *Allium* L. (Alliaceae) based on nuclear ribosomal DNA ITS sequences. *Aliso* 22: 372-395.

FRITSCH, R. M., and N. FRIESEN. 2002. Evolution, domestication and taxonomy. *In* H. D. Rabinovich AND L. Currah [eds.], *Allium* Crop Science – Recent Advances, vol. 1, 5-30. CABI Publishing.

GIVNISH, T. J., M. AMES, J. R. MCNEAL, M. R. MCKAIN, P. R. STEELE, C. W. DEPAMPHILIS, S. W. GRAHAM, et al. 2010. Assembling the tree of the monocotyledons: plastome sequence phylogeny and evolution of Poales. *Annals of the Missouri Botanical Garden* 97: 584-616.

GURUSHIDZE, M., R. M. FRITSCH, and F. R. BLATTNER. 2010. Species-level phylogeny of *Allium* subgenus *Melanocrommyum*: incomplete lineage sorting, hybridization and *trnF* gene duplication. *Taxon* 59: 829-840.

HANELT, P., J. SCHULTZE-MOTEL, R. M. FRITSCH, J. KRUSE, H. I. MAAB, H. OHLE, and K. PISTRICK. 1992. Infrageneric grouping of *Allium* - the Gatersleben approach. The genus *Allium*: taxonomic problems and genetic resources, Gatersleben, Germany: 107-123.

HEATH, T. A., S. M. HEDTKE, and D. M. HILLIS. 2008. Taxon sampling and the accuracy of phylogenetic analyses. *Journal of Systematics and Evolution* 46: 239-257.

KATOH, K., G. ASIMENOS, and H. TOH. 2009. Multiple alignment of DNA sequences with MAFFT. *Methods in molecular biology (Clifton, N.J.)* 537: 39-64.

KIM, S., and M. K. YOON. 2010. Comparison of mitochondrial and chloroplast genome segments from three onion (Allium cepa L.) cytoplasm types and identification of a trans-splicing intron of cox2. *Current Genetics* 56: 177-188.

LI, Q. Q., S. D. ZHOU, X. J. HE, Y. YU, Y. C. ZHANG, and X. Q. WEI. 2010. Phylogeny and biogeography of *Allium* (Amaryllidaceae: Allieae) based on nuclear ribosomal internal transcribed spacer and chloroplast rps16 sequences, focusing on the inclusion of species endemic to China. *Annals of Botany* 106: 709-733.

LISTER, R., B. D. GREGORY, and J. R. ECKER. 2009. Next is now: new technologies for sequencing of genomes, transcriptomes, and beyond. *Current Opinion in Plant Biology* 12: 107-118.

METZKER, M. L. 2010. Sequencing technologies the next generation. *Nature Reviews Genetics* 11: 31-46.

NGUYEN, N. H., H. E. DRISCOLL, and C. D. SPECHT. 2008. A molecular phylogeny of the wild onions (*Allium*; Alliaceae) with a focus on the western North American center of diversity. *Molecular Phylogenetics and Evolution* 47: 1157-1172.

NOCK, C. J., D. L. E. WATERS, M. A. EDWARDS, S. G. BOWEN, N. RICE, G. M. CORDEIRO, and R. J. HENRY. 2011. Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnology Journal* 9: 328-333.

OHRI, D., R. M. FRITSCH, and P. HANELT. 1998. Evolution of genome size in *Allium* (Alliaceae). *Plant Systematics and Evolution* 210: 57-86.

PARKS, M., R. CRONN, and A. LISTON. 2009. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biology* 7: 84.

PETTERSSON, E., J. LUNDEBERG, and A. AHMADIAN. 2009. Generations of sequencing technologies. *Genomics* 93: 105-111.

PIRES, J. C., I. J. MAUREIRA, T. J. GIVNISH, K. J. SYTSMA, O. SEBERG, G. PETERSON, J. I. DAVIS, et al. 2006. Phylogeny, genome size, and chromosome evolution of Asparagales. *Aliso* 22: 287-304.

RATAN, A. 2009. Assembly algorithms for next generation sequence data. PhD Dissertation, Pennsylvania State University.

SANGER, F., S. NICKLEN, and A. R. COULSON. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* 74: 5463-5467.

SHAW, J., E. B. LICKEY, E. E. SCHILLING, and R. L. SMALL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *Am. J. Bot.* 94: 275-288.

STAMATAKIS, A., P. HOOVER, and J. ROUGEMONT. 2008. A rapid bootstrap algorithm for the RAxML web-servers. *Systematic Biology* 75: 758-771.

STEELE, P. R., and J. C. PIRES. 2011. Biodiversity assessment: State-of-the-art techniques in phylogenomics and species identification. *American Journal of Botany* 98: 415-425.

STEELE, P. R., K. L. HERTWECK, D. MAYFIELD, J. LEEBENS-MACK, and J. C. PIRES. Submitted. Quality and quantity of data recovered from massively parallel sequencing: examples in Asparagales and Poaceae. *American Journal of Botany*.

WENDEL, J. F., and J. J. DOYLE. 1998. Phylogenetic incongruence: window into genome history and molecular evolution. *In* D. E. Soltis, P. S. Soltis, AND J. J. Doyle [eds.], Molecular systematics of plants II: DNA sequencing. Kluwer Academic Publishers.

WHEELER, E. J., S. MASHAYEKHI, D. W. MCNEAL, J. T. COLUMBUS, and J. C. PIRES. Submitted. Molecular systematics of *Allium* subgenus *Amerallium* (Amaryllidaceae) in North America. *American Journal of Botany*.

ZHANG, Y. J., P. F. MA, and D. Z. LI. 2011. High-throughput sequencing of six bamboo chloroplast genomes: phylogenetic implications for temperate woody bamboos (Poaceae: Bambusoideae). *PLoS ONE* 6.

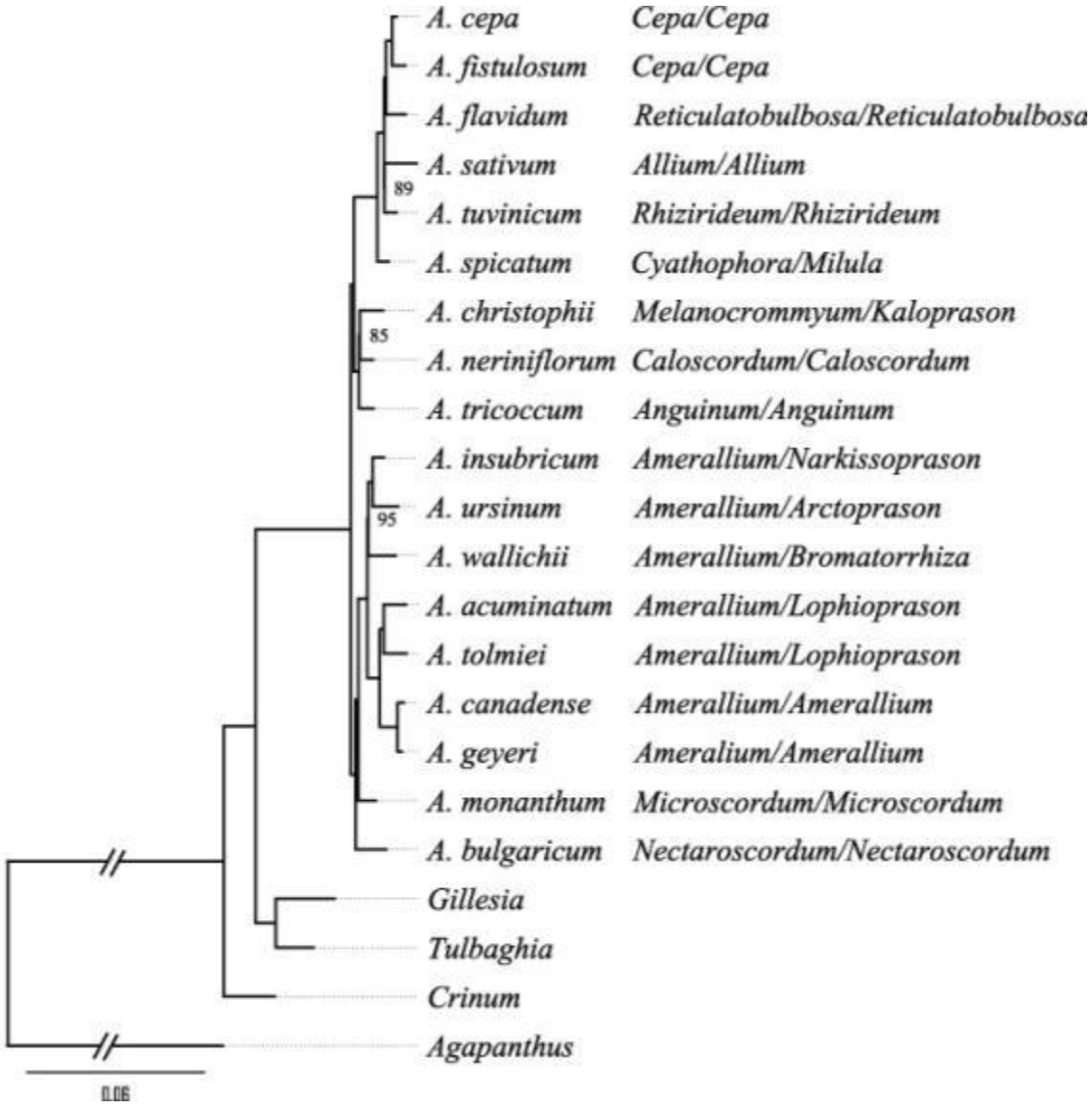

Figure 4.1 Maximum likelihood best tree (-ln = 181817.939680) showing relationships among *Allium* samples. Subgenus/Section are shown in the right column. ML bootstrap values <100% are shown; all others are 100%. Units of scale bar = substitutions per site.

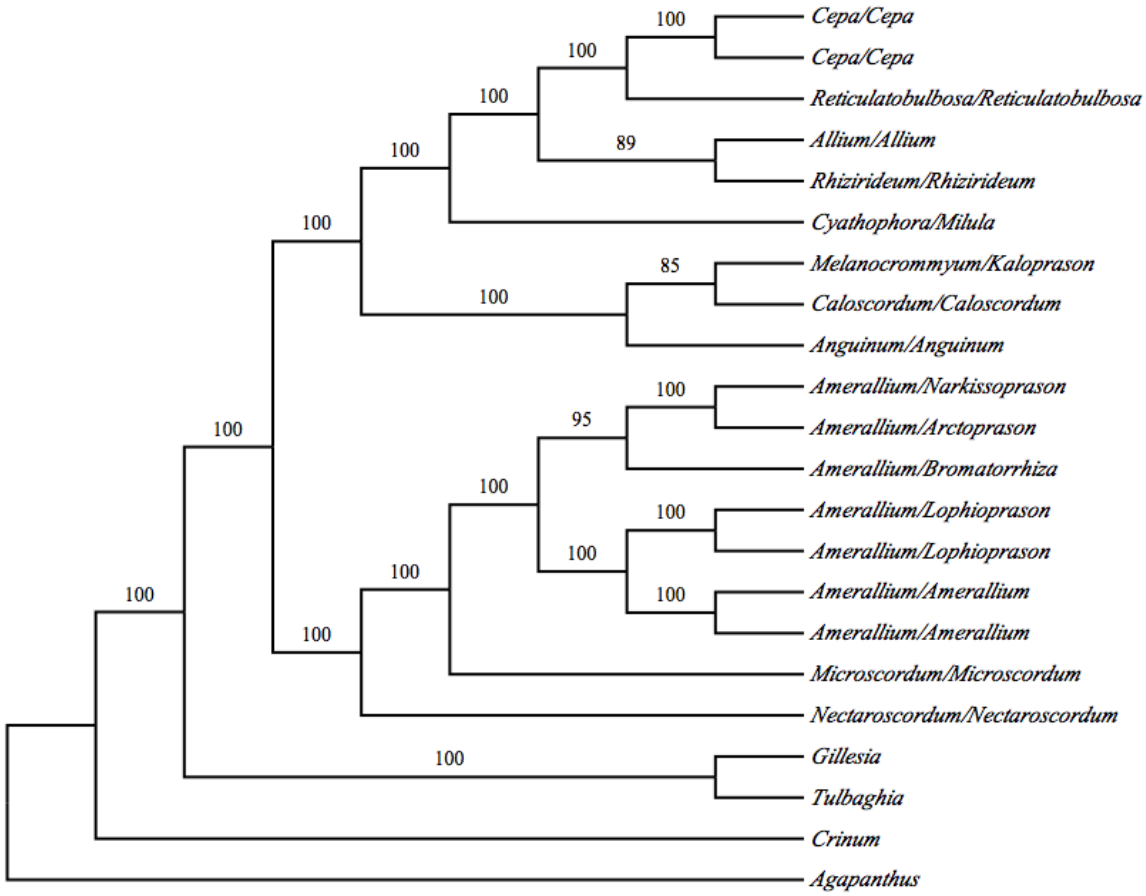

Figure 4.2 Cladogram showing evolutionary relationships among *Allium* subgenera and sections sampled recovered in maximum likelihood analysis.

Table 4.1. Species of *Allium* and outgroups sampled for this study and summary of the results from sequencing and assembly.

Species	Subgenus	Section	Voucher	Total # reads	% cp reads	Average length	Average coverage	# contigs
Allium canadense	*Amerallium*	*Amerallium*	EW-2009-062	4958673	2.11	77	63.3X	2
Allium geyeri	*Amerallium*	*Amerallium*	EW-2008-046	2170547	1.76	77	23.2X	17
Allium acuminatum	*Amerallium*	*Lophioprason*	EW-2008-036	3723924	3.43	77	77.3X	3
Allium tolmiei	*Amerallium*	*Lophioprason*	EW-2008-041	3985464	3.09	77	74.5X	9
Allium wallichii	*Amerallium*	*Bromatorrhiza*	NF 10-20-0007-20	3425721	1.31	117	41.3X	2
Allium ursinum	*Amerallium*	*Arctoprason*	NF 09-20-0005-20	3459311	0.49	117	15.6X	24
Allium insubricum	*Amerallium*	*Narkissoprason*	NF 02-30-0026-40	4298410	2.05	117	81.1X	4
Allium bulgaricum	*Nectaroscordum*	*Nectaroscordum*	NF Tax 3220	2896541	0.55	117	14.7X	32
Allium tricoccum	*Anguinum*	*Anguinum*	EW-2011-002	3457789	2.17	117	69.3X	4
Allium tuvinicum	*Rhizirideum*	*Rhizirideum*	NF 06-07-0030-10	3950772	0.84	117	30.7X	5
Allium christophii	*Melanocrommyum*	*Kaloprason*	NF 03-44-0016-20	10139265	0.96	97	74.4X	8
Allium flavidum	*Reticulatobulbosa*	*Reticulatobulbosa*	NF 08-31-0007-20	11255997	0.82	97	70.9X	6
Allium monanthum	*Microscordum*	*Microscordum*	NF 06-31-0131-20	9444914	1.86	97	134.3X	7
Allium neriniflorum	*Caloscordum*	*Caloscordum*	NF Tax 2797	8130786	0.3	97	18.4X	18
Allium sativum	*Allium*	*Allium*	NF 09-24-0002-80	8898325	3.15	97	214.2X	2
Allium spicatum	*Cyathophora*	*Milula*	NF 04-11-0111-10	9693231	1.62	97	120.1X	4
Allium fistulosum	*Cepa*	*Cepa*	Steele 1094	5098306	1.89	80	62.1X	13
Allium cepa	*Cepa*	*Cepa*	?	2795386	4.38	80	78.3X	1
Outgroups								
Tulbaghia violacea			?	2381172	3.08	80	46.8X	8
Gillesia graminea			MWC 450K	2915826	1.91	80	35.5X	3
Crinum asiaticum			Asiatica Nursery	2230364	2.7	80	38.1X	10
Agapanthus africanus			Steele 1099	1281941	3.1	80	25.3X	13

CHAPTER 5

CONCLUSIONS

Each of the three research projects I conducted for my dissertation used molecular phylogenetic or phylogenomic methods to investigate evolutionary relationships in the genus *Allium*, each at a different taxonomic scale, and each aimed at answering a different suite of questions. As such, the results from each chapter stand alone and there is no single conclusion to be drawn from these three studies, in support or refutation of a particular hypothesis. What unites these projects is my long-standing interest in the evolutionary history of the genus as a whole, and of North American species in particular. As is often the case, resolution of the problems at hand can lead to new questions. Below, I briefly summarize the results from each research project and identify some new research directions.

The molecular phylogeny of North American *Amerallium* species (Chapter 2) provides an evolutionary context with which to test the classification system proposed by Marion Ownbey, based on morphological characters (Saghir, Mann, and Ownbey, 1966). Although some of Ownbey's morphological alliances represent natural groups, other alliances are not monophyletic based on the molecular phylogeny. As a result, it will be necessary to revise the current classification system to some degree, as suggested in the discussion section of chapter two. However, since these alliances were proposed as informal groups which do not have taxonomic standing, we must look to the formal classification system of Hamilton Traub (1968c) to make changes. The classification systems of Ownbey's alliances and Traub's subsections are largely congruent, so this should be relatively straightforward for most clades. In collaboration with Dr. Dale W. McNeal, the taxonomic expert in North American *Allium*, I hope to publish these changes soon.

The molecular phylogeographic study of *Allium canadense* populations (Chapter three), demonstrates that pseudovivipary is a labile character in this species, and that the transition from seminifery to pseudovivipary is correlated with the transition to polyploidy. This is only the second study to examine and to document multiple origins of pseudovivipary in a phylogenetic context, and will hopefully provide an impetus for other researchers to examine this phenomenon in their particular taxonomic groups. Other species in the genus *Allium* also exhibit this reproductive strategy, and it would be interesting to expand this study to other *Allium* species to test the multiple origins hypothesis more broadly. As mentioned in the discussion section of chapter three, alternative splicing of the *gaLFY* mRNA transcript has been shown to be associated with the development of bulbils in the inflorescence of garlic (*A. sativum*) (Rotem et al., 2007; Rotem et al., 2011). Whether this mechanism is also responsible for pseudovivpary in *A. canadense* could be tested by repeating the experiments of Rotem et al. (2007), in which they looked for differential expression of alternatively spliced *gaLFY* transcripts during the development of seminiferous and pseudoviviparous inflorescences.

The molecular phylogenomic study of relationships among *Allium* subgenera and sections using GSS methods (chapter four) is best seen as a pilot study, which demonstrates that GSS methods can recover entire plastid genome sequences from species with large nuclear genomes, such as *Allium*. To make this research project meaningful to the systematics community, more species need to be added, especially species from the third evolutionary line in *Allium*, within which there remain unresolved relationships among subgenera and sections. As mentioned in the discussion section of chapter four, there are numerous research directions that could be pursued using the GSS data at hand and that we are currently generating. Of particular interest to me is the identification of variable coding and non-coding regions in the chloroplast that can be used for phylogenetic inference at different taxonomic levels in *Allium*. Although GSS methods may eventually replace the sequencing of individual loci in molecular phylogenetic studies, there are

still many questions that can be easily and perhaps best answered using good old Sanger sequencing methods. In addition, there are *Allium* systematists working in other parts of the world who have small research budgets and would benefit from the identification variable regions across the genus, and the publication of *Allium* specific primers.

Allium by no means characterizes a model system that can be easily subjected to genetic analysis in order to understand the inheritance of phenotypic traits; it takes between one and three years for most species to reach reproductive maturity, and the size of nuclear genome is among the largest in angiosperms (Bennett and Leitch, 2011). Among economically important cultivated species, however, these impediments have not stood in the way of genetic analysis using quantitative trait loci (Ohara et al., 2009), and monosomic addition lines (Shigyo et al., 1996). Among wild species and populations, there are a number of evolutionary phenomena that occur in *Allium* that make it an attractive research system for investigating standing phenotypic and genotypic variation.

First, the large nuclear genome size of *Allium* is of interest itself. The proliferation of repetitive elements is considered to play a significant role in the elevated nuclear genome size in *Allium* (Evans, James, and Barnes, 1983). Phylogenetic analysis of the relationships among repeat elements within the genus, and between *Allium* and other genera in the subfamily Allioideae, may provide insight into the evolution of repeat elements themselves, and into the pattern of repeat expansion in the genus *Allium*. We are currently generating GSS data for six more genera in subfamily Allioideae in order to investigate this.

Second, among North American *Allium* species, there is a high incidence of intraspecific ploidy variation, with 12% of species (10/81) maintaining both diploid and (largely) tetraploid individuals and/or populations (McNeal and Jacobsen, 2002). However, there is little phenotypic divergence between diploid and tetraploid races. This suggests that tetraploid races have arisen recently in these species, and therefore, there has not been sufficient time for phenotypic

divergence to occur among reproductively isolated races. Alternatively, there may be ongoing gene flow among ploidy levels, which would homogenize any phenotypic divergence among races. These alternative hypotheses could be tested through experimental interploidy crosses, aimed at assessing the level of reproductive isolation among ploidy levels.

Third, there appear to be high levels of genetic divergence among *Allium* species, accompanied by relatively little morphological diversification, given the age of the genus (Friesen, Fritsch, and Blattner, 2006). This is evident from examination of the relationships among the 15 genera of the subfamily Allioideae; the genus *Allium* is sister to the clade containing all other genera in the subfamily (Fay and Chase, 1996) i.e. whereas the sister lineage diversified into 14 morphologically distinct genera, the genus *Allium* remained relatively undifferentiated and morphologically unified. This stands in contrast to patterns seen in rapidly diverging angiosperm lineages, in which high levels of morphological diversification are accompanied by little genetic differentiation. Given an appropriate monocot fossil for calibration, it would be interesting to compare the rate and tempo of morphological differentiation among Allioideae genera.

In closing, the projects I have completed during my PhD research, have broadened and deepened my understanding of the genus *Allium*. I hope I will be lucky enough to continue conduct research in this genus in the future.

LITERATURE CITED

BENNETT, M. D., and I. J. LEITCH. 2011. Nuclear DNA amounts in angiosperms: Targets, trends and tomorrow. *Annals of Botany* 107: 467-590.

EVANS, I. J., A. M. JAMES, and S. R. BARNES. 1983. Organization and evolution of repeated DNA sequences in closely related plant genomes. *Journal of Molecular Biology* 170: 803-826.

FAY, M. F., and M. W. CHASE. 1996. Resurrection of Themidaceae for the Brodiaea alliance, and recircumscription of Alliaceae, Amaryllidaceae and Agapanthoideae. *Taxon* 45: 441-451.

FRIESEN, N., R. M. FRITSCH, and F. R. BLATTNER. 2006. Phylogeny and new intrageneric classification of *Allium* L. (Alliaceae) based on nuclear ribosomal DNA ITS sequences. *Aliso* 22: 372-395.

MCNEAL, D. W., and T. D. JACOBSEN. 2002. *Allium*, Flora of North America North of Mexico, vol. 26, 224-276. Oxford University Press, New York.

OHARA, T., H. TSUKAZAKI, Y. S. SONG, T. WAKO, K. I. YAMASHITA, and A. KOJIMA. 2009. Mapping of quantitative trait loci controlling seedling growth in bunching onion (Allium fistulosum L.). *Journal of the Japanese Society for Horticultural Science* 78: 436-442.

ROTEM, N., R. DAVID-SCHWARTZ, Y. PERETZ, I. SELA, H. RABINOWITCH, M. FLAISHMAN, and R. KAMENETSKY. 2011. Flower development in garlic: the ups and downs of *gaLFY* expression. *Planta* 233: 1063-1072.

ROTEM, N., E. SHEMESH, Y. PERETZ, F. AKAD, O. EDELBAUM, H. D. RABINOWITCH, I. SELA, et al. 2007. Reproductive development and phenotypic differences in garlic are associated with expression and splicing of LEAFY homologue gaLFY. *J. Exp. Bot.* 58: 1133-1141.

SAGHIR, A. R. B., L. K. MANN, and M. OWNBEY. 1966. Composition of volatiles in relation to taxonomy of American *Alliums*. *American Journal of Botany* 53: 477-484.

SHIGYO, M., Y. TASHIRO, S. ISSHIKI, and S. MIYAZAKI. 1996. Establishment of a series of alien monosomic addition lines of Japanese bunching onion (Allium fistulosum L.) with extra chromosomes from shallot (A. cepa L. Aggregatum group). *Genes and Genetic Systems* 71: 363-371.

TRAUB, H. P. 1968c. The subgenera, sections and subsections of *Allium* L. *Plant Life* 24: 147-163.

VITA

Erica Wheeler was born in Auckland, New Zealand, and grew up in Vancouver, British Columbia. After more than a decade working in the silviculture industry in British Columbia, Erica attended the University of Victoria in 1996 to pursue an undergraduate degree in the Department of Biology (2000), with a focus on botany. For two years after completing her B.Sc., she worked in a number of positions as a research technician on projects that included the molecular systematics of flowering plants, the response of plant communities to volcanic ash accumulation, and the effects of introduced amphibians on native species. In 2002, Erica began a Masters degree at the University of Victoria with Dr. Geraldine Allen, which she completed in 2006. During this time, she held a position as the Assistant Herbarium Curator at the University of Victoria Herbarium (UVIC), where she managed collections, trained volunteers, and provided technical support for undergraduate courses. Erica's research on the phylogeography of the rare slim-leaf onion (*Allium amplectens*) sparked her interest in the genus *Allium*, and inspired her to pursue a doctoral degree with this group of plants as a focus. In 2006, Erica moved to the University of Missouri to begin doctoral studies with J. Chris Pires, which she completed in 2011.

Erica's research interests center on the evolution and systematics of flowering plants, and on the application of molecular phylogenetic methods to investigate the evolution of traits such as asexuality and polyploidy. Although the genus *Allium* has been the focus of her graduate work, she is broadly interested in these questions across all flowering plants.