AN ABSTRACT OF THE DISSERTATION OF

Brian J. Knaus for the degree of Doctor of Philosophy in Botany and Plant Pathology presented on September 16, 2008.

Title: A Fistful of *Astragalus*: Phenotypic and Genotypic Basis of the Most Taxon Rich Species in the North American Flora.

Abstract approved:

_____

Rich Cronn Aaron Liston

The study of the infra-specific ranks (i.e., subspecies and variety) can be considered the study of the process of speciation. *Astragalus lentiginosus* Douglas ex Hooker (Fabaceae) is the most taxon rich species in the U.S. flora currently including 40 taxonomic varieties. These varieties were described using traditional taxonomic methods primarily through the work of M.E. Jones, P.A. Rydberg, and R.C. Barneby. Presently, three methodologies are employed to test the taxonomic hypotheses presented by these monographers as well as the more simplistic hypothesis of isolation by distance. These datasets include morphometric measurements made from herbarium samples, chloroplast simple sequence repeats (CpSSR), and amplified fragment length polymorphisms (AFLP). Morphometric analysis is an explicit test of the morphological characters employed to circumscribe taxa within the group. Analyses indicated cohesiveness to morphological varieties but an absence of significant discontinuity to define these taxa. This indicates that the varieties may make cohesive groups, but their distinction is arbitrary. Analysis of CpSSRs indicates that the greatest amount of variance is described by among population and within population components of variance as opposed to varietal or Rydbergian sectional components. Morphological and CpSSR data present an incongruence between

phenotypic and molecular data, a pattern which has been interpreted as the possiblesignal of selection. The AFLP dataset demonstrates a greater concordance of molecular data to morphological taxonomy, but also presents some interesting differences. An east-west transect at approximately 36.5° latitude demonstrates a pattern where *A. l.* var. *variabilis* occurs in the basins while *A. l.* var. *fremontii* occurs in the ranges. Genetic differentiation along this transect appears to support taxonomy despite a confounding geographic pattern, suggesting differential gene-flow among morphological varieties or that the AFLP dataset has captured the signal of selectively constrained loci (or regions linked to these loci). Multiple datasets provide multiple perspectives on evolution within *A. lentiginosus*. A significant pattern of morphological diversity accompanied by two molecular datasets which do not support this phenotypic diversity are interpreted as a potential instance of selective divergence among the varieties of *A. lentiginosus*.

A Fistful of *Astragalus*: Phenotypic and Genotypic Basis of the Most Taxon Rich Species in the North American Flora

by
Brian J. Knaus

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented September 16, 2008
Commencement June 2009

Doctor of Philosophy dissertation of Brian J. Knaus presented on September 16, 2008.
APPROVED:

Co-Major Professor, representing Botany and Plant Pathology

Co-Major Professor, representing Botany and Plant Pathology

Chair of the Department of Botany and Plant Pathology

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Brian J. Knaus, Author

ACKNOWLEDGEMENTS

The author would like to thank:

My family: John H. Hoffman†, Pat Hoffman Friedman†, M. Jane Knaus, Alan L. Knaus, Darin A., Rosie, Wyatt and Ellie Knaus, Kevin Knaus, Samuel N. Friedman.

My committee: Rich Cronn, Aaron Liston, Mitch Cruzan, Tom Kaye and Steven Sharrow.

The Cronn and Liston Lab groups: John Syring, Ann Wilyard, Rebecca Huot, Sue Huber, Angie Rodrieguez, Jane Marler, Paul Severns, Matt Parks and Stephen Meyers.

Department of Botany and Plant Pathology, OSU: Grad students, Lynda Ciuffetti, Diane Simpson and the office staff.

USDA FS Pacific Northwest Field station CSL: Matt Horning, Nancy Mandel, Randy Johnson, Chris Poklemba, and Brad St. Clair.

The Bassil Lab Group: Nahla Bassil, Wambui Nygren, and April Dring.

Botanical (and otherwise) inspiration: Dana A. York, Mark Fishbein, Lucinda McDade, Patrick and Christine Whitmarsh, Gretel Enck, Kathy Davis and Dell Heter, Jenna Straface, and Sue Haig.

The McCune Lab: Bruce McCune, Emily Holt, Sarah Jovan, Heather Root, and Heather Lintz.

The Herbarium Staff (OSC): Richard Halse, Ken and Henny Chambers, Katie Mitchell, and Barbara Wilson.

This work would not have been possible without their support!

CONTRIBUTION OF AUTHORS

Brian J. Knaus conceived the research topic, designed experiments, developed methodology, implemented the experimental design, analyzed the data and authored the present manuscript.

Rich Cronn contributed to the design of experiments, the development of methodology, the implementation of experiments and provided editorial contribution to the analysis and authoring of the present manuscript.

Aaron Liston contributed to the design of experiments, the development of methodology, the implementation of experiments and provided editorial contribution to the analysis and authoring of the present manuscript.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

1 Introduction

In biology the 'species' is considered 'distinct' and is often believed to be the fundamental unit of biology (Darwin 1859; Dobzhansky 1951; Stearn 1957; Fisher 1958; Mayr and Ashlock 1991; Coyne and Orr 2004; Rieseberg et al. 2006). While much rhetoric has been devoted to the debate on species concepts (e.g., Mayden 1997), a consensus has yet to present itself. Here I take the opinion that a species is somehow distinct from other species. The metric for defining 'distinct' (e.g., morphological, reproductive, genetic, phylogenetic) appears to be the source of confusion. If we take the species to be the unit of distinction, infra-taxa (the subspecies and the variety) are consequently non-distinct. The process in which a group of organisms diverge from being one cohesive group to becoming two or more distinct groups is the process of speciation. Here I explore infra-taxa as groups of organisms which may be on a path to speciation. I employ morphology, chloroplast simple sequence repeats and amplified fragment length polymorphisms to explore differing perspectives on the degree of discontinuity within the legume species *Astragalus lentiginosus* Douglas ex Hooker (Fabaceae).

1.1 INFRATAXA AS INCIPIENT SPECIES

Concepts used to define a species appear to be somewhat nebulous and vary from professional opinion to quantifiable measures. The following section reviews select species concepts relevant to the present study.

Linnaeus— The species concept employed by Linnaeus differed from other concepts covered here in that his system focused on the order of creation (Stearn 1957) rather than evolution. Linnaeus recognized the difficulty in identifying criteria for the definition of species and emphasized the importance of professional judgment. He also stressed the importance of constancy of character in identifying species. Varieties were optional to Linnaeus and included variation from the typical form which may be due to the effects of nature (e.g., phenotypic plasticity). He also suggested the recognition of different sexes of dioecious taxa as varieties. While infra-taxa were

present in Linnaeus' system, focus was at the levels of taxa such as the species and genus.

Darwin— Darwin recognized the difficulty in defining the species and, similar to Linnaeus, indicated it required a degree of professional judgment. While Darwin does mention the species as a 'distinct act of creation' (Darwin 1859) this seems conflicted with later statements pertaining to evolution. He recognized the presence of geographical races and that these may be interpreted differently by different taxonomists. Darwin's views on the infra-taxa are perhaps best summarized by the following statement:

> *And I look at varieties which are in any degree more distinct and permanent, as steps leading to more strongly marked and more permanent varieties; and at these latter, as leading to sub-species and, and to species.*

It is this view that may have first established the idea of infra-taxa as incipient species. Also important is the concept that the infra-taxa should not be considered as determinant, Darwin recognized that these intermediate states may go extinct. Extinction of intermediate forms in a morphological cline may lead to the distinction of entities which formerly intergraded. Furthermore, was the issue of whether varieties could be subsumed back into a whole; Darwin's view that varieties could persist in the wild contrasted with Wallace's view that they would revert to a relatively homogenous species. Darwin borrowed heavily from examples of domestic organisms to demonstrate selective divergence, which he felt was evidence of the initiation of speciation.

Wallace— Wallace felt varieties gave rise to successive variation, and that successful varieties would replace unsuccessful varieties to eventually become the species (1858). However, he felt that this was not true for domestic animals due to a relaxation of selective pressures encountered in the domestic environment. Wallace (1858) felt that domestic organisms would revert to a whole when returned to natural

conditions because the novelties of domestication did not confer an adaptive advantage in nature.

Clausen, Keck and Hiesey— Clausen Keck and Hiesey (1939) attempted to integrate the ideas of internal and external barriers to hybridization proposed by Dobzhansky (1951) with existing taxonomic concepts and their own pioneering work in common gardens and reciprocal transplants. They considered reproductive isolation to define the cenospecies which may correspond to taxonomic species or higher levels. Concepts such as allopatric versus sympatric distributions, coupled with hybrid vigor, defined four more categories below the cenospecies. These concepts attempted to capture process in their definitions. Yet the added complexity of unfamiliar terms combined with a general lack of information on interfertility present when many taxa were described may have prevented their widespread use.

Dobzhansky— Dobzhansky (1951) considered the determination of taxonomic rank to be the job of the taxonomist, which he was not. His perspectives on the nature of the species represent a divergence in perspective between evolutionary biologists, who are primarily interested in differences within the species, and taxonomists, who are primarily interested in the delimitation of species and higher level groupings. Dobzhansky described a continuum from races, varieties, to subspecies with a focus on process over pattern. Discussion on the importance of reproductive isolation to the process of speciation was present in Darwin's work. Dobzhansky helped formalize this discussion through categorization of different forms of reproductive isolation. Although allopatry prevents contemporary gene exchange it does not rule out the potential for future gene exchange. The ultimate form of reproductive isolation is the presence of intrinsic barriers (physical or physiological) that eliminate the potential for gene flow. This is important philosophically because it means that a species cannot share an evolutionary fate with another species through the sharing of genetic material. This identifies an idealized metric for defining species, but leaves the infra-taxon as groups of organisms which have some arbitrary level of interbreeding.

Fisher— Fisher (1958) viewed the species as genetically uniform where most loci were the lineal descendents of a favorable mutation. He posited a scenario where clinal diversity could be distributed from extremes of a species' range where a trait which is favorable at one extreme may be deleterious at the other extreme. Through gene flow this cline may be maintained, but Fisher characterized it as an unstable state, as if it were a piece of elastic stretched to its maximal extent before it fractured. The entities along this gradient could be considered varieties. Fission of this unstable state would result in the differentiation of these moieties into species.

Wright— Wright was largely concerned with issues of inbreeding and the correlation among genotypes due to descent; however, as an evolutionist he also addressed the formation of species (Wright 1978). Wright categorized processes the lead to species as including mutational pressure, random genetic drift, mass selection and selective differentiation, with the latter two being of primary importance. The idea of mass selection is largely attributed to Fisher. Wright is perhaps best known for his shifting balance theory, which he categorized under selective diffusion (Wright 1978). A critical difference between the ideas of Fisher and Wright was their assumptions on effective population size. Fisher considered effective population sizes to be relatively large suggesting that random genetic drift (i.e., sampling error) would be relatively inconsequential. Wright felt that effective population sizes were relatively small, indicating that random genetic drift played an important role in evolution. Wright described a scenario of slight inbreeding within a population, which leads to a continuous wobble around its adaptive optimum, the process of drift keeping it from occupying its peak (Wright 1932). This wobble may allow the population to drift to a point where it can be captured by a different peak, allowing movement across the adaptive landscape, from peak to peak and through adaptive valleys. If this peak is greater than the previous peak, small amounts of gene flow among populations at both peaks will slowly pull all populations toward the highest peak (Wright 1932). The result is the movement of populations from local optima towards greater optima.

When this process involves the fission of populations which ancestrally tracked a single optimum to populations which track different optima it results in the process of speciation.

Stebbins— In his treatment for the North American Flora (1993), Stebbins expressed that in practice species are delimited on the basis of morphological discontinuities between several visible traits. This idea was based on the concept that a species is any group of individuals (or of dried specimens) that an experienced taxonomist decides to call a species. This indicates that in practice, species concepts really haven't changed much since the time of Linnaeus, and that they should be largely considered as arguments based on professional opinion. Stebbins also included the idea that species are systems of populations which resemble each other yet contain genetically different ecotypes which could be arranged in a continuous series. These allopatric infraspecific categories are usually recognized as infra-taxa.

These species concepts appear uniform in the idea that a species is a relatively homogenous group which is somehow well differentiated from its congenerics. Conversely, the infra-taxon may represent subdivisions of these homogenous groups which are perhaps somewhat arbitrary. These non-discrete entities are sometimes recognized taxonomically at the infraspecific ranks (i.e., the subspecies and variety). A focal point of this dissertation is the characterization of the most taxon rich species in the U.S. flora as a potential instance of an entity in the process of speciation.

1.2 TAXONOMIC REVIEW OF *ASTRAGALUS LENTIGINOSUS*

Astragalus lentiginosus Douglas ex Hooker (Fabaceae) is a diverse set of plant forms varying from prostrate to erect, glabrous to tomentose, annual to perennial and with flowers from purple to white. A unifying character among most of the varieties is an inflated, bilocular pod with a unilocular beak. This unilocular beak dehisces at maturity to allow seeds to be dispersed from the fruit. Some varieties have slightly inflated pods where the abaxial suture does not extend all the way to the funicular

flange, resulting in a semi-bilocular condition. The epithet *lentiginosus* refers to the red mottling commonly (but not consitently) found on the pods, which resemble freckles. This legume is distributed primarily throughout the intermountain region of North America (Figure 1.1). It ranges from Northern Sonora in the south to southern British Columbia in the north, and from the California Coast Range to the western Rocky Mountains. While many varieties are relatively widespread, several are endemic to specialized habitats such as desert seeps, inland dunes, limestone habitats and montane ridges.

Many of what are currently know as varieties of *Astragalus lentiginosus* were originally described as species. Marcus E. Jones was the first to recognize the similarities among these taxa and arranged them as varieties of one species (Jones 1895, 1923). Per Axel Rydberg employed a very different species concept, stating that he did not believe in infra-taxa (Rydberg 1929b). This resulted in his raising Jones's varieties to species in the genera *Cystium* Steven and *Tium* Medikus (Rydberg 1929a). A novelty of Rydberg's treatment is the concept of sections which have been maintained in the keys of subsequent treatments, even if this was not explicitly stated. These sections are here referred to as Rybergian groups. Barneby (1945) is perhaps best seen as a moderator of the opinions of Jones and Rydberg. This resulted in the recognition of many of Rydberg's taxa at the rank of variety within the species *Astragalus lentiginosus*, as proposed by Jones (1895, 1923). Subsequent treatments include Barneby (1964, 1989), Isely (1998) and Welsh (2007). Each of these treatments was slightly different, containing between 36 and 42 taxa (Table 1.1). Here I recognize 40 taxa (Table 1.2; see Chapter 2 for discussion).

1.3 HYPOTHESES TO BE TESTED

In order to formalize this study in the context of the scientific method I propose several hypotheses. First I propose the null hypothesis of global panmixia. This phenomenon would be manifested in an observed pattern which is indistinguishable from a random pattern throughout the range of *A. lentiginosus*. An alternate

hypothesis is that of isolation by distance, where the marker system employed displays high correlation at small distances and low correlation at large distances. A second alternate hypothesis is sectional (regional) panmixia, characterized by significant among-section differentiation while within-section diversity is relatively random. A final hypothesis is that of varietal organization, characterized by among variety differentiation, which may or may not be nested within sections.

Over a century since Darwin's book (1859), the study of the processes leading to the diverse forms we call species remains an intellectual pursuit as interesting as it ws when Darwin first proposed it. Here I explore the taxon rich legume species *Astragalus lentiginosus* which, containing 40 varieties, appears to be a species with the potential to fracture into multiple new species. I provide three methodologies, each with differing taxonomic (Table 1.2) and geographic breadth to address the above hypotheses: morphology, chloroplast simple sequence repeats and amplified fragment length polymorphisms.

1.4 LITERATURE CITED

Barneby, R. C. 1945. *Pugillus Astragalorum* IV: the section diplocystium. Leaflets of Western Botany 4:65-147.

—. 1964. Atlas of North American *Astragalus*. Memoirs of the New York Botanical Garden 13:1-1188.

—. 1989, Fabales: Intermountain Flora, v. 3 part B. Bronx, New York Botanical Garden.

Clausen, J., D. D. Keck, and W. M. Hiesey. 1939. The concept of species based on experiment. American Journal of Botany 26:103-106.

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

Darwin, C. 1859, On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. *In* D.M. Porter and P.W. Graham *eds.* The Portable Darwin, 1993, Penguin Books.

Dobzhansky, T. 1951, Genetics and the Origin of Species. New York, Columbia University Press.

Fisher, R. A. 1958, The Genetical Theory of Natural Selection, Dover.

Isely, D. 1998, Native and Naturalized Leguminosae (Fabaceae) of the United Stated. Provo, Utah, Monte L. Bean Life Science Museum.

Jones, M. E. 1895. Contributions to Western Botany. No. 7. Proceedings of the California Academy of Sciences 5:611-732.

—. 1923, Revision of North American *Astragalus*. Salt Lake City, Utah.

Mayden, R. L. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem, Pages 381-422 *in* M. F. Claridge, H. A. Dawah, and M. R. Wilson, eds. Species: The Units of Biodiversity, Chapman & Hall.

Mayr, E., and P. D. Ashlock. 1991, Principles of Systematic Zoology, McGraw-Hill.

Rieseberg, L. H., T. E. Wood, and E. J. Baack. 2006. The nature of plant species. Nature 440:524-527.

Rydberg, P. A. 1929a. Astragalanae. North American Flora 24:251-462.

—. 1929b. Scylla and Charybdis. Proceedings of the International congress of plant sciences, Ithica, New York 2:1539-1551.

Stearn, W. T. 1957. Introduction Species Plantarum: a Facsimile of the First Edition, 1753. London, The Ray Society.

Stebbins, G. L. 1993. Concepts of Species and Genera *in* Flora of North America Editorial Committee, ed. Flora of North America. New York and Oxford, Oxford University Press.

USDA NRCS. 2006. The PLANTS database, National Plant Data Center, Baton Rogue, LA 70874-4490 USA.

Wallace, A. R. 1858. On the tendency of varieties to depart indefinitely from the original type *in* D. M. Porter, and P. W. Graham, eds. The Portable Darwin, 1993, Penguin Books.

Welsh, S. L. 2007, North American Species of *Astragalus* Linnaeus (Leguminosae). Prove, UT, Monte L. Bean Life Science Museum.

Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proceedings of the Sixth International Congress of Genetics 1:356-366.

—. 1978, Evolution and the genetics of populations. Volume 4: variability within and among natural populations. Chicago, University of Chicago Press.

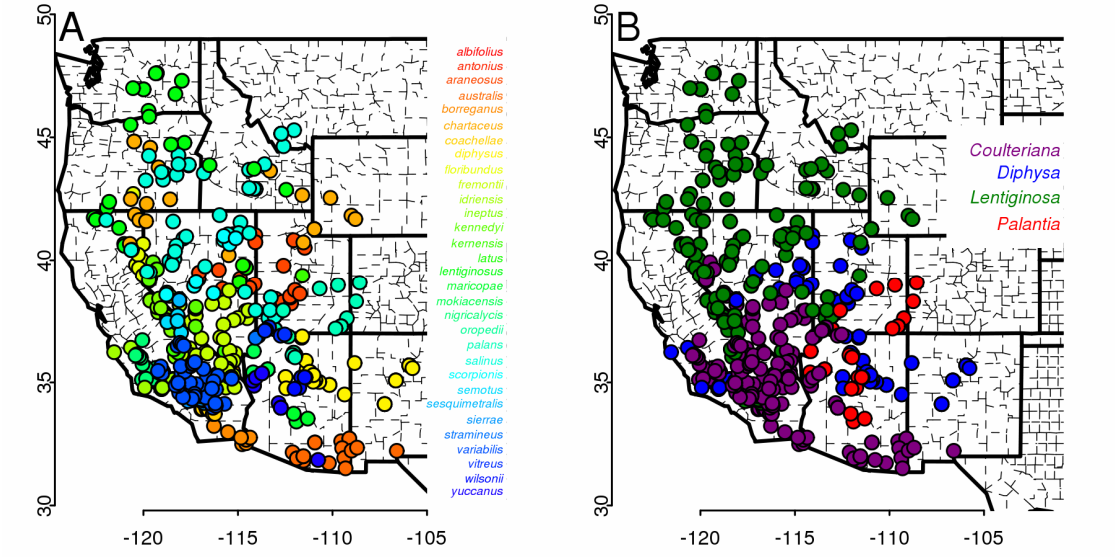


Figure 1.1. Distribution map of *Astragalus lentiginosus* based on specimens reported by Barneby (1945). Panel A is labeled by variety with taxonomy updated by Knaus (Table 1.2; Appendix A). Panel B is labeled by groups proposed by Rydberg 1929a.

Table 1.1. Taxonomic treatments of *Astragalus lentiginosus*.

Author	Year	Taxa
Jones, M.E.	1923	18
Rydberg, P.A	1929	36
Barneby, R.C.	1945	40
Barneby, R.C.	1964	36
Isely, D.	1998	40
USDA NRCS	2006	35
Welsh, S.L.	2007	42

Table 1.2. Summary of morphometric, CpSSR and AFLP samples. Barneby (1964) reports his personal collections (before the slash) as well as all specimens he viewed (after the slash). Entries for the current dataset include number of populations (before the slash) and total number of individuals (after the slash).

Variety	Rydbergian Group	Distribution	Barneby's sample	Morpho-metrics	CpSSR	AFLP
A. l. var. *ambiguus*	*Coulteriana*	AZ	1/4	-	1/4	-
A. l. var. *australis*	*Coulteriana*	AZ	5/31	-	1/4	-
A. l. var. *borreganus*	*Coulteriana*	CA, AZ; SON	4/31	19	2/8	2/30
A. l. var. *coachellae*	*Coulteriana*	CA	2/34	20	3/10	3/50
A. l. var. *fremontii*	*Coulteriana*	AZ, CA, NV, UT	14/101	21	8/31	6/60
A. l. var. *kennedyi*	*Coulteriana*	NV	3/31	19	2/8	2/20
A. l. var. *micans*	*Coulteriana*	CA	0/5	-	1/4	1/10
A. l. var. *nigricalycis*	*Coulteriana*	CA	2/56	-	1/4	1/9
A. l. var. *stramineus*	*Coulteriana*	NV	1/9	-	1/4	-
A. l. var. *variabilis*	*Coulteriana*	CA, NV	9/114	21	8/28	7/100
A. l. var. *vitreus*	*Coulteriana*	AZ, UT	2/16	-	1/4	-
A. l. var. *yuccanus*	*Coulteriana*	AZ	3/16	-	1/4	-
A. l. var. *araneosus*	*Diphysa*	NV, UT	8/39	20	3/12	-
A. l. var. *chartaceus*	*Diphysa*	NV, OR	8/56	10	2/8	-
A. l. var. *diphysus*	*Diphysa*	AZ, NM	10/66	16	-	-
A. l. var. *higginsii*	*Diphysa*	NM, TX	-	-	-	-
A. l. var. *idriensis*	*Diphysa*	CA	1/24	-	1/4	-
A. l. var. *latus*	*Diphysa*	NV	1/5	-	-	-
A. l. var. *multiracemosus*	*Diphysa*	NV	-	-	-	-
A. l. var. *negundo*	*Diphysa*	UT	-	-	-	-
A. l. var. *oropedii*	*Diphysa*	AZ	0/7	-	-	-
A. l. var. *piscinensis*	*Diphysa*	CA	-	-	1/4	-
A. l. var. *pohlii*	*Diphysa*	UT	-	-	-	-
A. l. var. *sesquimetralis*	*Diphysa*	CA, NV	0/1	-	2/7	-
A. l. var. *albifolius*	*Lentiginosa*	CA	3/16	-	1/4	-
A. l. var. *antonius*	*Lentiginosa*	CA	0/8	-	1/4	-
A. l. var. *floribundus*	*Lentiginosa*	CA, NV, OR	2/26	14	2/7	-
A. l. var. *ineptus*	*Lentiginosa*	CA	2/23	21	1/4	-
A. l. var. *kernensis*	*Lentiginosa*	CA, NV	0/8	-	1/4	-
A. l. var. *lentiginosus*	*Lentiginosa*	CA, NV, OR, WA, ID; B.C.	3/73	13	4/15	-
A. l. var. *salinus*	*Lentiginosa*	CA, ID, NV, OR, UT, WY	12/102	20	10/39	-
A. l. var. *scorpionis*	*Lentiginosa*	NV, UT	1/26	10	2/8	-
A. l. var. *semotus*	*Lentiginosa*	CA, NV	0/14	-	2/7	-
A. l. var. *sierrae*	*Lentiginosa*	CA	2/18	-	1/4	-
A. l. var. *maricopae*	*Palantia*	AZ	0/5	-	1/4	-
A. l. var. *mokiacensis*	*Palantia*	AZ, NV, UT	0/5	-	2/8	-
A. l. var. *palans*	*Palantia*	AZ, CO, UT	9/40	20	1/4	-
A. l. var. *trumbullensis*	*Palantia*	AZ	-	-	1/4	-
A. l. var. *ursinus*	*Palantia*	UT	0/1	-	1/4	-
A. l. var. *wilsonii*	*Palantia*	AZ	2/16	-	1/4	-
40 varieties			110/1027	244	71/272	22/279

2 Morphometric characterization of *Astragalus lentiginosus*.

ABSTRACT

The study of infra-taxa has historically been considered the study of incipient species. *Astragalus lentiginosus* Douglas ex Hooker (Fabaceae) is the most taxonomically complex species in the U.S. flora. The dramatic amount of morphological diversity contained within *A. lentiginosus* is reflected by its taxonomic history, where many taxa that are currently recognized as varieties were originally described as species. Morphometric data presented here indicate that the varieties lack clear regions of distinction, which is congruent with their circumscription as infra-taxa. *K*-means clustering was employed to determine the number of groups but failed to result in an optimal number of groups, suggesting that the varieties are clinal and can be divided into an arbitrary number of infra-taxa. Existing infra-specific circumscription is surprisingly similar to this statistical optimization. Significant correlations to climatic parameters suggest that the great diversity within *A. lentiginosus* may be due to local adaptation. The bewildering amount of diversity contained within the species *Astragalus lentiginosus* begs for decomposition, yet its clinal nature precludes it from division into discrete groups. However, divisions within this species should not be interpreted as discrete, nor should they necessarily be considered exclusive. As a species *A. lentiginosus* exists as an array of populations in a delicate balance between the cohesiveness for which it has been described as a species, and local adaptation.

Keywords: *Astragalus lentiginosus*; clines; ecotype; Fabaceae; Great Basin; Mojave Desert; morphometrics; speciation.

2.1 INTRODUCTION

The ‘species’ is considered the fundamental unit of biology (Stebbins 1950; Mayr and Ashlock 1991; Raven and Johnson 2002; Coyne and Orr 2004). The PLANTS database (USDA NRCS 2006) of United States vascular plants includes 33,383 species. It also includes 3,853 taxa at infraspecific ranks (USDA NRCS 2006), indicating that around 11% of the taxa in the U.S. flora are infra-taxa (taxa recognized at the ranks of subspecies or variety). If the species is the fundamental unit of biology, then of what value are infra-taxa and why do we have so many of them?

A unifying theme among species concepts is that the species is somehow ‘discrete’ (Mayden 1997; Coyne and Orr 2004), even though the metric is debatable (e.g., significant morphological distinctiveness, reproductive isolation, reciprocal monophyly, etc.). For example, the Biological Species Concept (Mayr and Ashlock 1991) indicates that a ‘species’ is a group of entities that are reproductively isolated from other ‘species.’ This is philosophically attractive because it implies that these entities (the ‘species’) no longer share a common evolutionary path due to an inability to share genetic material.

An important ‘unit’ of evolution may not necessarily require reproductive isolation. Theory indicates that adaptive divergence can occur in spite of gene flow (Wu 2001; Via 2002). This indicates that ‘groups’ of organisms can diverge to occupy different adaptive peaks even when reproductive barriers are incomplete or nonexistent. However, as long as there is a potential for the transfer of genetic material there is the possibility of intermediates. These intermediates may be present as poorly adapted individuals or individuals which are adapted to selective forces which are intermediate to the ends of the spectrum. A species can be seen as an array of populations which possess characters that confer cohesiveness in the face of local adaptation which works towards divergence.

If a species is given as discrete then how does one delineate infra-taxa, which therefore must be somehow non-discrete? Here I employ the most taxon-rich species in the U.S. flora (Table 2.1), *Astragalus lentiginosus* Douglas ex Hooker (Fabaceae), to explore the value and circumscription of infra-taxa.

The subspecies and variety as synonyms— The International Code of Botanical Nomenclature (Greuter et al. 2000) provides for the hierarchy of: species, subspecies, variety and form, as well as recommendations for other infraspecific ranks (e.g., subvariety). The zoological codes provide only for the rank of subspecies (Haig et al. 2006; Mallet 2007). Here I consider the ranks of 'subspecies' and 'variety' to be synonymous based on their current application (McDade 1995; Mallet 2007) and their legal interpretation in the United States (USFWS 1978). I use the term 'infraspecies' to refer to taxa circumscribed at the rank of either subspecies or variety.

A brief history of infra-taxa— Linnaeus is credited with providing the modern system of binominal nomenclature. He also employed the trinomial at the rank of 'variety' (Linnaeus 1753). Linnaeus considered the species to be the product of creation while the variety resulted from variation that had arisen since creation (Stearn 1957). This discreteness is perhaps inherited from the classical Greek concept of 'essence,' where species possess a unique 'essence' which helps define them (Mallet 2007). Modern nomenclature has adapted an evolutionary system; however, these systems share the concept that infraspecies are recently derived.

One path to speciation described by Darwin (1859) is the increase in variation to a point where the magnitude of variation is no longer maintainable, resulting in divergence that ends in distinct species. He subsequently relied on the multitude of artificial selection experiments performed informally by breeders to demonstrate this increase in variation, and how this variation can accumulate in a relatively short amount of time. These ideas were somewhat formalized by Fisher (1958), who described the idea of 'steady states' and their maintenance. These 'steady states'

could be maintained until the amount of variation contained within the system exceeded some critical point, whereupon this ‘steady state’ ruptured into new species.

Huxley (1938, 1939) addressed the study of large amounts of relatively continuous variation and sought to classify it. He tried to formalize the problem of clines by proposing the idea of discontinuous (stepped) clines (clines in which groups possess a shallower slope than the entire group), and continuous clines. This required the a priori determination of groups. Restriction of a ‘group’ to a population is unsatisfactory, as taxonomists are often interested in relationships among groups. This division of a number of geographic populations into a ‘group’ becomes a slippery slope into the contentious debate amongst ‘splitters’ and ‘lumpers’ (Mallet 2007).

The work of Clausen, Keck and Hiesey (1940) pioneered the reciprocal transplant experiment and shed new light on the nature of species. Through the growth of clones at different elevations they demonstrated dramatic local adaptation within species. These differences manifested themselves in a manner that was easily observed in the phenotype. Their work even encouraged a non-Linnaean terminology to describe this infraspecific variation (Clausen, Keck and Hiesey 1939).

Wilson and Brown (1953) criticized the subspecific rank. They argued that the naming of these infraspecific groups implied a discrete nature and that this detracted attention from the species, the rank which they felt should be the focus of biologists. They asserted that this implication is misleading because it is the species that is supposedly discrete. This leaves the subspecies as a group of entities whose divisions appear arbitrary and therefore may have little value.

These criticisms were addressed by Mayr (1982), who agreed that the infraspecific rank confused the importance of the species (which should be the focus of biology). However, he defended infraspecies (which he referred to as polytopic races) as an important record of infraspecific variation. He concluded that infraspecies represent

important evolutionary groups that may be incipient species and serve an important role in the documentation of instances of allopatric speciation (Mayr 1982; Mayr and Ashlock 1991).

A contemporary view of infra-taxa— Infra-taxa have been considered incipient species (Mayr and Ashlock 1991). Another important role that the infraspecific rank fulfills is the retention of names that were once considered species. As more data has become available, discontinuities between species have sometimes disappeared. These specific epithets are retained at the infraspecific level as a record of within species diversity (Jones 1923; Barneby 1964, 1989), as opposed to reduction to synonymy. Current theoretical discussion of the speciation process includes topics such as the coalescence (Hudson 1991; Nordborg 2001; Hudson and Coyne 2002; Felsenstein 2004) or adaptive divergence (Wu 2001; Via 2002; Dieckmann et al. 2004). The discrete nature of molecular genetic data (e.g., A, T, G or C) appears to promise a discrete answer, however these authors present theoretical rationale for the existence of genetic intermediates (e.g., the coalescence). It appears that these infraspecific ranks are excellent opportunities for empirical tests of these theories. However, recognition of taxa at the infraspecific ranks remains contentious (Zink 2004; Haig et al. 2006).

The most taxonomically complex species in the U.S. flora— *Astragalus lentiginosus* Douglas ex Hooker (Fabaceae) contains more infra-taxa than any other species in the U.S. flora (Table 2.1). The species is distributed throughout the arid regions of western North America (Figure 2.1), where it frequently occupies disturbed, saline, or otherwise marginal habitats. Many of the varieties were originally described as species (Hooker 1833; Gray 1856, 1863; Sheldon 1894). As collections increased intermediate forms became apparent, this led to the reduction of these species to varieties (Jones 1895, 1923). Per Axel Rydberg employed a very different species concept (Rydberg 1929b), elevating the varieties of *A. lentiginosus* to species in the genera *Cystium* Steven (inflated pods) and *Tium* Medikus (slightly inflated pods)

(Rydberg 1929a), as well as describing new species himself. Within the genus *Cystium* he included the groups *Lentiginosa,Coulteriana* and *Diphysa* (Rydberg 1929a) which were separated based on inflorescence length, flower size and flower color. Similarly, within the genus *Tium* he included the group *Palantia* (Rydberg 1929a). This grouping is no longer formally recognized with names but is reflected in the modern keys to the group (Barneby 1945, 1964, 1989; Spellenberg 1993; Isely 1998; Welsh et al. 2003; Welsh 2007). Barneby (1945) returned the group to a single species with numerous varieties. Through time several varieties have been reduced to synonymy (Barneby 1964, 1989) while new varieties have also been described (Barneby 1977; Welsh 1981; Welsh and Barneby 1981; Welsh and Atwood 2001). Alexander (2005, 2007) proposed taxonomic revisions within *A. lentiginosus* and related taxa. His morphological and molecular analyses resulted in no significant findings and are therefore not addressed here. The membership of *A. lentiginosus* has waxed and waned through time and taxonomic opinion. Jones (1923) initiated our current concept of a species with many varieties by recognizing 18 varieties. However, as many as 42 varieties have been recognized at once (Welsh 2007).

Several systems of taxonomy for *A. lentiginosus* have recently been proposed (Barneby 1964, 1989; Isely 1998; USDA NRCS 2006; Welsh 2007), all of which present differences in recognized taxa. It is therefore important to identify a system of taxonomy prior to investigations within this group. Here I recognize a system of 40 varieties (Table 2.2). Philosophically this system largely follows the comprehensive treatment of Barneby (1964) with subsequent modifications (Barneby 1989). Unfortunately, both of these treatments are currently out of date. Welsh (2007) contributed the most current treatment, but included taxonomic differences to the works of Barneby (1964, 1989). Here I recognize a hybrid system with an attempted adherence to the philosophy of Barneby (1964, 1989) with the updates of Welsh (2007). This system consists of the taxa recognized by Welsh (2007) with exceptions. The variety *chartaceus* is recognized with the variety *platyphyllidius* considered a synonym (Barneby 1989). The variety *macrolobus* is considered a synonym to the

variety *salinus* (Barneby 1989). The variety *toyabensis* is considered a synonym to the variety *scorpionis* (Barneby 1989). The variety *wahweapensis* is considered a synonym to the variety *diphysus* (Barneby 1989). These varieties have been organized into the four groups proposed by Rydberg (see above), however because taxa have been described since Rydberg's work some of these organizations reflect the present author's judgement.

A number of studies have investigated genetic relationships within the New World members of *Astragalus* (Karron 1989; Liston 1990; Liston 1992a, b; Wojciechowski et al. 1993; Travis et al. 1996; Wojciechowski et al. 1999; Alexander et al. 2004; Allphin et al. 2005; Knaus et al. 2005). Studies employing the internally transcribed spacer region (ITS) of ribosomal DNA alone (Wojciechowski et al. 1993) and ITS concatenated with the chloroplast *trn*L intron (Wojciechowski et al. 1999) demonstrated low levels of molecular diversity among the New World, aneuploid (Spellenberg 1976; Wojciechowski et al. 1999), members of *Astragalus*, with *A. lentiginosus* appearing nearly identical to *A. purshii* (Wojciechowski et al., 1993; Wojciechowski et al., 1999). One interpretation of this data is that the aneuploid members of *Astragalus* have seen a rapid radiation into the New World, leading to numerous morphological species (419 species; USDA NRCS 2006) with relatively little molecular divergence. Knaus et al. (2005) utilized amplified fragment length polymorphisms within *A. lentiginosus* to discern infraspecific molecular variation. They concluded that there appeared to be greater significance to varietal groupings than to geographic groupings, suggesting a molecular basis to the varieties.

Here I choose to focus on the morphometrics of a varietal complex as an important facet in determining what a species is. The phenotype has many obvious relations to the genotype (Falconer and Mackay 1996; Waitt and Levin 1997; Walsh 2001) and is of great relevance to the delineation of species (Rieseberg et al. 2006). The vast majority of plant taxa have been described based on the Linnaean (morphological) species concept (Mayden, 1997), based on its ease of application and relatively long

history. Morphometric studies have become prevalent within the plant kingdom (Rieseberg et al. 2006), as well as at the specific rank within the Fabaceae (Small et al. 1984; Small and Brookes 1990; Chandler and Crisp 1998; Sheidai et al. 2001; Conti 2007; Kropf 2008). These studies present varying degrees of resolution among species and infra-taxa. One potential reason for varying success among species and infraspecific discernment may be that many researchers report significant findings based on differences in means as opposed to diagnosability (Patten and Unitt 2002), resulting in taxa which lack regions of discontinuity (i.e., have overlapping 95% confidence intervals) which could facilitate the assignment to exclusive categories. The quantification of the morphological aspects of a taxon of evolutionary interest is therefore a logical first step in gaining inference into processes that may be active within the group.

2.2 METHODS

Morphometric measurements— Specimens of *Astragalus lentiginosus* from major western herbaria (JEPS, NESH, NY, ORE, OSC, POM, RENO, RM, RSA, and UC) were measured. A goal of 20 specimens per variety, having both fruit and flower, was attained for several varieties of *A. lentiginosus* (Table 2.3). Many endemic varieties occur only at a few localities (e.g., varieties *albifolius*, *sesquimetralis* and *piscinensis*) which resulted in poor representation in collections, these taxa were therefore omitted. Some varieties, which occur distantly from the population center of southern California, were similarly poorly represented by herbarium collections and were also omitted. This opportunistic sample represents the well collected varieties of *A. lentiginosus*, which largely contains the widespread or common varieties.

Fourteen linear characters were chosen from the keys of Barneby (1945, 1964 1989) and measured with a ruler, electronic caliper, or ocular micrometer (Table 2.3). Whenever possible, three measurements were made of each structure and the arithmetic mean of these values was recorded. Measurements were made from different parts of the plant (i.e., different stems or racemes) whenever possible. If

multiple plants were present on a single sheet measurements were taken from as many plants as possible. This was an attempt to capture the maximal amount of variation contained within a plant or specimen. The fourteen measured characters were: stem internode length, leaf rachis length, leaf petiole length, leaflet number, leaflet width, leaflet length, peduncle length, floral axis in fruit, keel length, calyx tooth length, calyx tube length, pod length, pod height, pod valve thickness and beak length. All data were examined for univariate normality and heteroscedacity through histograms and scatterplots using the generic functions 'hist' and 'plot' (R Development Core Team 2007). The characters floral axis in fruit and pod valve thickness were natural log transformed to improve normality (Table 2.3).

Assessment of infraspecific structure— Discriminant function analysis was performed using the function 'lda' in the R package 'MASS' (R Development Core Team 2007) to explore structure given the a priori grouping as varieties both with and without the use of latitude and longitude as additional explanatory variables (Figure 2.2). Discriminant function analysis seeks to build multivariate functions that best discriminate among a priori groups (Everitt 2005; Tabachnick and Fidell 2007). These functions can then be plotted in ordination space. All characters were standardized by standard deviations in order to equalize the magnitude of each character.

In order to assess the optimal number of groups, *K*-means clustering was performed on the data using the function 'kmeans' in the R package 'stats' (R Development Core Team 2007). *K*-means analysis uses a predefined number of groups (but not group membership) and utilizes an optimality criterion to fit the data within these groups (Everitt 2005; Tabachnick and Fidell 2007). A sum of squares can then be calculated to assess the fit. Note that as the number of groups increase the sum of squares is expected to decrease; therefore researchers usually examine plots for a breakpoint in the data where additional groups no longer appear to dramatically decrease the sum of squares. Standardization by standard deviation was performed to equalize the contribution of each trait. In order to explore the sensitivity of the data to the

algorithm, several algorithms were employed (Hartigan-Wong, Lloyd, Forgy and MacQueen; R Development Core Team 2007).

Phenological standardization— In order to explore climatic trends in morphology the PRISM dataset (PRISM Group, Oregon State University, http://www.prismclimate.org) was used. Specimens were assigned a latitude and longitude by referencing specimen label information to a place-name database (topozone.com) or converted from township and range when this data was available (www.esg.montana.edu/gl/trs-data.html). The spatial join command in ESRI's ArcView Spatial Analyst (Redlands, California) was used to extract elevation, monthly minimum and maximum temperature and monthly precipitation from the PRISM dataset.

In order to assign phenologically meaningful data to the monthly PRISM dataset, a sine wave was fit to each specimen's annual set of monthly temperature data using:

$$y = A \sin (x - \omega t)$$

where y is degrees Celsius, A is [max(temp)-min(temp)]/2 and scales the amplitude of the wave to the maximum and minimum temperatures recorded for a site (in radians the unscaled maximum and minimum amplitudes are 1 and -1), x is days from January 1, ω is angular frequency and here is set to 1, and t indicates the initial phase where $\sin(\pi/2) = 1$ radian and here is transformed to 91.25 days (see conversion below). The initial phase was set by averaging the two greatest temperatures and subtracting 91.25 to determine the day of greatest temperature. Trigonometric functions were performed in radians and were converted to days (instead of 360 degrees) with radians = days*π/182.5. The fitting of sine waves was performed using custom scripts executed in the R statistical programming language (R Development Core Team, 2007).

Calculation of two climatic parameters was of particular interest. Dormant season precipitation was calculated by plotting sine waves of minimum temperature for each individual. The dormant season was then defined as days between the first frost-free

day and after the last frost-free day. Precipitation data was only available in a monthly format, and it did not appear appropriate to fit sine waves to these data. The daily data was used to determine frost-free months and the cumulative precipitation from these months was calculated. Days above 10˚ Celsius were calculated by plotting sine waves of mean daily temperature and summing the days that were above 10˚ Celsius.

Simple linear regression was performed using the R function 'lm' (R Development Core Team 2007). Climatic parameters used as independent variables and morphological characters as well as principle components were used as dependent variables to explore the relative importance of morphology with respect to climate. Principle components analysis was performed using the function 'princomp' (R Development Core Team 2007) on a matrix of correlations to explore patterns of structure in the group. Principle components analysis is an eigen analysis used to explore data without the a priori assignment of groups (Everitt 2005; Tabachnick and Fidell 2007). A matrix of correlations was chosen to give each character equal weighting in the analysis.

Distribution of keel lengths— In order to explore the distribution of a representative character, a plot of normal distributions was made. Because of its taxonomic importance (Jones 1923; Rydberg 1929a; Barneby 1945, 1964, 1989) and potential importance in pollinator success (Green and Bohart 1975), keel length was chosen as a representative character. A mean and standard deviation was calculated for each variety, which were then used to plot normal distributions for each variety's keel length.

2.3 RESULTS

Analysis of the varieties of *Astragalus lentiginosus* resulted in a lack of discontinuity. Varieties and Rydbergian groups occupied contiguous but non-discrete regions of morphospace. A search for an optimal number of failed to result in a clearly optimal value. Existing taxonomic circumscription was very similar to the statistical

optimization of group membership. Regression of multivariate morphological characters and climatic parameters resulted in significant correlations.

Naming of infra-taxa, as well as the provision of infraspecific keys, implies that these groups have a discrete nature (Wilson and Brown 1953). These names and keys also imply a degree of diagnosability. Names and keys have both been applied to the varieties of *Astragalus lentiginosus* (Jones 1923; Rydberg 1929a; Barneby 1945, 1964, 1989). In order to explore the implications of these categories discriminant function analysis (Figure 2.2) was performed on taxonomically important characters.

Discriminant function analysis (Figure 2.2a) demonstrates cohesiveness to Rydbergian groups, and the varieties contained within these groups. However, an absence of regions of discontinuity among Rydbergian groups or varieties demonstrates these categories to be somewhat arbitrary. These results suggest a clinal nature among the Rydbergian groups and varieties. The inclusion of latitude and longitude as explanatory factors results in the initiation of clustering among Rydbergian groups (Figure 2.2b); however these clusters still lack regions of discontinuity to define them. This analysis demonstrates that a lack of morphometric clustering can be improved by the knowledge of geographic position, indicating that the nature of the Rydbergian groups and varieties of *A. lentiginosus* is largely geographic.

Optimal number of groups— In order to assess the optimal number of groups within this species, and thus the number of infraspecific taxa, *K*-means clustering was performed (Figure 2.3). There does not appear to be a 'natural' break point in the data. This indicates that no clear optimal number of groups exists within the exceptional amount of diversity contained within this species. The groups of Rydberg (1929b) and the varieties of Barneby (1964) are included for comparison. These two taxonomic groupings are surprisingly similar to the statistical optimization presented here. This suggests that while statistical optimizations may result in slightly 'better' groupings, the existing system is remarkable good.

Climatic correlations— In order to explore climatic correlations, principle components were compared to climatic parameters. Principle component one is loaded with vegetative characters (Table 2.4) and is negatively correlated with dormant season precipitation (Figure 2.4a; $r^2 = 38.38\%$, $p < 2.2 * 10^{-16}$). Principle component two is loaded with floral characters (Table 2.4) and is highly correlated with growing season temperature (Figure 2.4b. $r^2 = 20.9\%$, $p = 1.082 * 10^{-10}$). These statistics imply a significant and dramatic proportion of the variation in multivariate diversity is attributable to climatic factors. These results are particularly interesting in light of the error associated with the estimate of geographic position and interpolation of climatic parameters to these positions.

2.4 DISCUSSION

A species can be seen as an array of populations which exhibit both cohesive and divergent characters. Cohesive processes, such as gene flow and the retention of ancestral traits, help define higher levels of grouping, such as the variety or Rydberg's groups (1929a). Divergent processes, such as genetic drift and local adaptation, lend distinctiveness to lower levels of organization, such as the population. Cohesive processes active within *Astragalus lentiginosus* provides diagnosability to the species. Divergent processes have led to a large degree of heterogeneity within the species which has been recognized through several perspectives on taxonomic circumscription (Jones 1923; Rydberg 1929a; Barneby 1964).

The importance of the morpho-species— I have chosen to focus on the morpho-species for largely practical reasons, primarily due to its ease of application with available herbarium specimens. Evolutionary biologists frequently refer to the Biological Species Concept (Mayr and Ashlock 1991; Coyne and Orr 2004), which defines the species as a group of organisms that are reproductively isolated from other species. This is philosophically attractive because it delimits the 'species' as groups which can no longer share evolutionary fates due to their inability to share genetic

material. Of the 33,000 species included in the PLANTS database (USDA NRCS 2006), it seems doubtful that a large percentage of these taxa have been subjected to the relatively laborious tests of inter-fertility. Morphologically circumscribed species are often confirmed as biological species and play an important role in building evidence prior to tests of reproductive incompatibility (Mayr and Ashlock 1991). The primary utility of morphology is that it is how we perceive organisms; it is therefore an important first step in characterizing a species.

Theory predicts that divergence can progress despite gene flow (Wu 2001; Via 2002; Dieckmann et al. 2004). Given the multiple paths to speciation that have been proposed (Stebbins 1950; Dobzhansky 1951; Coyne and Orr 2004), it seems reasonable to accept that the nature of a species may be multi-faceted, and therefore require several methodological approaches to ascertain, particularly in the case of multi-taxon complexes. Selection acts on the phenotype, and includes processes that may lead to quantitative divergence which exceeds the genetic mark of reproductive isolation (Spitze 1993; Podolsky and Holtsford 1995; Merilä and Crnokrak 2001; Wilding et al. 2001; Wu 2001; Storz 2002; Via 2002; Gomez-Mestre and Tejedo 2004; Beaumont 2005). Morphometric measurements of herbarium samples may be considered a measure of the phenotype which is confounded by an unknown environmental contribution. Phenotypic measurements are considered to result in weaker correlations to selective pressures compared to measurements where the environmental contribution to the phenotype controlled (Waitt and Levin 1997). Therefore, these measurements may represent an important first step in exploring quantitative differences (Meagher et al. 1978; Storfer 1996), particularly for taxa that may be difficult to propagate or have other practical limitations.

Infraspecific structure— Naming of infraspecific taxa, and the provision of keys, implies that infra-taxa have a discrete nature (Wilson and Brown 1953). This is in conflict with the idea of the species as the fundamental unit of biology, where the 'species' is considered discrete. Here I've demonstrated that the varietal complex *A.*

lentiginosus does not contain 'discrete' varieties. Instead, these taxa occupy cohesive regions of ordination space but lack clear or 'natural' breaks. The varieties of *A. lentiginosus* fall along a cline of morphometric diversity. However, this is a complex cline which must be viewed from distribution-wide perspective. Due to patchiness in the distribution of *A. lentiginosus* all of the intermediates along the cline may not occur geographically proximal. Instead a large portion of the range of *A. lentiginosus* must be surveyed in order to capture all the intermediates. I feel this is in agreement with the hierarchical system of nomenclature where the infra-taxon (e.g., subspecies or variety) is subordinate to the species and the species is discrete. This is consistent with the presented data where the varieties of *A. lentiginosus* are contiguous in morphospace but lack discontinuity.

Optimal number of groups— The sampled varieties of *A. lentiginosus* do not fall into an optimal number of groups. This is consistent with discriminant function analysis and *K*-means clustering failing to find discrete groups within *A. lentiginosus*. Instead of falling into easily classified groups, the varieties of *A. lentiginosus* fall into regions along a continuum. The morphometric diversity within the species begs decomposition, as its taxonomic history reflects. Current circumscription separates populations with white versus purple flowers, glabrous versus tomentose vestiture and prostrate versus erect habits as different varieties. Despite this great morphological diversity, the species *A. lentiginosus* appears relatively cohesive (Figure 2.2).

Comparison of Rydberg's four groups (1929a) and Barneby's 14 varieties (1964) (Figure 2.3) demonstrates a large coherence between these taxonomic systems and the presented statistical analysis. While statistical methods do demonstrate an improvement upon existing taxonomy, it is unclear how dramatic this improvement is. The delimitations of Rydberg (1929a) and Barneby (1964) represent close approximations to the presented statistical analysis (Figure 2.3). A major concern of any statistical optimization is sensitivity to methods. Different methods may result in slightly but not dramatically different groups, similar to how the present optimization

does not dramatically improve upon existing circumscription. In the interest of stability in the taxonomic system it seems that any 'improvement' on the system needs to present dramatically novel results and justify its superiority to those methods that have already been implemented.

Much methodological interest has been recently focused on the identification of optimal group number within a dataset as well as the membership of these groups (Pritchard et al. 2000; Falush et al. 2003; Guillot et al. 2005; Guillot et al. 2005; Corander and Marttinen 2006). This interest has arisen not only in the field of population genetics, but is also shared by those working in fields such as community ecology (Austin 1985). Many statistical texts include a decision tree to help researchers choose appropriate statistical tests (Dytham 1999; Tabachnick and Fidell 2007) however, these trees frequently start with the question of whether there are a priori groups (which leads to ANOVA type analyses) or whether a relationship is sought (which leads to regression methods). This circumvents the question of whether groups should be delimited or not. Determination of appropriate group number and membership is has a complex and sometimes controversial problem that thus far has no simple solution.

Climatic correlations— Morphologies of *A. lentiginosus* have significant correlation to climatic parameters (Figure 2.4), which suggests but does not confirm causation. Climatic change during the beginning of the Holocene has been implicated in changing the distributions of plants (Thompson and Mead 1982; Grayson 1993; Davis and Shaw 2001). The distribution of *A. lentiginosus* includes inland sand dunes and desert seeps, as well as regions such as the Lahontan Basin (the site of Pleistocene Lake Lahontan). These habitats have changed dramatically since the last glacial maximum and these changes have undoubtedly played a role in the evolution of *A. lentiginosus*. Here the role of climate as a potential selective force that may be responsible for the diversity in morphology currently expressed in *A. lentiginosus* has been explored and results indicate that localized specialization is prevalent and

significant (Figure 2.4). Meta-analyses suggest that these correlations may be improved upon if the environmental component of variance is controlled (Waitt and Levin 1997).

'Auxiliary' systems of taxonomy— A vocabulary to describe clinal relationships among infra-taxa exists outside the terminology of the International Code of Botanical Nomenclature (Greuter et al. 2000). This includes the 'rassenkreis' (Endler 1977) (this term is attributed to Rensch; see Mallet 2007), clines (Huxley 1938, 1939) and ecotypes (Clausen, Keck and Hiesey, 1939). This vocabulary has not become prevalent in the taxonomic literature, and hasn't been mentioned in the literature of *A. lentiginosus*. This may be due to several reasons. Naming implies homogeneity within a group and distinctiveness among groups; this perhaps is reflected in the Linnaean background where the 'species' is a product of creation. Typification seems to reinforce this idea by suggesting that one specimen is representative of an entire taxon. Conversely, many authors insist that infraspecific diversity is the rule (Mayr and Ashlock 1991) and these claims have been supported by a wealth of empirical studies (Spitze 1993; Podolsky and Holtsford 1995; Storz 2002; Roff and Mousseau 2005; St Clair et al. 2005; Lankau and Strauss 2007). This is, at least in part, an issue of perspective. Unfamiliar organisms may appear uniform while taxa with which we are more familiar may appear to host a wealth of diversity. The more we study something the more detail we discover, and as a species is more closely investigated a wealth of infraspecific diversity will undoubtedly be discovered.

As an example of the morphological diversity within *A. lentiginosus*, I've presented keel length data (Figure 2.5). Keel length is not only important taxonomically (Jones 1923; Barneby 1945, 1964, 1989), but may also be tied to pollinator success (Green and Bohart 1975), suggesting this character may be important for natural selection. The keel lengths of the sampled varieties of *A. lentiginosus* range from 5.8 to 15 mm, giving values that range by a factor of almost three. It seems inappropriate to categorize characters spanning this magnitude of difference. In addition, the

possibility that this is a character tied to reproductive success suggests a barrier to gene flow. Yet the clinal nature of this trait does not rule out the possibility gene flow through a stepping stone process (Rieseberg and Burke, 2001).

Accounting of the morphological diversity contained within *A. lentiginosus* is a complex problem. While these 'auxiliary' systems of taxonomy have not been thoroughly embraced by the taxonomic community, they seem to fulfill an appropriate role in not only naming but describing the nature of infra-taxon diversity. I agree with Wilson and Brown's argument that the practice of naming groups implies that these groups are homogenous groups and such naming creates confusion. I disagree with them in that I do feel that infraspecific variation is important to recognize (Mayr 1982). After all, if biologists are interested in studying the process of speciation they must study within species diversity. If a record exists for this diversity, through nomenclature, it facilitates the transfer of information and has thus fulfilled one of the most important roles of nomenclature. Yet I am also faced with the troubling reality that the use of these 'auxiliary' systems of taxonomy, which seem very appropriate in the case of *A. lentiginosus*, have not seen wide acceptance. This is perhaps indicative of the unique nature of *A. lentiginosus*. It appears obvious that *A. lentiginosus* represents another example where the term 'cline' needs to be applied and emphasized alongside the infra-taxon names that have already been applied.

Complex characters— The current dataset consists of relatively simple linear measurements. Many studies have employed similar measurements and these linear characters are frequently employed in floristic keys due to their ease of use. This has failed to capture some of the complexities within *A. lentiginosus*. Characters such as flower color and degree of hairiness have all been important taxonomic characters. But in order to include an unbiased measure of these traits, complex methodology or apparatus is frequently necessary (Ehleringer 1981; Sandquist and Ehleringer 1997; Ackerly et al. 2000; Reig et al. 2001). For example, reflectance could capture floral color or degree of hairiness while 3-D modeling could be employed to capture the

shape of the banner. These complex methodologies frequently lead the researcher to focus attention on a single structure due to the increased effort involved with this sort of data capture. Here I've employed a more holistic perspective of the plant with the hope that later research can use this to justify a focus on these complex, and possibly diagnostic, characters.

Conclusion— The 'species' is considered the fundamental unit of biology; however, infra-taxa are prevalent in the U.S. flora (Table 2.1). *Astragalus lentiginosus* (Fabaceae) is presented here as the most taxonomically complex species in the U.S. flora (Table 2.1) and an exemplar for the study of infra-taxa. A brief history of infra-taxa has been presented to address why we have infra-taxa. Theory in speciation has also been explored to understand processes relevant to the study of infra-taxa as incipient species. Finally, the nomenclature of clines has been introduced as a manner in which to describe these biologically important patterns.

Several lines of evidence have been presented to support infraspecific nomenclature in *A. lentiginosus*. First, precedence exists for the naming of parts of clines (Huxley 1938; Clausen, Keck and Hiesey 1939; Huxley 1939; Clausen, Keck and Hiesey 1940). Second, a distribution spanning over one thousand kilometers suggests that gene flow is not occurring throughout this species' range; at a minimum the ends of the spectrum may be in the process of allopatric speciation. Third, climatic correlations suggest that local adaptation may be important. Finally, infra-specific nomenclature facilitates one of the most basic purposes of nomenclature; it assigns names to allow conversation of these morphologically diverse forms. It remains to be elucidated whether the apparent cohesiveness among the varieties of *A. lentiginosus* is due to shared ancestral polymorphism or a stepping stone process of gene flow throughout its range, for a few important characters. Demonstration of intrinsic barriers to reproduction would strengthen these claims. Future research should take into account the clinal nature of characters reported here, and should expect similar patterns in reproductive incompatibility.

A species is an array of populations that are evolving collectively. This collectiveness can be facilitated through some level of gene flow, shared retention of ancestral traits or similar responses to selective pressures. This collectiveness must occur despite the pressures of local adaptation and drift, or else speciation occurs. I present the idea that an infraspecific complex is an array of populations where the processes that promote collective evolution are in a tenuous balance with more localized processes. The lack of discontinuity among the varieties of *A. lentiginosus* is demonstrative of the collective processes that have historically held this group together, leading to their taxonomic description as a single species. Yet the great amount of diversity within *A. lentiginosus* coupled with significant climatic correlations suggest the group is beginning to diverge due to these local processes. The species *A. lentiginosus* is perhaps a multifaceted example of a plastic mass being stretched across the western United States; it may contract leading to homogenization, or it may begin to fracture, eventually leading to new species.

2.5 LITERATURE CITED

Ackerly, D. D., S. A. Dudley, S. E. Sultan, J. Schmitt, J. S. Coleman, C. R. Linder, D. R. Sandquist et al. 2000. The evolution of plant ecophysiological traits: recent advances and future directions. BioScience 50:979-995.

Alexander, J. A. 2005. The taxonomic status of *Astragalus mokiacensis* (Fabaceae), a species endemic to the southwestern United States. Brittonia 57:320-333.

—. 2007. A Taxonomic Revision of *Astragalus mokiacensis* and Allied Taxa within the *Astragalus lentiginosus* Complex of Section *Diphysi*. Ph.D. thesis, Oregon State University.

Alexander, J. A., A. Liston, and S. J. Popovich. 2004. Genetic diversity of the narrow endemic *Astragalus oniciformis* (Fabaceae). American Journal of Botany 91:2004-2012.

Allphin, L., N. Brian, and T. Matheson. 2005. Reproductive success and genetic divergence among varieties of the rare and endangered *Astragalus cremnophylax* (Fabaceae) from Arizona, USA. Conservation Genetics 6:803-821.

Austin, M. P. 1985. Continuum concept: ordination methods, and niche theory. Annual Review of Ecology and Systematics 16:39-61.

Barneby, R. C. 1945. *Pugillus Astragalorum* IV: the section *Diplocystium*. Leaflets of Western Botany 4:65-147.

—. 1964. Atlas of North American *Astragalus*. Memoirs of the New York Botanical Garden 13:1-1188.
—. 1977. *Dragma hippomanicum* III: *novitates Californicae*. Brittonia 29: 376-381.
—. 1989, Fabales: Intermountain Flora, v. 3 part B. Bronx, New York Botanical Garden.
Beaumont, M. A. 2005. Adaptation and speciation: what can F_{ST} tell us? Trends in Ecology & Evolution 20:435-440.
Chandler, G. T., and M. D. Crisp. 1998. Morphometric and phylogenetic analisis of the *Daviesia ulicifolia* complex (Fabaceae, Mirbelieae). Plant Systematics and Evolution 209:93-122.
Clausen, J., D. D. Keck, and W. M. Hiesey. 1939. The concept of species based on experiment. American Journal of Botany 26:103-106.
Clausen, J., D. D. Keck, and W. M. Hiesey. 1940, Experimental Studies on the Nature of Species. Washington D.C., Carnegie Institution of Washington.
Conti, F. 2007. Morphometric study and taxonomy of *Genista pulchella* Vis. *s.l.* (Fabaceae), a south European species. Botanical Journal of the Linneaen Society 153:245-254.
Corander, J., and P. Marttinen. 2006. Bayesian identification of admixture events using multilocus molecular markers. Molecular Ecology 15:2833-2843.
Coyne, J. A., and H. A. Orr. 2004, Speciation. Sunderland, Mass., Sinauer Associates.
Darwin, C. 1859, On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. *In* D.M. Porter and P.W. Graham *eds.* The Portable Darwin, 1993, Penguin Books.
Davis, M. B., and R. G. Shaw. 2001. Range shifts and adaptive responses to Quaternary climate change. Science 292:673-679.
Dieckmann, U., M. Doebeli, J. A. J. Metz, and D. Tautz. 2004, Adaptive speciation: Cambridge studies in adaptive dynamics ; 3. Cambridge, UK ; New York, Cambridge University Press.
Dytham, C. 1999, Choosing and Using Statistics: a Biologist's Guide, Blackwell Science.
Ehleringer, J. R. 1981. Leaf absorptances of Mojave and Sonoran Desert plants. Oecologia 49: 366-370.
Endler, J. A. 1977, Geographic Variation, Speciation, and Clines. Princeton, New Jersey, Princeton University Press.
Everitt, B. 2005, An R and S-plus companion to multivariate analysis. London, Springer.
Falconer, D. S., and T. F. C. Mackay. 1996, Introduction to Quantitative Genetics, Pearson.
Falush, D., M. Stephens, and J. K. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567-1587.
Felsenstein, J. 2004, Inferring Phyogenies, Sinauer.
Fisher, R. A. 1958, The Genetical Theory of Natural Selection, Dover.

Gomez-Mestre, I., and M. Tejedo. 2004. Contrasting patterns of quantitative and neutral genetic variation in locally adapted populations of the natterjack toad, *Bufo calamita*. Evolution 58:2343-2352.
Gray, A. 1856. *Astragalus fremontii in* J. Torrey, ed. Pacific Rail Road Reports, Vol. 4., Report on the Botany of the Expedition. Washington, D.C.
—. 1863. Revision and arrangement (mainly by the fruit) of the North American species of *Astragalus* and *Oxytropis*. Proceedings of the American Academy of Arts and Sciences 6:188-237.
Grayson, D. K. 1993, The desert's past: a natural prehistory of the Great Basin. Washington, Smithsonian Institution Press.
Green, T. W., and G. E. Bohart. 1975. The pollination ecology of *Astragalus utahensis* (Leguminosae). American Journal of Botany 62:379-386.
Greuter, W., J. McNeill, F. R. Barrie, H.-M. Burdet, V. Demoulin, T. S. Filgueiras, D. H. Nicolson et al. 2000, International Code of Botanical Nomenclature (St. Louis Code). Konigstein, Koeltz Scientific Books.
Guillot, G., A. Estoup, F. Mortier, and J. F. Cosson. 2005. A spatial statistical model for landscape genetics. Genetics 170:1261-1280.
Guillot, G., F. Mortier, and A. Estoup. 2005. Geneland: a program for landscape genetics. Molecular Ecology Notes 5:712-715.
Haig, S. M., E. Beever, S. M. Chambers, H. M. Draheim, B. D. Dugger, S. Dunham, E. Elliot-Smith et al. 2006. Taxonomic considerations in listing subspecies under the U.S. endangered species act. Conservation Biology 20: 1584-1594.
Hooker, W. J. 1833, Flora Boreali-Americana, v. 1. London, Treuttel & Wurtz.
Hudson, R. R. 1991. Gene geneologies and the coalescent process, Pages 1-44 *in* D. J. Futuyma, and J. Antonovics, eds. Oxford Surveys in Evolutionary Biology. Oxford, UK, Oxford UNiversity Press.
Hudson, R. R., and J. A. Coyne. 2002. Mathematical consequences of the genealogical species concept. Evolution 56:1557-1565.
Huxley, J. 1938. Clines: an auxiliary taxonomic principle. Nature 142:219-220.
—. 1939. Clines: an auxiliary method in taxonomy. Bijdragen tot de Dierkunde (Contributions to Zoology) 27:491-520.
Isely, D. 1998, Native and Naturalized Leguminosae (Fabaceae) of the United Stated. Provo, Utah, Monte L. Bean Life Science Museum.
Jones, M. E. 1895. Contributions to Western Botany. No. 7. Proceedings of the California Academy of Sciences 5:611-732.
—. 1923, Revision of North American *Astragalus*. Salt Lake City, Utah.
Karron, J. D. 1989. Breeding systems and levels of inbreeding depression in geographically restricted and widespread species of *Astragalus* (Fabaceae). American Journal of Botany 76:331-340.
Knaus, B. J., R. C. Cronn, and A. Liston. 2005. Genetic characterization of three varieties of *Astragalus lentiginosus* (Fabaceae). Brittonia 57:334-344.
Kropf, M. 2008. Intraspecific patterns of European mountain plants: a morphometric analysis confirms molecular results in the submediterranean oreophyte *Anthyllis montana* L. (Fabaceae). Taxon 57:511-524.

Lankau, R. A., and S. Y. Strauss. 2007. Mutual feedbacks maintain both genetic and species diversity in a plant community. Science 317:1561-1563.
Linnaeus, C. 1753, Species Plantarum. Facsimile of the first edition with introduction by W.T. Stearn. 1957. London, The Ray Society.
Liston, A. 1990. An evolutionary study of *Astragalus* sect. *Leptocarpi* subsect. *californici* (Fabaceae). Ph.D. Thesis thesis, Claremont Graduate School, Claremont.
Mallet, J. 2007. Subspecies, semispecies, superspecies, Pages 523-526 *in* S. Levin, and et al., eds. Encyclopedia of Biodiversity, Academic Press.
Mayden, R. L. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem, Pages 381-422 *in* M. F. Claridge, H. A. Dawah, and M. R. Wilson, eds. Species: The Units of Biodiversity, Chapman & Hall.
Mayr, E. 1982. Of what use are subspecies? Auk 99:593-595.
Mayr, E., and P. D. Ashlock. 1991, Principles of Systematic Zoology, McGraw-Hill.
McDade, L. A. 1995. Species concepts and problems in practice: insights from botanical monographs. Systematic Botany 20:606-622.
Meagher, T. R., J. Antonovics, and R. Primack. 1978. Experimental ecological genetics in *Plantago*. III. Genetic variation and demography in relation to survival of *Plantago cordata*, a rare species. Biological conservation 14:243-257.
Merilä, J., and P. Crnokrak. 2001. Comparison of genetic differentiation at marker loci and quantitative traits. Journal of Evolutionary Biology 14:892-903.
Nordborg, M. 2001. Coalescent theory, Pages 179-212 *in* D. J. Balding, M. J. Bishop, and C. Canings, eds. Handbook of Statistical Genetics. Chichester, UK, John Wiley and Sons.
Patten, M. A., and P. Unitt. 2002. Diagnodability versus mean differences of Sage Sparrow subspecies. The Auk 119:26-35.
Podolsky, R. H., and T. P. Holtsford. 1995. Population structure of morphological traits in *Clarkia dudleyana* I. Comparison of F_{ST} between allozymes and morphological traits. Genetics 140:733-744.
Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.
R Development Core Team. 2007.R: a language and environment for statistical computing.R foundation for statistical computing, Vienna, Austria.
Raven, P. H., and G. B. Johnson. 2002, Biology, McGraw-Hill.
Reig, S., M. J. Daniels, and D. W. Macdonald. 2001. Craniometric differentiation within wild-living cats in Scotland using 3D morphometrics. Journal of Zoology 253:121-132.
Rieseberg, L. H., and J. M. Burke. 2001. The biological reality of species: Gene flow, selection, and collective evolution. Taxon 50:47-67.
Rieseberg, L. H., T. E. Wood, and E. J. Baack. 2006. The nature of plant species. Nature 440:524-527.
Roff, D. A., and T. Mousseau. 2005. The evolution of the phenotypic covariance matrix: evidence for selection and drift in *Melanoplus*. Journal of Evolutionary Biology 18:1104-1114.

Rydberg, P.A. 1929a. *Astragalanae*. North American Flora 24:251-462.
—. 1929b. Scylla and Charybdis. Proceedings of the International Congress of Plant Sciences, Ithica, New York 2:1539-1551.
Sandquist, D. R., and J. R. Ehleringer. 1997. Intraspecific variation in leaf pubescence and drought response in *Encelia farinosa* associated with contrasting desert environments. New Phytologist 135:635-644.
Sheidai, M., Z. Yazdanbakhsh, M. Assadi, and M. Moussavi. 2001. Cytology and morphometric study of *Alhagi* (Leguminosae) species in Iran. Nordic Journal of Botany 21:83-91.
Sheldon, E. P. 1894. A preliminary list of the North American species of *Astragalus*. Minnesota Botanical Studies 1:116-175.
Small, E., and B. Brookes. 1990. A numerical taxonomic analysis of the *Medicago littoralis - M. truncatula* complex. Canadian Journal of Botany 68:1667-1674.
Small, E., W. F. Grant, and C. W. Crompton. 1984. A taxonomic study of the *Lotus corniculatus* complex in Turkey. Canadian Journal of Botany 62:1044-1053.
Spellenberg, R. 1976. Chromosome numbers and their cytotaxonomic significance for North American *Astragalus* (Fabaceae). Taxon 25:463-476.
—. 1993. *Astragalus*, Pages 583-605 *in* J. C. Hickman, ed. The Jepson Manual: Higher Plants of California. Berkeley, CA, University of California Press.
Spitze, K. 1993. Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. Genetics 135:367-374.
St Clair, J. B., N. L. Mandel, and K. W. Vance-Borland. 2005. Genecology of douglas fir in western Oregon and Washington. Annals of Botany 96:119-1214.
Stearn, W. T. 1957. Introduction *Species Plantarum*: a Facsimile of the First Edition, 1753. London, The Ray Society.
Stebbins, G. L. J. 1950, Variation and Evolution in Plants. New York, Columbia University Press.
Storfer, A. 1996. Quantitative genetics: a promising approach for the assessment of genetic variation in endangered species. Trends in Ecology and Evolution 11:343-348.
Storz, J. F. 2002. Contrasting patterns of divergence in quantitative traits and neutral DNA markers: analysis of clinal variation. Molecular Ecology 11:2537-2551.
Tabachnick, B. G., and L. S. Fidell. 2007, Using Multivariate Statistics, Pearson.
Thompson, R. S., and J. I. Mead. 1982. Late quaternary environments and biogeography in the great basin. Quaternary Research 17:39-55.
Travis, S. E., J. Maschinski, and P. Keim. 1996. An analysis of genetic variation in *Astragalus cremnophylax* var. *cremnophylax*, a critically-endangered plant, using AFLP markers. Molecular Ecology 5:735-745.
USDA NRCS. 2006. The PLANTS database, National Plant Data Center, Baton Rogue, LA 70874-4490 USA.
USFWS. 1978. Determination that 11 plant taxa are endangered and 2 plant taxa are threatened species. Federal Register 43:17910-17916.
Via, S. 2002. The ecological genetics of speciation. The American Naturalist 159:S1-S7.

Waitt, D. E., and D. A. Levin. 1997. Genetic and phenotypic correlations in plants: a botanical test of Cheverud's conjecture. Heredity 80:310-319.
Walsh, B. 2001. Quantitative genetics in the age of genomics. Theoretical Population Biology 59:175-184.
Welsh, S. L. 1981. New taxa of western plants - in tribute. Brittonia 33:294-303.
—. 2007, North American Species of *Astragalus* Linnaeus (Leguminosae). Prove, UT, Monte L. Bean Life Science Museum.
Welsh, S. L., and N. D. Atwood. 2001. New taxa and nomenclatural proposals in miscellaneous families - Utah and Arizona. Rhodora 103:71-95.
Welsh, S. L., N. D. Atwood, S. Goodrich, and L. C. Higgins. 2003, A Utah Flora. Provo, Brigham Young University.
Welsh, S. L., and R. C. Barneby. 1981. *Astragalus lentiginosus* (Fabaceae) revisited - a unique new variety. Iselya 2:1-2.
Wilding, C. S., R. K. Butlin, and J. Grahame. 2001. Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. Journal of Evolutionary Biology 14:611-619.
Wilson, E. O., and W. L. Brown. 1953. The subspecies concept and its taxonomic application. Systematic Zoology 2:97-111.
Wojciechowski, M. F., M. J. Sanderson, B. G. Baldwin, and M. J. Donoghue. 1993. Monophyly of aneuploid *Astragalus* (Fabaceae): evidence from nuclear ribosomal DNA internal transcribed spacer sequences. American Journal of Botany 80:711-722.
Wojciechowski, M. F., M. J. Sanderson, and J.-M. Hu. 1999. Evidence on the monophyly of *Astragalus* (Fabaceae) and its major subgroups based on nuclear ribosomal DNA ITS and chloroplast DNA *trnL* intron data. Systematic Botany 24:409-437.
Wu, C.-I. 2001. The genic view of the process of speciation. Journal of Evolutionary Biology 14:851-865.
Zink, R. M. 2004. The role of subspecies in obscuring avian biological diversity and misleading conservation policy. Proceedings of the Royal Society of London B 271:561-564.

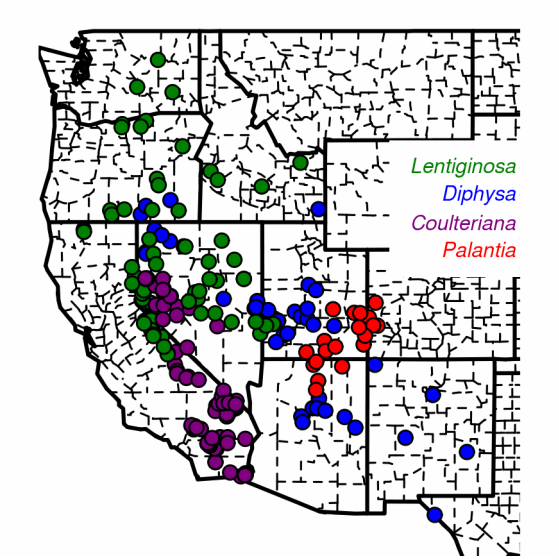


Figure 2.1. Map of the morphometric sample. Color represents Rydbergian groups (see text).

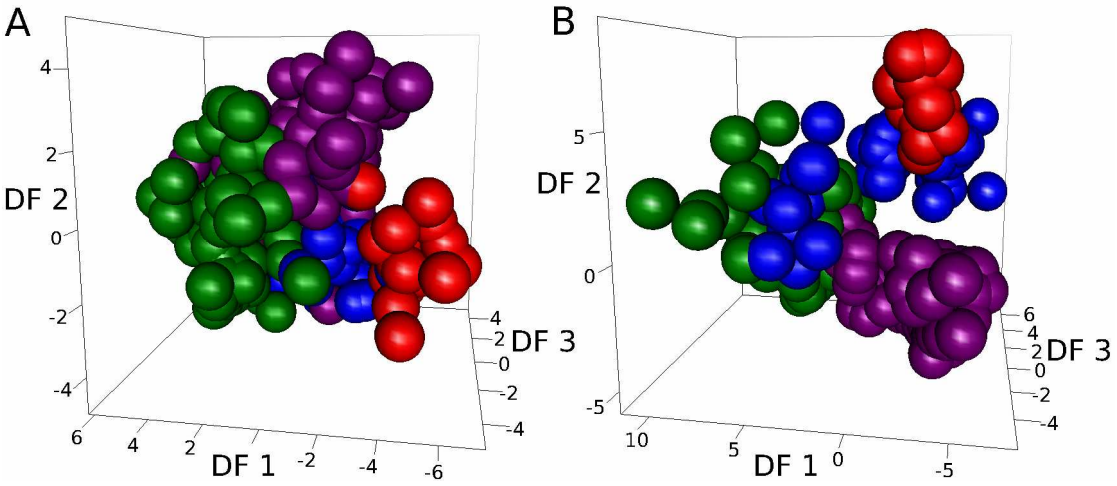


Figure 2.2. Plot of discriminant functions. Colors follow Figure 2.1. Pane A represents an analysis of only morphometric characters. Pane B represents an analysis of morphometric characters as well as latitude and longitude.

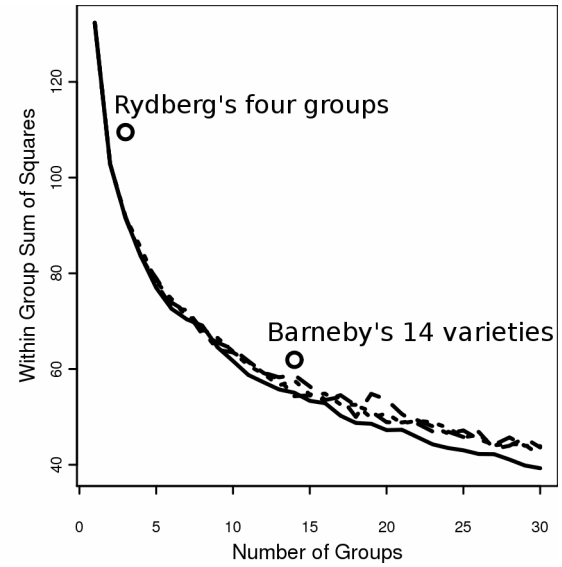


Figure 2.3. *K*-means clustering diagnostics. Lines represent the optimality criteria Hartigan-Wong, Lloyd, Forgy, and MacQueen. Sum of squares for Rydberg's four groups and Barneby's 14 varieties are provided for comparison.

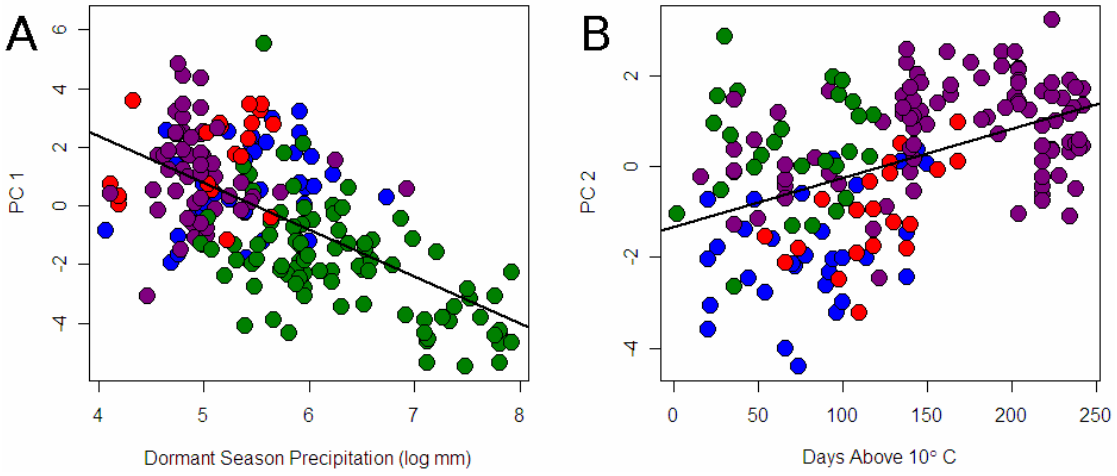


Figure 2.4. Correlations of principle components to climatic parameters. Pane A is PC1 as a function of dormant season precipitation. Pane B is PC2 as a function of days above 10° C. Color follows Figure 2.1. PC1 is correlated to dormant season precipitation with $r^2 = 38\%$ ($F_{1, 196} = 123.7$, $p < 2.2 * 10^{-16}$). PC2 is correlated to dormant season precipitation with $r^2 = 21\%$ ($F_{1, 174} = 47.23$, $p = 1.082 * 10^{-10}$).

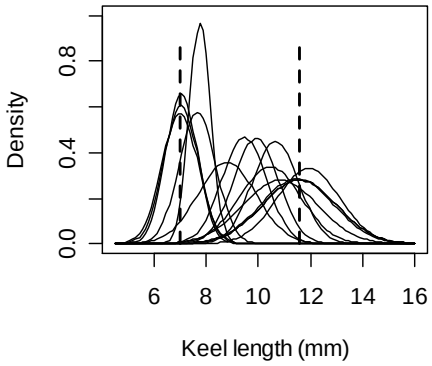


Figure 2.5. Normal distributions of keel lengths. Distributions were plotted from the mean and variance for each variety sampled (Table 2.2). Broken lines mark the mean keel length of *A. l.* vars. *floribundus* and *araneosus*. Using *A. l.* var. *araneosus* as a reference (μ = 11.6 mm, σ = 1.4 mm) the keel lengths of *A. lentiginosus* span just over three standard deviations to *A. l.* var. *floribundus* (μ = 7.0 mm, σ = 0.7 mm).

Table 2.1. Species of North American plants with 10 or more infra-taxa (USDA NRCS 2006).

Family	Scientific Name	Infra-rank	count
Fabaceae	***Astragalus lentiginosus***	**var.**	**35**
Polygonaceae	*Eriogonum umbellatum*	var.	30
Asteraceae	*Ericameria nauseosa*	ssp. & var.	22
Asteraceae	*Hymenopappus filifolius*	var.	13
Malvaceae	*Sidalcea malviflora*	ssp.	13
Polygonaceae	*Eriogonum nudum*	var.	13
Asteraceae	*Ericameria parryi*	var.	12
Asteraceae	*Eriophyllum lanatum*	var.	12
Brassicaceae	*Lepidium montanum*	var.	12
Asteraceae	*Achillea millefolium*	var.	11
Caryophyllaceae	*Arenaria congesta*	var.	11
Fabaceae	*Trifolium longipes*	ssp.	11
Rosaceae	*Potentilla glandulosa*	ssp.	11
Asteraceae	*Machaeranthera canescens*	ssp. & var.	10
Brassicaceae	*Descurainia pinnata*	ssp.	10
Fabaceae	*Oxytropis campestris*	var.	10
Onagraceae	*Camissonia claviformis*	ssp.	10
Polygonaceae	*Eriogonum heermannii*	var.	10
Polygonaceae	*Eriogonum ovalifolium*	var.	10

Note: There are 33,383 species in the PLANTS database, 1,330 subspecies, and 2,523 varieties.

Table 2.2. The varieties of *Astragalus lentiginosus*. Sampled taxa are in bold.

Variety	Section	Distribution	Barneby's Sample	Barneby's Specimens	Current Sample
A.l. var. *ambiguus*	*Coulteriana*	endemic	4	1	NA
A.l. var. *australis*	*Coulteriana*	widespread	31	5	NA
A.l.* var. *borreganus	***Coulteriana***	**widespread**	**31**	**4**	**19**
A.l.* var. *coachellae	***Coulteriana***	**endemic**	**34**	**2**	**20**
A.l.* var. *fremontii	***Coulteriana***	**widespread**	**101**	**14**	**21**
A.l.* var. *kennedyi	***Coulteriana***	**widespread**	**31**	**3**	**19**
A.l. var. *micans*	*Coulteriana*	endemic	5	0	NA
A.l. var. *nigricalycis*	*Coulteriana*	widespread	56	2	NA
A.l. var. *stramineus*	*Coulteriana*	endemic	9	1	NA
A.l.* var. *variabilis	***Coulteriana***	**widespread**	**114**	**9**	**21**
A.l. var. *vitreus*	*Coulteriana*	endemic	16	2	NA
A.l. var. *yuccanus*	*Coulteriana*	endemic	16	3	NA
A.l.* var. *araneosus	***Diphysa***	**widespread**	**39**	**8**	**20**
A.l.* var. *chartaceus	***Diphysa***	**widespread**	**56**	**8**	**10**
A.l.* var. *diphysus	***Diphysa***	**widespread**	**66**	**10**	**16**
A.l. var. *higginsii*	*Diphysa*	endemic	NA	NA	NA
A.l. var. *idriensis*	*Diphysa*	endemic	24	1	NA
A.l. var. *latus*	*Diphysa*	endemic	5	1	NA
A.l. var. *multiracemosus*	*Diphysa*	endemic	NA	NA	NA
A.l. var. *negundo*	*Diphysa*	endemic	NA	NA	NA
A.l. var. *oropedii*	*Diphysa*	endemic	7	0	NA
A.l. var. *piscinensis*	*Diphysa*	endemic	NA	NA	NA
A.l. var. *pohlii*	*Diphysa*	endemic	NA	NA	NA
A.l. var. *sesquimetralis*	*Diphysa*	endemic	1	0	NA
A.l. var. *albifolius*	*Lentiginosa*	endemic	16	3	NA
A.l. var. *antonius*	*Lentiginosa*	endemic	8	0	NA
A.l.* var. *floribundus	***Lentiginosa***	**widespread**	**26**	**2**	**14**
A.l.* var. *ineptus	***Lentiginosa***	**widespread**	**23**	**2**	**21**
A.l. var. *kernensis*	*Lentiginosa*	endemic	8	0	NA
A.l.* var. *lentiginosus	***Lentiginosa***	**widespread**	**73**	**3**	**13**
A.l.* var. *salinus	***Lentiginosa***	**widespread**	**102**	**12**	**20**
A.l.* var. *scorpionis	***Lentiginosa***	**endemic**	**26**	**1**	**10**
A.l. var. *semotus*	*Lentiginosa*	endemic	14	0	NA
A.l. var. *sierrae*	*Lentiginosa*	endemic	18	2	NA
A.l. var. *maricopae*	*Palantia*	endemic	5	0	NA
A.l. var. *mokiacensis*	*Palantia*	widespread	5	0	NA
A.l.* var. *palans	***Palantia***	**widespread**	**40**	**9**	**20**
A.l. var. *trumbullensis*	*Palantia*	endemic	NA	NA	NA
A.l. var. *ursinus*	*Palantia*	endemic	1	0	NA
A.l. var. *wilsonii*	*Palantia*	endemic	16	2	NA
40 varieties			1027	110	244

Note: Barneby (1964) reports the number of specimens (Barneby's samples) viewed during preparation of his monograph as well as the number which were his own collections (Barneby's specimens).

Table 2.3. Precision and transformation of characters.

Character	Units	Transformation
Floral		
peduncle length	0.5 mm	-
Fl axis in fruit	0.5 mm	ln
keel length(mm)	0.5 mm	-
calyx tooth length	0.5 mm	-
calyx tube length	0.5 mm	-
Fruit		
pod length	0.5 mm	-
pod height	0.5 mm	-
pod valve thickness	0.01 mm	ln
beak length	0.5 mm	-
Vegetative		
Stem internode length	0.5 mm	-
Leaf rachis length	0.5 mm	-
Leaf petiole length	0.5 mm	-
Leaflet number	count	-
Leaflet width	0.5 mm	-
Leaflet length	0.5 mm	-

Table 2.4. Loadings for discriminant functions and principle components.

Character	**LD1**	**LD2**	**LD3**	**LD4**	**LD5**	**PC 1**	**PC 2**	**PC 3**	**PC 4**	**PC 5**
Peduncle length	-3.0628	2.3170	1.6900	-1.7522	-1.2688	0.3658	0.2069	0.0409	0.1855	0.0599
Floral axis in fruit	-0.4969	3.0506	0.2208	-2.9768	3.5591	0.2937	0.3250	0.0267	0.2569	-0.0601
Keel length	-4.8703	1.3598	2.3959	0.5514	-2.5785	0.2956	-0.3821	0.2145	0.0518	0.1036
Pod length	-4.3427	-1.6489	-2.9375	-3.8215	1.2224	0.2543	-0.2265	-0.3597	-0.0146	-0.2220
Pod height	3.5225	2.4580	6.5343	2.3027	-1.1104	0.0993	0.0835	-0.5256	-0.0238	0.0625
Pod valve thickness	-2.5271	2.7385	-0.4439	4.4077	-1.4871	0.0866	-0.1807	0.4458	-0.2919	-0.0139
Beak length	3.0125	-0.7117	2.6313	1.6680	1.5023	0.0663	-0.2751	-0.5270	-0.1199	-0.2473
Calyx tooth length	-1.5068	0.1470	-0.6577	3.4267	5.7861	0.2741	-0.3029	0.0415	-0.2239	-0.0528
Calyx tube length	-0.6633	-7.0489	2.2687	-2.2037	-0.5266	0.2577	-0.4654	0.0713	0.0142	0.1306
Stem internode length	1.1577	0.6324	0.3236	-0.8491	-2.1032	0.3047	0.2498	-0.0130	0.1416	-0.1239
Leaf rachis length	-2.4533	1.8196	0.6022	1.3225	-1.5102	0.3683	0.1681	0.0536	0.1309	0.1656
Petiole length	0.7016	-0.1877	-0.1590	0.0034	0.5311	0.0242	0.0828	-0.2333	-0.3237	0.8459
Leaflet number	0.9904	-1.5872	0.0869	0.5291	-0.3161	0.0850	-0.2800	-0.0470	0.6515	0.2710
Leaflet width	-0.9819	2.5898	-1.5847	1.7318	0.6733	0.3381	0.1692	0.0552	-0.3395	-0.0358
Leaflet length	2.8475	-2.2680	0.5134	0.6629	1.2639	0.3375	0.1614	-0.0020	-0.2540	-0.1211

3 Chloroplast simple sequence analysis of *Astragalus lentiginosus*

3.1 INTRODUCTION

The study of infra-specific taxa has been considered by many to be the study of incipient species (Darwin 1859; Dobzhansky 1951; Fisher 1958; Knaus Chapter 2). A species can be viewed as an array of populations, each diverging from the others due to local adaptation and neutral genetic drift. Gene flow mediated via migration (pollen or seed flow) provides a homogenizing force among the populations, and allows for selectively advantageous alleles to spread throughout the species (Slatkin 1976; Reiseberg and Burke 2001). Divergence via neutral genetic drift or adaptive divergence provides a heterogenizing force among populations. This 'push and pull' of local divergence moderated by gene-flow leads to collective divergence among entities we recognize as species.

The use of molecular markers in taxonomy shares similarities with the biological species concept. The biological species concept (Mayr and Ashlock 1991) defines a species as having a physiological barrier to reproduction, preventing gene flow, while molecular markers test for gene flow. The biological species concept is philosophically attractive because definite reproductive isolation represents independent evolutionary fates. The use of molecular markers therefore integrates well into existing taxonomic concepts and may provide important perspectives on taxonomic hypotheses.

Here I employ the concept of 'genetic identity' to refer to an abstraction of identical genotypes or haplotypes dependant on the marker system employed. The concept of genetic identity is a function of evolutionary divergence and also a function of our ability to discern it (i.e., the marker system employed). Molecular divergence is the result of extinction of shared ancestral genetic identities while mutation introduces new genetic identities. Because mutation in isolated (independent) populations is unlikely to result in genetic identities which are identical in state (based on a sufficient

number of markers), these new mutations are likely to be characteristic of new lineages. As evolutionary time increases among these lineages, genetic identity unique to each lineage is expected to increase, leading to increasingly divergent sampled populations. Rate of divergence is a function of the strength of genetic drift (which removes genetic identities) and mutation (which adds new genetic identities). Therefore, the observed divergence among populations is a function of the mutation rate, the size of the population and the time since actual divergence. This process has been referred to in the literature as lineage sorting (Avise 2004) or the coalescence (Rosenberg and Nordborg 2002; Felsenstein 2004).

I employ five chloroplast simple sequence repeats to infer levels of reproductive isolation within the most taxon rich species in the U.S. flora (Knaus Chapter 2). This marker system is used to test the hypotheses of panmixia at four levels of hierarchy: 1) global panmixia (distribution wide), 2) Rydbergian sections (regional panmixia), 3) varietal panmixia (varieties described by Barneby 1964) and 4) distinct populations. Global panmixia represents a null hypothesis evident as a pattern of random breeding throughout the range of the taxon of interest (*Astragalus lentiginosus*). Regional panmixia is defined as a pattern of significant genetic differentiation among Rydbergian sections (Rydberg 1929; Table 3.1) but haplotypic sharing among varieties. Varietal panmixia is defined as a pattern of minute differentiation among regions (Rydbergian sections) but significant divergence among varieties (Barneby 1964; Table 3.1) and sharing of diversity within varieties. Distinct populations are defined as a pattern where among-population differences account for the majority of the observed diversity and this diversity is not organized hierarchically within sections or varieties.

Astragalus (Fabaceae) is considered to be the largest genus of vascular plant (Mabberly 1987; Lock and Shrire 2005), consisting of approximately 2,500 species worldwide. The aneuploid New World species of *Astragalus* are considered to be a monophyletic group based on an aneuploid chromosome number (Spellenberg 1976)

as well as sequence data based on the nuclear ribosomal internal transcribed spacer (ITS; Wojciechowski et al. 1993) and concatenated data including the ITS and the *trnL* intron (Wojciechowski et al.1999). Scherson et al. (2008) has demonstrated some structure within the New World *Astragalus* based on ITS and two non-coding chloroplast regions. These studies demonstrate the aneuploid New World species of *Astragalus* to be monophyletic, however, resolution is lacking within the group.

Although the molecular phylogeny of the New World species of *Astragalus* remains to be resolved (Wojciechowski et al.1999; Wojciechowski et al.1999; Scherson et al. 2008), population level studies have resulted in significant inferences. Researchers have employed enzymes (Karron 1988; Liston 1992; Morris et al. 2002; Allphin et al. 2005), inter-simple sequence repeats (Alexander et al. 2004) and amplified fragment length polymorphisms (Travis et al. 1996; Knaus et al. 2005). Molecular generalities within *Astragalus* are lacking due a large range of among-population divergences having been reported.

The desert legume *Astragalus lentiginosus* Douglas ex Hooker (Fabaceae) contains more infra-taxa than any other species in the U.S. flora (Knaus Chapter 2). Utilizing amplified fragment length polymorphisms, Knaus et al. (2005) have suggested that the formation of varieties may be due to a significant level of reproductive isolation. Morphometric analysis (Knaus Chapter 2) has indicated clinality correlated with climate among the varieties, with a lack of regions of discontinuity to distinguish the varieties. Here I provide the most comprehensive exploration to date into the evolutionary relationships of *A. lentiginosus* utilizing a molecular marker system. Included in the sample are specimens representing almost all of the geographic range of *A. lentiginosus* (Figure 3.1) as well as most of its described taxa (Table 3.1).

3.2 METHODS

Sampling Strategy— A sampling strategy to represent the genetic constitution of *A. lentiginosus* throughout its range in the western United States as well as its taxonomic

diversity was employed. A goal of three to four samples per population was used to attain a sample containing 33 varieties, 71 populations, and 272 individuals (Table 3.1). For phenetic analysis samples from other species of *Astragalus* were employed. These taxa were *A. amphioxus, A. iodanthus, A. platytropis, A. pseudiodanthus, A. purshii* and *A. utahensis*. The species *A. iodanthus* and *A. pseudiodanthus* were placed in the same section as *A. lentiginosus* by Barneby (1964) and Welsh (2007). Woijciechowski et al. (1993) and Woijciechowski et al. (1999) found *A. purshii* and *A. utahensis* to be closely related to *A. lentiginosus* based on the nuclear ribosomal internal transcribed spacer (ITS), as well as ITS combined with the chloroplast *trnL* intron DNA sequence data. Rydberg (1929) has considered *A. platytropis* to be closely related to *A. lentiginosus*. The species *A. amphioxys* is from the closely related section *Argophyllii* (Barneby 1964). To visualize the geospatial extent of the sample, a map was created using the R package 'maps' (Becker and Wilks 1993).

Leaf tissue was collected by the author from 2003-2005. Tissue was preserved on silica gel, dry ice, or in an herbarium press until its return to Corvallis, OR. Samples of *A. l.* vars. *australis, borreganus* (Font's Wash), and *piscinensis* were propagated from seed in petri dishes on moist filter paper. When the cotyledons had fully emerged these individuals were homogenized and DNA was extracted. Desiccated leaf tissue was also contributed by J.A. Alexander (Utah Valley State University) for samples from eastern NV, UT and AZ. Extraction methods followed Knaus et al. (2005), fastDNA (Qbiogene), or the Puregene (Gentra Systems) methods.

Haplotyping— Amplification protocols followed Weising and Gardner (1999) with the exceptions that reactions performed at 10 μl and the concentration of *Taq* polymerase was lowered to 0.04U/ μl. Primers ccmp2, ccmp5, ccmp6 and ccmp10 were screened as producing amplicons and as polymorphic. The ccmp10 primers were found to produce an amplicon containing several indels in a fragment ranging from 200-273 bp. Because of this complexity, a pair-wise distance was chosen for subsequent analysis as it is perceived as being relatively agnostic to mutation models.

The remaining three loci produced amplicons consistent with their being mononucleotide runs. Preliminary sequencing efforts (Knaus, unpublished data) revealed a mononucleotide run in the *trnT/L* region. This region was amplified using the reverse primer of Taberlet et al. (1991; primer b) and the novel primer trnTL839f (5'-CTT TGT CCT GTA ATC TCA TTA TTC-3'). All primers (except primer b of Taberlet et al.1991) included 'PIGtailing' to minimize stutter (Brownstein et al. 1996). Loci were amplified individually with primers labeled with HEX, FAM, or NED fluorescent molecules. Samples were multiplexed for fragment analysis which occurred at the Oregon State University Core Labs facility. Fragment analysis was performed on ABI3100 genotypers with POP4 polymer and gsROX500 size standard. Fragments were scored using Genotyper 3.7 (Applied Biosystems).

Statistical Methods— Distribution maps and histograms of allele distribution and haplotypes were created in the statistical programming language R (R Development Core Team 2007). Geneland convergence diagnostics and barplot of probability of group membership were also created in R.

Neighbor-Joining Tree— In order to visualize relationships among the taxa and haplotypes, a neighbor-joining tree was created. Because the marker ccmp10 was inferred to be a compound indel, and the marker trnTL839f had an anomalously small allele observed for *A. amphioxus* (Figure 3.2 D), a simple pair-wise distance was employed. This distance scores differences based on identity and is relatively agnostic to the mutation model. The function 'dist.gene' in the R package 'ape' (Paradis 2006) was used to generate a distance matrix and functions in the R package 'ape' were used to plot the tree. For visualization the taxon *Astragalus amphioxys* was used to root the tree, however this tree should be interpreted as unrooted.

Hierarchical *F*-statistics— To explore the partitioning of molecular variance among hierarchical levels within *A. lentiginosus*, *F*-statistics were calculated. Hierarchical levels of among-section, among-varieties, among-populations and within populations

were explored using the R package 'hierfstat' (Goudet 2005; de Meeûs and Goudet 2007). Because the chloroplast is considered to be a haploid, non-recombining chromosome, the alleles at each marker were combined into haplotypes and univariate *F*-statistics were computed. The high polymorphism observed at microsatellite loci decreases the maximum possible *F*-statistic below its theoretical limit of one (Hedrick 2005). To correct for this we rescored the data into three datasets to compute a maximum *F*-statistic for each level of hierarchy (Criscione and Blouin 2007). For example, to correct at the level of variety each dataset was rescored so that no variety shared the same allele while the allelic frequency within each population was retained. This modified dataset was used to calculate a maximum *F*-statistic for that level of hierarchy. This process was repeated for each level of hierarchy.

Bayesian spatial clustering— In order to infer optimal group number within our sample and their geographic distribution the software Geneland was used (Guillot et al. 2005, Guillot et al. 2005). The authors describe the algorithm in the software to be identical to the popular software Structure (Pritchard et al. 2000, Falush et al. 2003, 2007) except that it allows for the use of an x and y coordinate in the model. Here latitude and longitude were used. The Geneland function 'PlotTessellation' was modified to gain greater control over plotting features including the plotting of state boundaries from the R package 'maps' (Becker and Wilks 1993).

The five loci were scored as a homozygous diploid organism with decimal latitude and longitude used as x and y coordinates. First the software was used to determine the optimal number of groups in the dataset. Markov Chain Monte Carlo (MCMC) settings were: Rate.max = 273 (the number of individuals in the dataset), delta.coord = 0, npopmin = 1, npopinit = 71 (the number of populations), npopmax = 100, nb.nuclei.max = 819 (three times the number of individuals), nit (number of iterations) = 1000000, thinning = 50, freq.model = "Dirichlet", varnpop = "T" and spatial = "T" (see Geneland documentation). This resulted in a MCMC simulation run for one million iterations. To avoid serial autocorrelation the chain was thinned every 50

iterations, resulting in 20,000 iterations saved for subsequent processing. Shorter runs were also run to explore the stability of the chain (data not presented). Histograms and line plots were used to explore convergence of the chain. Several runs of this simulation determined that the optimal number of groups was seven. Next the software was used to determine membership of these seven groups and to plot this data spatially. The MCMC was rerun with the same settings above with the following changes: npopmin = 6, npopinit = 7, npopmax = 7 and nit = 10000. This resulted in 200 iterations being saved for subsequent analysis. The chains were post-processed with the function 'PostProcessChain', the posterior probabilities of group membership were plotted with 'PlotTessellation,' and a summary of these plots was created with 'PosteriorMode.' Modifications to plots beyond the default output were performed in R (R Development Core Team 2007).

3.3 RESULTS

Summary statistics— Histograms of allele frequencies indicate that per marker polymorphism is relatively infrequent (Figure 3.2). Most loci are represented by a common allele and other infrequent alleles. The combination of these alleles into haplotypes seemed to distribute haplotype frequencies (Figure 3.2 F) more uniformly than allele frequencies (Figure 3.2 A-E) but still retained a pattern of a few common haplotypes and many infrequent haplotypes. A total of 57 haplotypes were observed in *Astragalus lentiginosus* with an additional eight observed in non-*A. lentiginosus* species of *Astragalus*. No haplotypes were shared among *A. lentiginosus* and non-*A. lentiginosus* species of *Astragalus* (Figure 3.2 F).

Neighbor-joining tree— A neighbor-joining tree was constructed using a pair-wise distance (Figure 3.3). A general lack of correspondence of taxonomy to haplotypes or sections was observed. Haplotypes were clearly shared among varieties and clustering did not correspond to varietal or sectional arrangements. Two clusters of non-*A. lentiginosus* species of *Astragalus* were observed, one containing *A. amphioxys*, *A. platytropis* and *A. purshii* while another contained *A. utahensis*, *A. iodanthus* and *A.*

pseudiodanthus. Regardless of how the tree was rooted, one of these clusters of non-*A. lentiginosus* taxa will nest within the rest of *A. lentiginosus*.

Hierarchical F-statistics— Hierarchical *F*-statistics indicated significant differentiation at the levels of section, variety and population (Table 3.2). However, the major components of variance occurred at the among-population (55%) and within population (25%) levels. Variance at taxonomic levels was relatively small, with 17% of the variation accounting for the among-variety component and 4% accounting for the among section component (Table 3.2).

Bayesian spatial clustering— Results of the Geneland analysis of optimal group number of *Astragalus lentiginosus* based on five CpSSRs indicated that the optimal number of groups was seven (Figure 3.4). This number does not correspond to the sections proposed by Rydberg (1929) or the varieties proposed by Barneby (1964).

3.4 DISCUSSION

Chloroplast simple sequence repeat (CpSSR) haplotyping resulted in 57 observed haplotypes among four sections, 33 varieties, 71 populations and 272 individuals. The observed abundance of genetic polymorphism relative to section or variety number suggests that ample diversity exists to distinguish among these levels of hierarchy if divergence has occurred. While significant differences occur at all levels of hierarchy, the majority of the variance is partitioned at the among-population level, indicating that regional and varietal groupings explain relatively small amounts of the total variance (Table 3.2). This indicates that the most important level of hierarchy within this system is the population, rather than the variety or section. Clustering based on the neighbor-joining method and a Bayesian clustering method did not support regional or varietal taxonomic hypotheses. Bayesian clustering did not support a hypothesis of isolation by distance (IBD) in the data set. Based on the rejection of taxonomic and IBD hypotheses, I propose a new hypothesis of incomplete coalescence.

Performance of Hierarchical Hypotheses— Hypotheses for genetic structure have been presented at several levels of hierarchy: global, Rydbergian section, variety and population. The hypothesis of varietal structure appears to have been rejected. This is perhaps best illustrated by the neighbor-joining tree (Figure 3.3) where the variety *salinus*, the best represented taxon in the dataset (Table 3.1), appears almost ubiquitously throughout the dendrogram. If the variety represented a good genetic group it would be expected that all populations of *A. l.* var. *salinus* would be more closely related to one another than to other varieties. This would also be evident by the varietal grouping accounting for a large proportion of variance in the dataset, and the varieties would form clusters (Figure 3.3; Figure 3.5) of geographically proximal groups (Figure 3.6). Instead we find that the largest component of variance is not the varietal component (Table 3.2), varieties do not cluster together (Figure 3.3), hierarchical Bayesian clustering results in fewer clusters than varieties (Figure 3.4), and these cluster do not correspond well to varieties (Figure 3.5). Furthermore, while the varieties appear to have a strong correspondence to geography (Knaus Chapter 2; Figure 3.1), genetic clustering does not consistently correspond to geography (Figure 3.6 B, D, E and G). It is therefore concluded that the hypothesis of varieties as good genetic groups based on the current sample, its measure (five CpSSRs) and the presented analyses is rejected.

The hypothesis of Rydbergian sections (Rydberg 1929; Table 3.1) also appears to have been rejected. The smallest variance component, although significant, was explained by sectional grouping (Table 3.2). Hierarchical Bayesian clustering resulted in seven clusters (Figure 3.4), almost twice as many as sections (Table 3.1). These clusters had a poor correspondence to sections (Figure 3.5). While the sections appear to have a strong correspondence to geography (Figure 3.1B), genotypic clustering not only does not correspond to geography, but does not correspond to sectional divisions either (Figure 3.6). It is therefore concluded that the hypothesis of Rydbergian sections is rejected by the presented molecular dataset.

The hypothesis of distinct populations appears to be the best supported hypothesis. This is perhaps best illustrated by hierarchical *F*-statistics (Table 3.2), where the greatest component of variance is explained by among-population differentiation. A potential weakness of the current study is that in order to attain broad taxonomic and geographic coverage a relatively small sample size of three to four individuals per population were used. This may have upwardly biased the estimate of among population differentiation as the sample is likely to only contain common haplotypes. Less common haplotypes which may be shared among populations might have been observed if within population sample sizes were increased; this would have the effect of decreasing the estimate of among population variance. With 55% of the variance explained by among population difference (Table 3.2), it seems that the among-population component is large enough that it could withstand substantial erosion and still support the hypothesis that the among population component of variance has the greatest explanatory effect.

Finding of no discernable pattern— With the rejection of the taxonomic hypotheses we are left with the null hypothesis of global panmixia and the best supported hypothesis of distinct populations. These two scenarios appear to be very different phenomena; however, their representation in our experimental design may be difficult to disentangle. These two hypotheses share the pattern of no discernable pattern. Global panmixia may be ideally observed as all haplotypes occurring in each of the populations. Based on the observed number of 57 haplotypes this would require a relatively large within population sample, which the present study lacks. While it is probably not necessary to sample 57 individuals per population to demonstrate relative global panmixia, it seems that the results of the hierarchical *F*-statistics (Table 3.2) are the best summary of the differences between global panmixia and distinct populations. These *F*-statistics demonstrate that 55% of the variance is accounted for by among-population differences, while 25% is accounted for by within population variance. A large percentage of variance explained by the within-population component would

argue that global panmixia was the best supported hypothesis. A large percentage of variance explained by the among-population component would argue that the distinct populations hypothesis is the best supported hypothesis. The finding that the among population component accounts for by far the greatest amount of variance, accompanied with the observation that many populations only had a single haplotype, indicates that the distinct populations hypothesis is best supported by the given data.

CpSSRs as inappropriate markers for inference of lineage— There have been several articles which have suggested that CpSSRs are not appropriate for the inference of lineage (Doyle et al. 1998; Hale et al. 2004; Navascués and Emerson. 2005). A recurring question in this argument is whether or not the mutation rate for CpSSRs is too, leading to haplotypes which are identical in state not due to descent (i.e., homoplasy or recurrent mutation). The problem of identity in state not due to descent is relatively independent of mutation rate but is related to mutational divergence among existing haplotypes. If a population consists of two haplotypes separated by ten mutational steps, it is unlikely that a new mutation will create a haplotype which is identical in state to one of the existing haplotypes. If, on the other hand, a population consists of two haplotypes separated by a single mutational step then it is relatively likely that a new mutation will lead to a new haplotype which is identical in state to an existing haplotype. Therefore the potential for genetic states which are identical in state not due to descent is related to the relative divergence of observed haplotypes. Lineages in the early stages of divergence are likely to exhibit allelic states that are susceptible to mutations which lead to identity in state not due to descent as a function of close identity of allelic state (i.e., small genetic distance) rather than due to mutation rate. It is also important to recognize that identity in state not due to descent is a transient state. If two isolated lineages share a haplotype which is identical in state not due to descent, mutation will likely replace these haplotypes with ones that are unique in state in evolutionary time. Genetic markers with relatively fast mutational rates will pass through this period rapidly while markers with slow mutational rates will take longer amounts of evolutionary time to traverse

this state. Salient question include what is the relative divergence of observed genetic states and how might this relate to the question of identity in state not due to descent, as opposed to the question of mutation rate.

Doyle et al. (1998) present issues of homoplasy (identity due to state but not descent) in a system of two CpSSR loci within the genus *Glycine* (Fabaceae). The present study utilizes five loci, which dramatically increases the number of allelic combinations that contribute to each haplotype, suggesting the issues presented by Doyle et al. (1998) may not be relevant to the current study.

Hale et al. (2004) compare seven CpSSRs to 2545 bp of chloroplast sequence (*trnL* intron, *trnL-F*, *atpB-rbcL*, and *accD-psaL*) in the genus *Clusia* (Clusiaceae). Based on the assumption that the sequence tree is the correct species tree, they demonstrate multiple instances where identical alleles have evolved in independent lineages (allelic identity by state not due to descent). It is important to note that while individual markers have mutated to identical states in different lineages the allelic combinations (haplotypes) are unique to each lineage (see Hale et al. 2004, Figures 1 and 2). This means that the assessment of lineage identity is reasonably estimated by CpSSRs while the relationship among lineages based on allelic differentiation may have been affected by issues of identity in state not due to descent. This suggests that lineage (haplotypic) identity may be robust to the problem of allelic identity in state not due to descent while inference of among lineage relationships may be susceptible to issues of identity in state not due to descent. Implications for the present study are that inference of lineage relatedness may be obscured by identity in state not due to descent (Figs. 3.3, 3.5, 3.6) but analysis focusing on haplotypic state should be robust to issues of allelic identity in state not due to descent (Table 3.2).

Navascués and Emerson (2005) employed a coalescence simulation to explore evolution of CpSSR haplotypes using a comparison between an infinite allele model and a stepwise mutation model to gain inference on levels of electromorph size

homoplasy. Because they employed a simulation, breeding structure was known, and their implementation was able to accurately assign new haplotypes different identities even if the haplotype were already present in the dataset (Navascués and Emerson 2005). This allowed them to use the decrease in haplotypes due to use of the stepwise mutation model compared to the infinite alleles model as an index of homoplasy. Many empirical datasets, such as the present dataset, lack knowledge of breeding structure making direct comparisons to Navascués and Emerson's challenging. Another complication in comparing the simulations of Navascués and Emerson is their focus on conifers, which have a relatively long generational time. This caused them to limit their simulations to 250 generations.

The present dataset contains numerous haplotypes which are a single mutational step apart (Figure 3.3) suggesting that even if mutation to haplotypes identical in state not due to descent has not been an issue in the past it is likely to be in the future. One should be cautious about interpreting these differences as phylogenetically informative and should perhaps see them as phylogenetically misleading. A short branch should be seen as a short branch, nothing more. Two important features of the present dataset are the taxon *A. l.* var. *salinus*, the most highly represented taxon in the dataset, and the two clusters of outgroups (Figure 3.3). That there are two clusters of outgroups positively indicates that no matter how the tree is rooted, non-*A. lentiginosus* members of *Astragalus* will be nested within *A. lentiginosus*. These non-*A. lentiginosus* haplotypes are not shared with *A. lentiginosus* yet their nesting suggests that if *A. lentiginosus* is an exemplar of lack of coalescence this issue may be widespread within the New World species of *Astragalus*. That *A. l.* var. *salinus* appears in multiple clusters (including outgroup clusters) indicates that this taxon is represented by relatively divergent haplotypes. Even if a majority of the observed haplotypes were the result of identity in state not due to descent, the haplotypes of the taxon *A. l.* var. *salinus* are sufficiently divergent to indicate that this pattern isn't simply due to identity in state not due to descent, but perhaps to the presence of haplotypes which predate the divergence of morphological taxa.

Lack of coalescence hypothesis— Rejection of the taxonomic hypotheses at both the varietal level (Barneby 1964) and the sectional level (Rydberg 1929) indicates either a lack of pattern (global panmixia) or a pattern where each sampled unit (the population) is highly differentiated (distinct populations). Hierarchical *F*-statistics differentiate among these hypotheses, suggesting that populations are the most distinct hierarchical unit. If this differentiation was due to the mutational process, it seems reasonable to expect a pattern of isolation by distance, however hierarchical Bayesian clustering suggests that isolation by distance is not a significant pattern. An alternate hypothesis is that the haplotypes observed are older than the lineages (morphological varieties) which were the focus of this research. This is synonymous with the idea that we are observing incomplete lineage sorting or a lack of coalescence. If this alternate hypothesis is correct, there does not appear to be a reasonable expectation to observe cladogenesis among the varieties or sections of *A. lentiginosus*. Significant patterns in morphological diversity (Knaus Chapter 2) indicate that taxonomic circumscription, while not discrete, is correlated with climatic parameters which may have led to the (at least initial stages of) divergence. Observed molecular diversity does not correspond to these morphological taxa. This suggests that the gene genealogy inferred in this study may not be tracking phenotypic divergence. Knaus et al. (2008) employed coalescence simulations to demonstrate that, given the observed haplotypic diversity in the observed dataset, cladogenesis of Rydbergian groups may take 0.5-1.0 million generations since the establishment of reproductive isolation. A new hypothesis is proposed. The varieties of *A. lentiginosus* appear to form a morphological cline in response to climate (which may be adaptive). From a neutral molecular standpoint they have ancestral haplotypes which have yet to sort into varietal lineages, or be replaced by mutations characteristic to these lineages. These empirical results are likely due to a recent radiation of *A. lentiginosus* into numerous varieties from an ancestral population of large effective size and without a bottleneck leading to these new lineages.

How is the hypothesis of incomplete lineage sorting a falsifiable hypothesis? The morphometric analysis of phenotypic traits (Knaus Chapter 2), which are assumedly available for selection to act upon, resulted in a significant pattern suggesting that current varietal hypotheses (Barneby 1964; Welsh 2007) are supported, although the differences among varieties appear to be clinal in nature as opposed to being discrete (Knaus Chapter 2). This indicates that the phenotypic signal of lineage divergence may be present. The apparent lack of correspondence of chloroplast simple sequence repeat haplotypes to morphological varieties (Figs. 3.3, 3.5, 3.6 and Table 3.2) provides a qualitative test of assumedly neutral molecular divergence against phenotypic divergence. The result is a lack of congruence. Phenotypic opinion (Barneby 1964; Isely 1989; Welsh 2007) and morphometric analysis appear to support the idea of non-discrete, clinal morphological entities. This presents the hypothesis of lineage based on taxonomy (Barneby 1964; Isely 1989; Welsh 2007) and morphometric analysis (Knaus Chapter 2) as a refutable hypothesis. The present CpSSR dataset represents a lack of discernable pattern. There are two interpretations of this result: either the morphological taxonomy is unsupported or that molecular diversity predates the lineages of interest. Given previous research (Knaus Chapter 2), it is reasonable to assume that the hypothesis of phenotypic divergence without molecular support is due to a lack of coalescence.

Conclusion— Observed haplotypic diversity within *A. lentiginosus* does not support taxonomic hypotheses at either the varietal (Barneby 1964) or the sectional level (Rydberg 1929). This supports hypotheses which lack discernable pattern, such as global panmixia and distinct populations. Hierarchical *F*-statistics support the hypothesis of distinct populations over that of global panmixia, despite caveats in experimental design. Based on a lack of isolation by distance, the pattern of distinct populations appears to be due to the sorting of ancestral haplotypes, as opposed to the creation of new haplotypes subsequent to lineage divergence. The apparent recent radiation of *Astragalus* in the New World (Wojciechowski et al.1999) appears to have included the more recent radiation of *A. lentiginosus*. This radiation appears to have

been produced from an ancestral population of large haplotypic diversity, which has yet to sort into sectional or varietal lineages, assuming these morphological lineages represent evolutionary lineages. The observed lack of correspondence of assumedly neutral haplotypic diversity to the presumptuously selected phenotype (Knaus Chapter 2) suggests that morphologically described varieties of *A. lentiginosus* may have resulted from a greater signal of selection relative to neutral divergence (Lewontin and Krakauer 1973; Wu 2001; Via 2002; Beaumont and Balding 2004; Beaumont 2005).

3.5 LITERATURE CITED

Alexander, J. A., A. Liston, and S. J. Popovich. 2004. Genetic diversity of the narrow endemic *Astragalus oniciformis* (Fabaceae). American Journal of Botany 91:2004-2012.

Allphin, L., N. Brian, and T. Matheson. 2005. Reproductive success and genetic divergence among varieties of the rare and endangered *Astragalus cremnophylax* (Fabaceae) from Arizona, USA. Conservation Genetics 6:803-821.

Avise, J. C. 2004, Molecular markers, natural history, and evolution. Sunderland, Mass., Sinauer Associates.

Barneby, R. C. 1964. Atlas of North American *Astragalus*. Memoirs of the New York Botanical Garden 13:1-1188.

Beaumont, M. A. 2005. Adaptation and speciation: what can F_{ST} tell us? Trends in Ecology & Evolution 20:435-440.

Beaumont, M. A., and D. J. Balding. 2004. Identifying adaptive genetic divergence among populations from genome scans. Molecular Ecology 13:969-980.

Becker, R. A., and A. R. Wilks. 1993. Maps in S. AT&T Bell Laboratories Statistics Research Report 95.2.

Brownstein, M. J., J. D. Carpten, and J. R. Smith. 1996. Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. BioTechniques 20:1004-1010.

Criscione, C. D., and M. S. Blouin. 2007. Parasite phylogeographical congruence with salmon host evolutionary significant units: implications for salmon conservation. Molecular Ecology 16:993-1005.

Darwin, C. 1859, On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. *In* D.M. Porter and P.W. Graham *eds.* The Portable Darwin, 1993, Penguin Books.

de Meeûs, T., and J. Goudet. 2007. A step-by-step tutorial to use HierFstat to analyse populations hierarchically structured at multiple levels. Infection, Genetics and Evolution 7:731-735.

Dobzhansky, T. 1951, Genetics and the Origin of Species. New York, Columbia University Press.

Doyle, J. J., M. Morgante, S. V. Tingey, and W. Powell. 1998. Size homoplasy in chloroplast microsatellites of wild perennial relatives of soybean (*Glycine* subgenus *Glycine*). Molecular Biology and Evolution 15:215-218.

Falush, D., M. Stephens, and J. K. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567-1587.

—. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular Ecology Notes 7:574-578.

Felsenstein, J. 2004, Inferring Phyogenies, Sinauer.

Fisher, R. A. 1958, The Genetical Theory of Natural Selection, Dover.

Goudet, J. 2005. HIERFSTAT, a package for R to compute and test hierarchical *F*-statistics. Molecular Ecology Notes 5:184-186.

Guillot, G., A. Estoup, F. Mortier, and J. F. Cosson. 2005. A spatial statistical model for landscape genetics. Genetics 170:1261-1280.

Guillot, G., F. Mortier, and A. Estoup. 2005. Geneland: a program for landscape genetics. Molecular Ecology Notes 5:712-715.

Hale, M. L., A. M. Borland, M. H. G. Gustafsson, and K. Wolff. 2004. Causes of size homoplasy among chloroplast microsatellites in closely related *Clusia* species. Journal of Molecular Evolution 58:182-190.

Hedrick, P. W. 2005. A standardized genetic differentiation measure. Evolution 59:1633-1638.

Karron, J. D., Y. B. Linhart, C. A. Chaulk, and C. A. Robertson. 1988. Genetic structure of populations of geographically restricted and widespread species of *Astragalus* (Fabaceae). American Journal of Botany 75:1114-1119.

Knaus, B. J., R. C. Cronn, and A. Liston. 2005. Genetic characterization of three varieties of *Astragalus lentiginosus* (Fabaceae). Brittonia 57:334-344.

—. 2008. A fistful of *Astragalus*: incipient speciation in the American West? Presentation at Botany 2008.

Lewontin, R. C., and J. Krakauer. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. Genetics 74:175-195.

Liston, A. 1992. Isozyme Systematics of Astragalus Sect. Leptocarpi Subsect. Californici Fabaceae. Systematic Botany 17:367-379.

Lock, J. M., and B. D. Schrire. 2005. Galegeae, Pages 475-487 *in* G. Lewis, B. Schrire, B. MacKinder, and M. Lock, eds. Legumes of the World. Kew, Royal Botanical Gardens.

Mabberley, D. J. 1987, The plant-book : a portable dictionary of the higher plants utilising Cronquist's An integrated system of classification of flowering plants (1981) and current botanical literature arranged largely on the principles of editions 1-6 (1896/97-1931) of Willis's A dictionary of the flowering plants and ferns. New York, Cambridge University Press.

Mayr, E., and P. D. Ashlock. 1991, Principles of Systematic Zoology, McGraw-Hill.

Morris, A. B., R. S. Baucom, and M. B. Cruzan. 2002. Stratified analysis of the soil seed bank in the cedar glade endemic *Astragalus bibullatus*: evidence for historical changes in genetic structure. American Journal of Botany 89:29-36.

Navascués, M., and B. C. Emerson. 2005. Chloroplast microsatellites: measures of genetic diversity and the effect of homoplasy. Molecular Ecology 14:1333-1341.

Paradis, E. 2006, Analysis of phylogenetics and evolution with R. New York, Springer.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

R Development Core Team. 2007.R: a language and environment for statistical computing.R foundation for statistical computing, Vienna, Austria.

Rieseberg, L. H., and J. M. Burke. 2001. The biological reality of species: Gene flow, selection, and collective evolution. Taxon 50:47-67.

Rosenberg, N. A., and M. Nordborg. 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. Nature Reviews Genetics 3:380-390.

Rydberg, P. A. 1929. *Astragalanae*. North American Flora 24:251-462.

Scherson, R. A., R. Vidal, and M. J. Sanderson. 2008. Phylogeny, biogeography, and rates of diversification of New World *Astragalus* (Leguminosae) with an emphasis on South American radiations. American Journal of Botany 95:1030-1039.

Slatkin, M. 1976. The rate of spread of an advantageous allele in a subdivided population *in* S. Karlin, and E. Nevo, eds. Population Genetics and Ecology. New York, NY, Academic Press, Inc.

Spellenberg, R. 1976. Chromosome numbers and their cytotaxonomic significance for North American *Astragalus* (Fabaceae). Taxon 25:463-476.

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 17:1105-1109.

Travis, S. E., J. Maschinski, and P. Keim. 1996. An analysis of genetic variation in *Astragalus cremnophylax* var. *cremnophylax*, a critically-endangered plant, using AFLP markers. Molecular Ecology 5:735-745.

Via, S. 2002. The ecological genetics of speciation. The American Naturalist 159:S1-S7.

Weising, K., and R. C. Gardner. 1999. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. Genome 42:9-19.

Welsh, S. L. 2007, North American species of *Astragalus* Linnaeus (Leguminosae), a taxonomic revision. Prove, UT, Monte L. Bean Life Science Museum.

Wojciechowski, M. F., M. J. Sanderson, B. G. Baldwin, and M. J. Donoghue. 1993. Monophyly of aneuploid *Astragalus* (Fabaceae): evidence from nuclear ribosomal DNA internal transcribed spacer sequences. American Journal of Botany 80:711-722.

Wojciechowski, M. F., M. J. Sanderson, and J.-M. Hu. 1999. Evidence on the monophyly of *Astragalus* (Fabaceae) and its major subgroups based on nuclear ribosomal DNA ITS and chloroplast DNA *trnL* intron data. Systematic Botany 24:409-437.

Wu, C.-I. 2001. The genic view of the process of speciation. Journal of Evolutionary Biology 14:851-865.

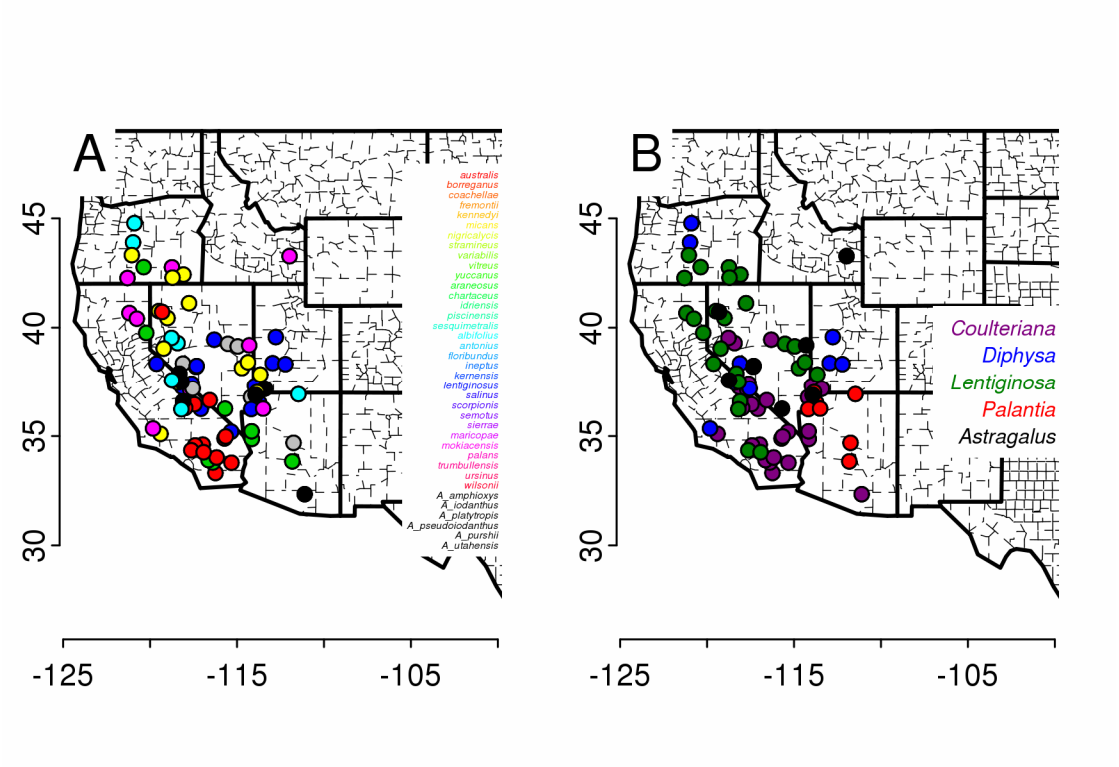


Figure 3.1. Map of the CpSSR sample. Panel A is organized by variety (and out-group species). Panel B is labeled by sections proposed by Rydberg (1929).

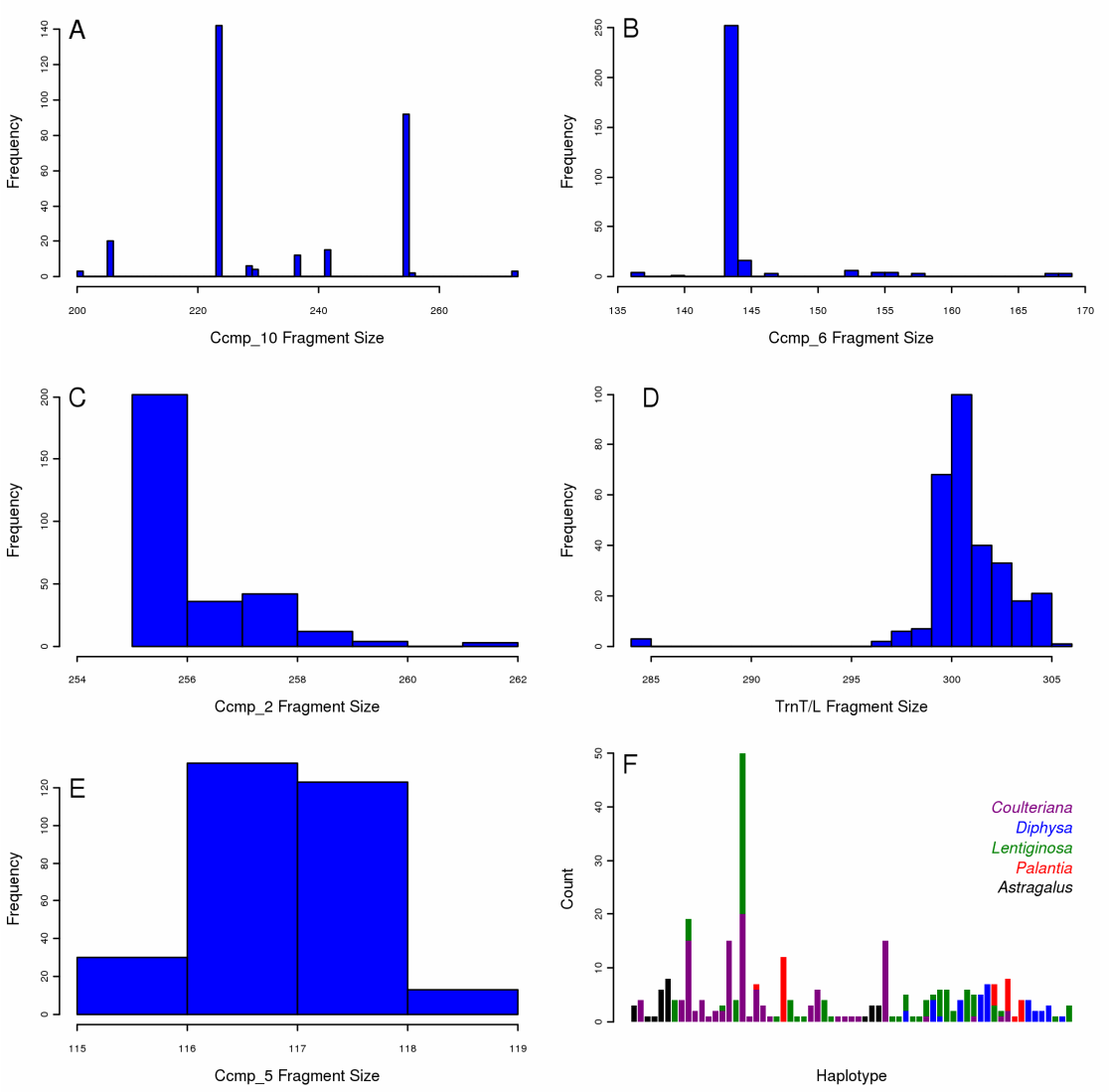
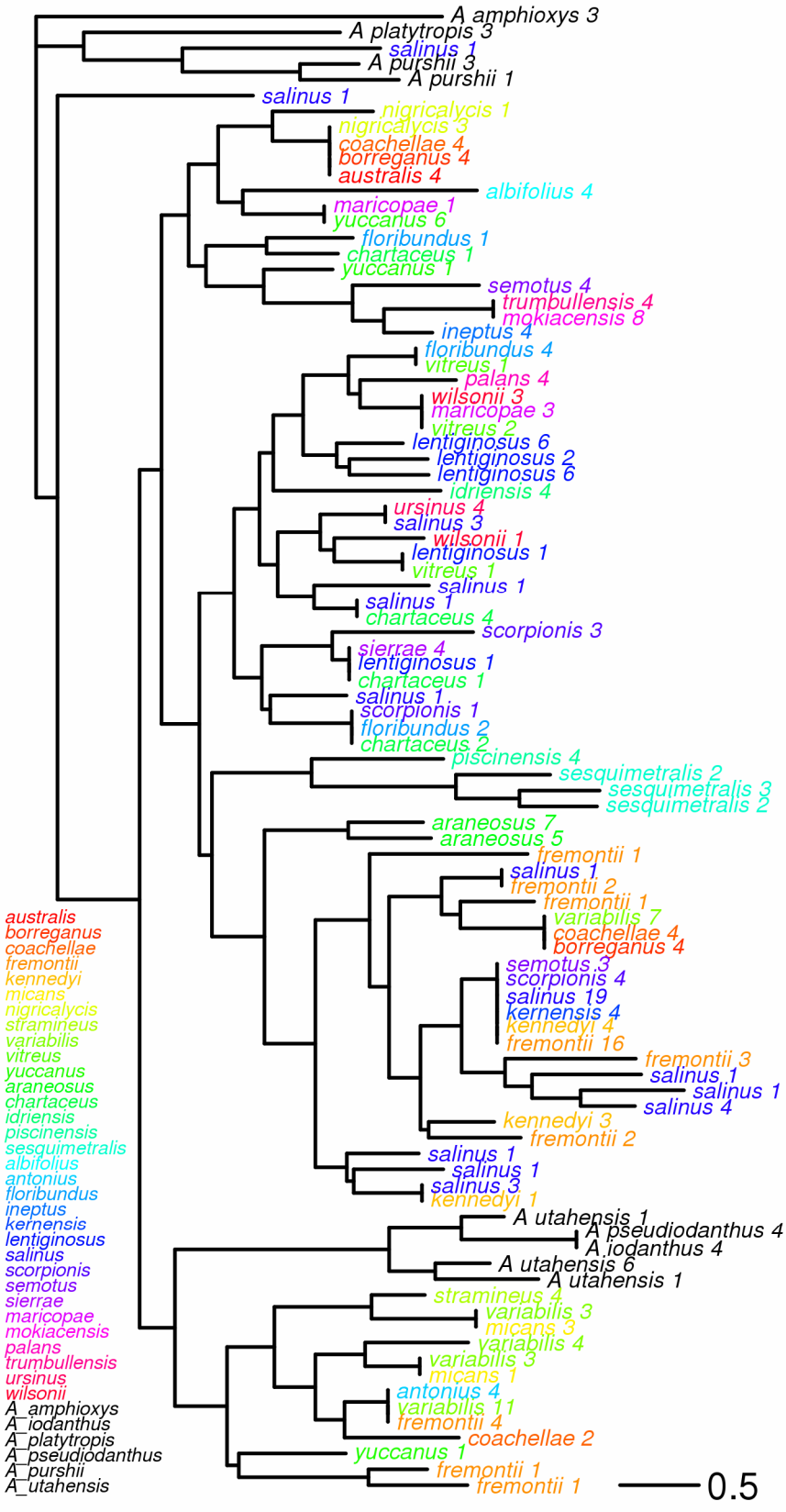


Figure 3.2. Histograms of CpSSR allele frequencies. The marker ccmp10 (Pane A) shows a pattern that is not representative of a mononucleotide repeat, but of a compound indel. The marker trnTL839F (pane D) shows an anomolously small allele at 285 base pairs observed in *Astragalus amphioxus*. Figure 3.2F shows haplotypes on the x-axis which are unordered.

Figure 3.3. Neighbor-joining tree based on CpSSRs. A total of 57 haplotypes were observed within four sections, 33 varieties and 71 populations of *Astragalus lentiginosus*. Eight haplotypes were observed in outgroup species of *Astragalus*. Numbers following tip labels indicate number of individuals sharing that haplotype. A pairwise distance computed from five CpSSR markers was used.


A amphioxys 3
A platytropis 3
salinus 1
A purshii 3
A purshii 1
salinus 1
nigricalycis 1
nigricalycis 3
coachellae 4
borreganus 4
australis 4
albifolius 4
maricopae 1
yuccanus 6
floribundus 1
chartaceus 1
yuccanus 1
semotus 4
trumbullensis 4
mokiacensis 8
ineptus 4
floribundus 4
vitreus 1
palans 4
wilsonii 3
maricopae 3
vitreus 2
lentiginosus 6
lentiginosus 2
lentiginosus 6
idriensis 4
ursinus 4
salinus 3
wilsonii 1
lentiginosus 1
vitreus 1
salinus 1
salinus 1
chartaceus 4
scorpionis 3
sierrae 4
lentiginosus 1
chartaceus 1
salinus 1
scorpionis 1
floribundus 2
chartaceus 2
piscinensis 4
sesquimetralis 2
sesquimetralis 3
sesquimetralis 2
araneosus 7
araneosus 5
fremontii 1
salinus 1
fremontii 2
fremontii 1
variabilis 7
coachellae 4
borreganus 4
semotus 3
scorpionis 4
salinus 19
kernensis 4
kennedyi 4
fremontii 16
fremontii 3
salinus 1
salinus 1
salinus 4
kennedyi 3
fremontii 2
salinus 1
salinus 1
salinus 3
kennedyi 1
A utahensis 1
A pseudiodanthus 4
A iodanthus 4
A utahensis 6
A utahensis 1
stramineus 4
variabilis 3
micans 3
variabilis 4
variabilis 3
micans 1
antonius 4
variabilis 11
fremontii 4
coachellae 2
yuccanus 1
fremontii 1
fremontii 1
0.5
australis
borreganus
coachellae
fremontii
kennedyi
micans
nigricalycis
stramineus
variabilis
vitreus
yuccanus
araneosus
chartaceus
idriensis
piscinensis
sesquimetralis
albifolius
antonius
floribundus
ineptus
kernensis
lentiginosus
salinus
scorpionis
semotus
sierrae
maricopae
mokiacensis
palans
trumbullensis
ursinus
wilsonii
A_amphioxys
A_iodanthus
A_platytropis
A_pseudiodanthus
A_purshii
A_utahensis

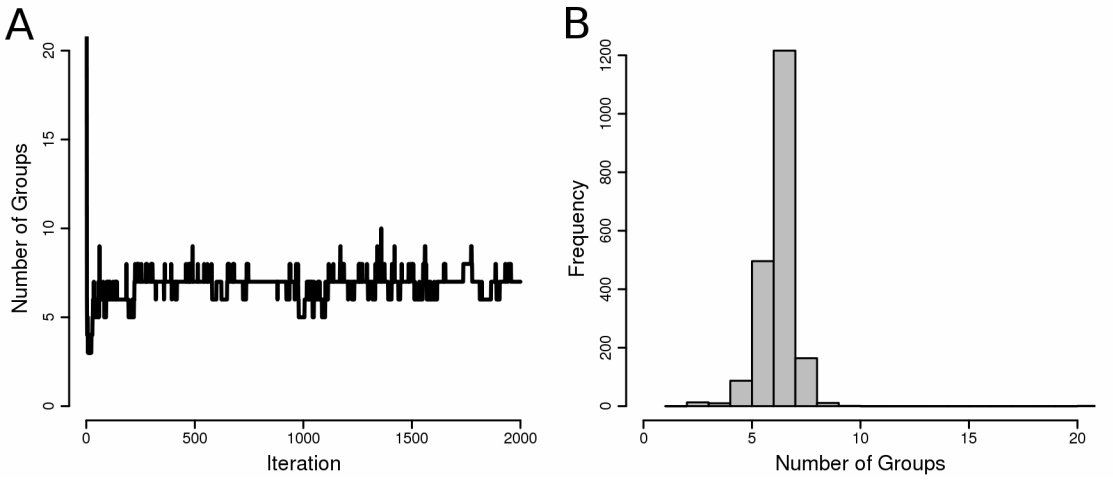


Figure 3.4. Convergence diagnostics for Bayesian clustering. Pane A plots number of groups by MCMC iteration. Pane B is a histogram plotting density of number of groups.

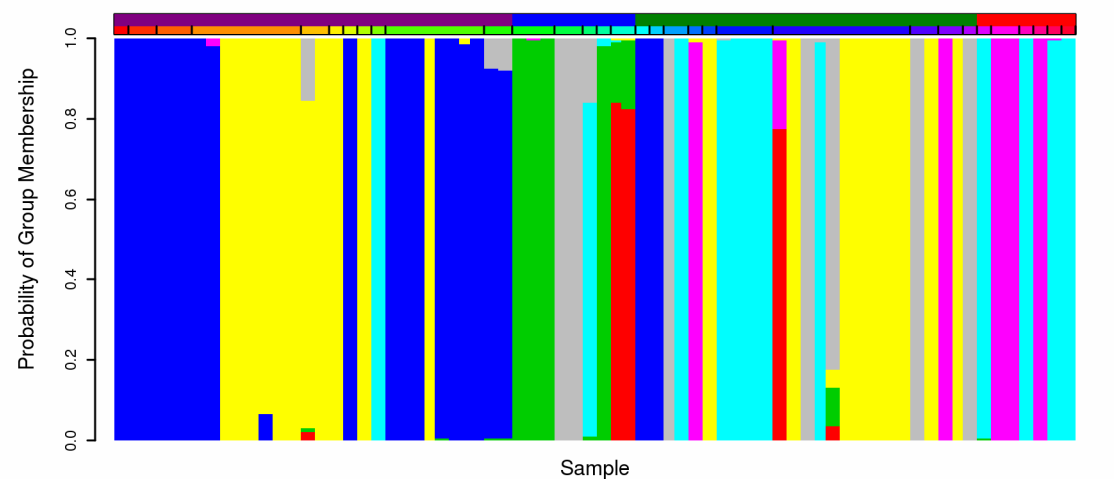
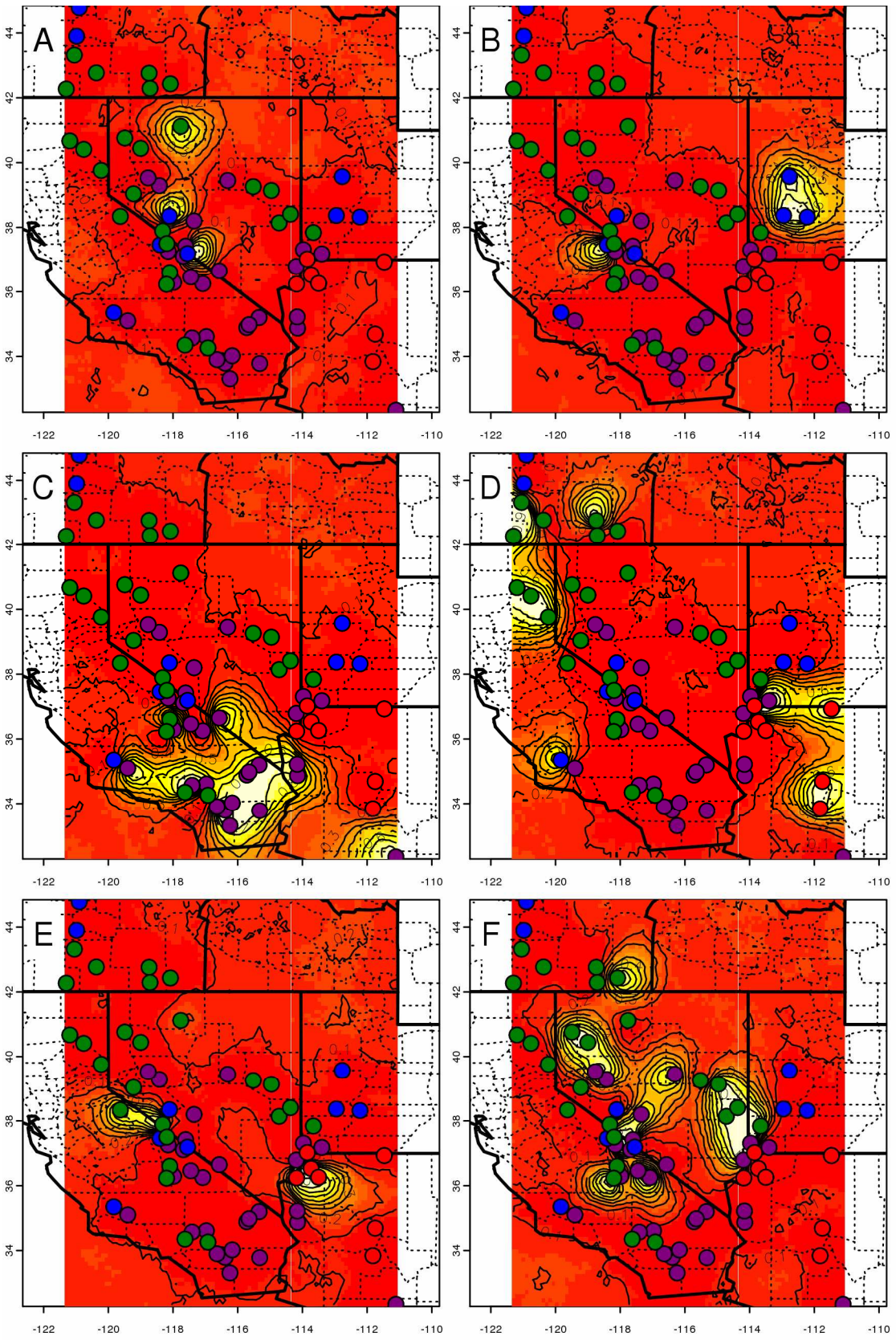


Figure 3.5. Barplot of probability of group membership from Bayesian clustering. Vertical bars represent individual samples (n=272) and colors represent the probability an individual belongs to a group. First horizontal row above y=1 represents varietal hypotheses. Second horizontal line above y=1 represents sections proposed by Rydberg (1929) colored to follow Figure 3.1.


A
B
C
D
E
F
44
42
40
38
36
34
-122
-120
-118
-116
-114
-112
-110

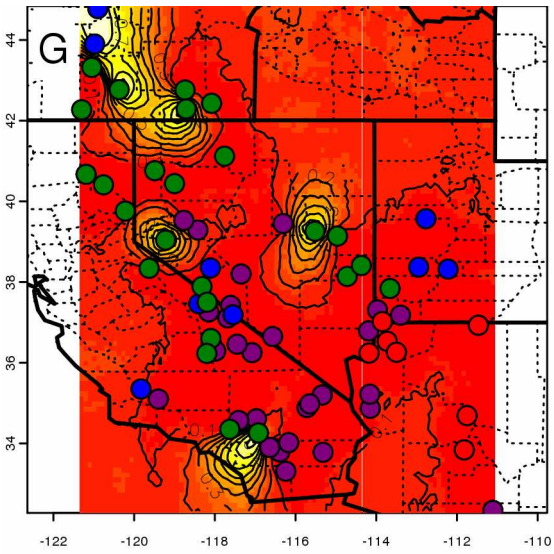


Figure 3.6. Spatial interpolations of posterior probability of group membership. Color of points follows Figure 3.1B. Probability of group membership ranges from zero (red) to one (white). The x-axis is longitude while the y-axis is latitude.

Table 3.1. Summary of the CpSSR sample. Number of populations precedes the slash while total number of individuals follows the slash. Barneby (1964) reports the number of specimens (Barneby's samples) viewed during preparation of his monograph as well as the number which were his own collections (Barneby's specimens).

Variety	Rydbergian Section	Distribution	Barneby's Sample	Barneby's Specimens	CpSSR pops
A.l. var. *ambiguus*	*Coulteriana*	endemic	4	1	1/4
A.l. var. *australis*	*Coulteriana*	widespread	31	5	1/4
A.l. var. *borreganus*	*Coulteriana*	widespread	31	4	2/8
A.l. var. *coachellae*	*Coulteriana*	endemic	34	2	3/10
A.l. var. *fremontii*	*Coulteriana*	widespread	101	14	8/31
A.l. var. *kennedyi*	*Coulteriana*	widespread	31	3	2/8
A.l. var. *micans*	*Coulteriana*	endemic	5	0	1/4
A.l. var. *nigricalycis*	*Coulteriana*	widespread	56	2	1/4
A.l. var. *stramineus*	*Coulteriana*	endemic	9	1	1/4
A.l. var. *variabilis*	*Coulteriana*	widespread	114	9	8/28
A.l. var. *vitreus*	*Coulteriana*	endemic	16	2	1/4
A.l. var. *yuccanus*	*Coulteriana*	endemic	16	3	1/4
A.l. var. *araneosus*	*Diphysa*	widespread	39	8	3/12
A.l. var. *chartaceus*	*Diphysa*	widespread	56	8	2/8
A.l. var. *diphysus*	*Diphysa*	widespread	66	10	NA
A.l. var. *higginsii*	*Diphysa*	endemic	NA	NA	NA
A.l. var. *idriensis*	*Diphysa*	endemic	24	1	1/4
A.l. var. *latus*	*Diphysa*	endemic	5	1	NA
A.l. var. *multiracemosus*	*Diphysa*	endemic	NA	NA	NA
A.l. var. *negundo*	*Diphysa*	endemic	NA	NA	NA
A.l. var. *oropedii*	*Diphysa*	endemic	7	0	NA
A.l. var. *piscinensis*	*Diphysa*	endemic	NA	NA	1/4
A.l. var. *pohlii*	*Diphysa*	endemic	NA	NA	NA
A.l. var. *sesquimetralis*	*Diphysa*	endemic	1	0	2/7
A.l. var. *albifolius*	*Lentiginosa*	endemic	16	3	1/4
A.l. var. *antonius*	*Lentiginosa*	endemic	8	0	1/4
A.l. var. *floribundus*	*Lentiginosa*	widespread	26	2	2/7
A.l. var. *ineptus*	*Lentiginosa*	endemic	23	2	1/4
A.l. var. *kernensis*	*Lentiginosa*	endemic	8	0	1/4
A.l. var. *lentiginosus*	*Lentiginosa*	widespread	73	3	4/15
A.l. var. *salinus*	*Lentiginosa*	widespread	102	12	10/39
A.l. var. *scorpionis*	*Lentiginosa*	endemic	26	1	2/8
A.l. var. *semotus*	*Lentiginosa*	endemic	14	0	2/7
A.l. var. *sierrae*	*Lentiginosa*	endemic	18	2	1/4
A.l. var. *maricopae*	*Palantia*	endemic	5	0	1/4
A.l. var. *mokiacensis*	*Palantia*	endemic	5	0	2/8
A.l. var. *palans*	*Palantia*	widespread	40	9	1/4
A.l. var. *trumbellensis*	*Palantia*	endemic	NA	NA	1/4
A.l. var. *ursinus*	*Palantia*	endemic	1	0	1/4
A.l. var. *wilsonii*	*Palantia*	endemic	16	2	1/4
40 varieties			**1027**	**110**	**71/272**

Table 3.2. Hierarchical *F*-statistics.

	d.f.	variance	% variance	*F*	*F*-max	*F′*	p-value*
among sections	3	0.0341	3.51	0.035	0.062	0.562	0.001
among varieties within sections	28	0.1647	16.92	0.175	0.224	0.781	0.001
among populations witihin varieties	39	0.5339	54.83	0.689	0.756	0.911	0.001
within populations	202	0.2409	24.74				

*based on 1,000 permutations

4 Amplified fragment length polymorphisms

4.1 INTRODUCTION

Infrataxa in the study of speciation— Understanding the process of speciation has been a subject of philosophical discourse since its now widely accepted proposal (Darwin 1859), through the Modern Synthesis (Dobzhansky 1951; Fisher 1958; Wright 1978) and into the contemporary literature (Lewontin and Krakauer 1973; Nordborg 2001; Orr 2001; Wu 2001; Rieseberg & Burke 2001; Hudson & Coyne 2002; Rosenburg & Nordborg 2002; Via 2002; Beaumont & Balding 2004; Coyne & Orr 2004; Morjan & Rieseberg 2004; Beaumont 2005; Rieseberg et al. 2006; Syring et al. 2007). Ideally, species are entities which have independent evolutionary fates, a trait perhaps best characterized by the lack for the potential of gene flow. This criterion is represented by the Biological Species Concept (Mayr & Ashlock 1991; Coyne & Orr 2004). However, currently recognized taxa frequently fail to live up to our philosophical expectations. This is perhaps best evidenced by a large laundry list of species concepts (Mayden 1997), as well as increasing empirical studies (Syring et al. 2007). This may, at least in part, be due to the many pathways which lead to reproductive isolation such as allopatry, polyploidy, chromosomal inversion, Dobzhansky-Muller incompatibilities or adaptive divergence.

If we view a species as an array of populations inhabiting non-identical habitats, it becomes reasonable to assume these populations are experiencing the diversifying forces of neutral genetic drift and local adaptation as well as the homogenizing force of geneflow. This heterogeneity can result in morphological differences which can be captured at the infraspecific ranks (i.e., the subspecies and variety). The study of infra-taxa can then be seen as the study of incipient speciation. If diversifying forces are great relative to homogenizing forces, the result may be new evolutionary lineages. Here, amplified fragment length polymorphisms are utilized to explore the genetic architecture of the most taxon rich species in the U.S. flora. The large number of taxa

included within this species presents itself as heterogeneous species with a potential to fracture into numerous evolutionary lineages.

A. lentiginosus as a potential for speciation?— With 40 taxonomic varieties, the desert legume *Astragalus lentiginosus* Douglas ex Hooker (Fabaceae) contains more taxa than any other species of vascular plant in the U.S. flora (Knaus Chapter 2). If each of these varieties are considered as incipient species, the group displays a potential for mass speciation. As a member of the most species rich genus of vascular plant (Mabberley 1987; Lock and Schrire 2005), *A. lentiginosus* may also be seen as an exceptionally diverse species within an exceptionally diverse genus. Current occupation of geologically young habitats (e.g., inland dune systems, desert seeps, montane ridges) lends credence to the hypothesis of recent radiation (i.e., post-Pleistocene). Previous research has demonstrated morphological divergence correlated with climatic parameters but a lack discontinuity which may be required for diagnosability (Knaus Chapter 2). Molecular analysis of *A. lentiginosus* utilizing chloroplast simple sequence repeats (CpSSRs) demonstrated a lack of correspondence of molecular diversity to morphological taxa (Knaus Chapter 3). Coalescence simulation has demonstrated that based on the observed CpSSR diversity within *A. lentiginosus* (Knaus Chapter 3), it may take between 0.5 and 1.0 million generations from reproductive isolation before cladogenesis may be observed (Knaus et al. 2008). Investigation into the species *Astragalus lentiginosus* presents potential insight into an actively radiating species which may include elements of local adaptation, diversifying neutral drift and collective evolution at the hierarchical levels of species, Rydbergian group (Knaus Chapter 2, 3; Rydberg 1929) and variety (Barneby 1964) and population.

Fidelity of the AFLP method— Amplified fragment length polymorphisms have been employed to provide high marker content, dominant, nuclear perspective on varieties of *A. lentiginosus* endemic to the Mojave Desert and its immediate surroundings. This marker has been used to address hypotheses at the level of species and below within

other parts of the range of *A. lentiginosus* (Knaus et al. 2005), in other species of *Astragalus* (Travis et al. 1996) and in other legumes such as *Oxytropis campestris* (Chung et al. 2004; Schönswetter et al. 2004) and *Anthyllis montanum* (Kropf et al. 2002). The AFLP method has been reported as a high fidelity marker by many researchers (Jones et al. 1997; Jones et al. 1998; Amsellem et al. 2000) and has been successfully used to score heterozygosity in pedigreed samples (Rouppe van der Voort et al. 1997), although this is not recommended for wild populations. Informally, the AFLP method appears to have a somewhat questionable reputation. Whitlock et al. (2008) report an average of 25 peaks per profile above a threshold of 50 rfus resulting from the genotyping of water, demonstrating issues associated with scoring samples lacking expected fragment sizes. Bonin et al. (2004) reported large variation among different people scoring AFLP bands. This did not result in large differences in results among different scorings of the data, lending credence to the idea that AFLPs have a high signal to noise ratio. Here I explore how to characterize genotyping error in large datasets and how to manage its inevitable manifestation.

Interpretation of clustering results— Hierarchical Bayesian model based clustering tools such as Structure (Pritchard et al. 2000; Falush et al. 2003; Falush et al. 2007), BAPS (Corander and Marttinen 2006) and Geneland (Guillot et al. 2005; Guillot et al. 2005) have become popular tools for exploring the optimal number of groups contained in a sample, as well as membership to these groups (untrained clustering). Interpretation of what is ‘optimal’ can be obfuscated by model parameterizations which result in similar posterior probabilities. Stated differently, there may be no clearly ‘optimal’ number of groups, but rather several candidates of similar likelihood. Evanno et al. (2005) employed simulation to suggest that using the rate of change in the likelihood of the model to infer 'major' structure within a sample may be the most appropriate interpretation of optimal group number. Here I advocate a perspective where a range of ‘optimal’ number of groups is explored (Rosenberg et al. 2002). The recommendation of Evanno et al. (2005) may be employed as a lower bound to appropriate values of group number to be explored, and I provide a method for

assigning an upper bound on optimal group number which appears to have a straightforward interpretation.

Amplified fragment length polymorphisms are employed to investigate evolutionary relationships within the taxon-rich desert legume *Astragalus lentiginosus* as an exploration into a taxon which may be diversifying into species. Issues in the scoring of AFLP phenotypes are addressed, as well as their use in the clustering program Structure (Falush et al. 2007). The relative strengths and weaknesses of these methods are discussed and a biological interpretation of the results is proposed.

4.2 METHODS

Experimental design—In order to explore evolutionary relationships within *A. lentiginosus*, I chose to sample varieties within Rydberg's (1929) section *Coulteriana*, a group distributed over southern California and southern Nevada (Table 4.1, Figure 4.1). Leaf tissue was collected in the field and preserved either on silica gel or on dry ice. The number of populations sampled corresponds largely to the distribution of each taxon. The varieties *fremontii* and *variabilis* are widespread and therefore occur frequently in the dataset. The variety *micans* occurs only at Eureka Valley, CA and thus occurs only once in the dataset. The variety *nigricalycis* occurs in the San Juaquin Valley, CA, a region of intensive agriculture. Collection activities were successful in locating only one population. Therefore this taxon occurs as one population in the dataset. The variety *borreganus* is geographically restricted but occurs in two disjunct regions: the Sonoran Desert of California and the mid-Mojave desert. Samples from both regions occur in the dataset. The sample from the Sonoran desert (Font's Wash) was propagated from seed contributed by the Desert Legume Program (Tucson, AZ). These seeds were scarified with sand paper and germinated on filter paper. When the cotyledons were fully emerged the samples were homogenized for DNA extraction. Because these samples were from seed, it is possible they are maternal siblings (i.e., collected from one plant). The variety *coachellae* is listed as endangered under the U.S. Endangered Species Act (USFWS

1998). This taxon is restricted to the Coachella Valley. Due to its federal listing it has special interest and is included as three populations.

The sample resulting in the highest DNA concentration after extraction was selected as a control. This turned out to be a sample of *A. l.* var. *fremontii* from the central Mojave Desert of California (Figure 4.1; Table 4.1). This sample was included six times as a control, twice in each of the three 96-well plates in which the reactions were performed. This control sample was also included as a sample associated its population of origin, for a total of seven occurrences in the experiment. This control was the product of a single DNA extraction but independent AFLP chemistry.

Amplified fragment length polymorphism genotyping— Amplified fragment length polymorphism (AFLP) chemistry was performed on a total of 288 samples (three 96-well plates) according to the protocols of Vos et al. (1995). Primer pairs were initially screened in the closely related *A. utahensis* (Wojciechowski et al. 1993; Wojciechowski et al. 1999) in a parallel project and selected for high band number. Selective amplifications (+3) consisted of: [FAM]-*Eco*RI+AGC *Mse*I+GAA, [HEX]-*Eco*RI+ACA *Mse*I+GAC, [FAM]-*Eco*RI+ACA *Mse*I+GAT, [HEX]-*Eco*RI+AGC *Mse*I+GAT, [FAM]-*Eco*RI+ACG *Mse*I+GAT and [HEX]-*Eco*RI+ACA *Mse*I+GAA. Chemistry was performed in three 96-well plates. Samples were multiplexed by combining one FAM and one HEX fluorophore (a total of nine plates) and submitted to the Oregon State University Core Laboratories facility for genotyping on ABI 3100 capillary DNA genotypers (http://corelabs.cgrb.oregonstate.edu/genotype).

ABI traces were scored using Genotyper 3.7 (Applied Biosystems) with the criteria of all peaks between 200 and 490 base pairs in size and 50 to 6,000 rfus in height for each fluorophore. The data were exported as tab delimited tables for further processing.

Automated Scoring— In order to achieve a standardized and repeatably scored data set, a rule based system was implemented using scripts written in the R language (R Development Core Team 2007) and modified functions in the R package genomatic v0.1 (available at: http://oregonstate.edu/~knausb/). Bins for each primer pair were determined by concatenating all of the bands called for the primer pair by Genotyper into a vector and looking for a gap parameterized in base pairs. Every time this gap was encountered a new bin was created. Here this gap was parameterized to 0.3 base pairs (an arbitrary choice which experience indicated was appropriate). Summary statistics for each bin were calculated including the mean size, standard deviation, minimum and maximum values in units of base pairs. Bins that had more bands than there were samples (i.e., 288) were discarded. Bins which had an excessive standard deviation (i.e., the bin spanned several base pairs) were also discarded. In order to explore the effect of bin width on the scored dataset, maximum bin widths of 0.2, 0.4, 0.5, 0.6, 0.8, 1.0, 2.0 and 3.0 standard deviations were used to define bins.

A quality control step was also performed. Samples which amplified less than three bands for any of the primer pairs were removed from the entire dataset. This also provided a summary for quality of each primer pair: *Eco*RI+AGC *Mse*I+GAA had 23 failures, *Eco*RI+ACA *Mse*I+GAC had 29 failures, *Eco*RI+ACA *Mse*I+GAT had 3 failures, *Eco*RI+AGC *Mse*I+GAT had 40 failures, *Eco*RI+ACG *Mse*I+GAT had 72 failures and *Eco*RI+ACA *Mse*+GAA had 7 failures. Because of the high incidence of failed samples for *Eco*RI+AGC *Mse*I+GAT and *Eco*RI+ACG *Mse*I+GAT, these two primer pairs were omitted from subsequent analyses.

In order to explore the performance of the controls a matrix of Jaccard distances and percent percent pair-wise distances was constructed using the R packages ade4 (Chessel et al. 2004) and ape (Paradis et al. 2004; Paradis 2006), respectively. The proportion of the dataset containing ones (band presence) was calculated for all tested bin widths by summing the entire matrix and dividing by the number of cells. Lastly,

the number of bins resulting from each of the tested bin width was summed to provide a summary of the number of bins resulting from each bin width parameterization.

Band occurrence was summarized with a barplot constructed in R (R Development Core Team 2007). A neighbor-joining tree was constructed from Jaccard distances using the R packages ade4 (Chessel et al. 2004) and ape (Paradis 2006).

Optimal group number— The software Structure 2.2 (Pritchard et al. 2000; Falush et al. 2003, 2007) was used to explore optimal group number and the membership of these groups. This software implements a model based clustering algorithm to assign individuals to a parameterized number of groups. This hierarchical Bayesian analysis has been recently extended to dominant markers by treating the second allele at each locus as a parameter to be estimated (Falush et al. 2007). Data was formatted for input into Structure using the software AFLPDAT (Ehrich 2006). Values of *K*, the number of groups, ranged from 2 to 17. In order to generate a mean and standard deviation for each parameterization of *K*, each value of *K* was run sixteen times. The Pr($X|K$) was used to determine convergence of the simulation, when these values appeared to asymptote, the simulation was considered to have converged. The simulation was run with a burnin of 100,000 generations with 30,000 subsequent generations kept. Occasionally a chain would not converge; a phenomenon that appeared to correspond to the magnitude of *K*. These simulations were omitted from further analysis. The decision to run each simulation 16 times resulted in at least 10 converged simulations per value of *K*. The summary statistic delta *K* was calculated according to Evanno et al. (2005). It was observed that at high values of *K*, Structure would return groups which had no members (empty groups). In order to explore this phenomenon a Sunflower plot was created in R (R Development Core Team 2007). Barplots for a single simulation were constructed in R (R Development Core Team 2007). For $K = 6$ and 9 inverse distance weighted spatial interpolations were performed for each group using the R package gstat (Pebesma 2004).

4.3 RESULTS

Assessment of Controls— The controls were not sensitive to increases in bin width above 0.8 standard deviations (Figure 4.2). Analyses utilizing Jaccard (Figure 4.2 A) and mean pair-wise distances (Figure 4.2 B) led to very different results (Table 4.2). For example, *Eco*RI + ACA *Mse*I + GAT has a large distance among controls as measured by a Jaccard distance but has a low distance based on a pair-wise distance. Below a standard deviation of 0.6, Jaccard distances were sensitive to bin width, however not all primer-pairs responded in the same manner. The percentage of ones in the dataset as well as the number of bins increased sharply below a standard deviation of 0.8. The performance of the primer pairs *Eco*RI+AGC *Mse*I+GAT and *Eco*RI+ACG *Mse*I+GAT were poor (Figure 4.2), which corroborates rationale for their removal from the dataset (also see Methods). In order to maximize the percent of ones and the number of bins in the dataset while minimizing the mean Jaccard distance, a bin width of 0.5 standard deviations was chosen for subsequent analyses.

Summary statistics and phenetic analysis— The data set was binned using a maximum bin width of 0.5 standard deviations. Samples with less than three bands for any primer pair were removed. This resulted in a dataset containing 231 individuals scored for 398 bins from four primer pairs (Figure 4.3; Table 4.1) and resulted in a data matrix containing 14.7 percent ones. A neighbor-joining tree was constructed from Jaccard distances (Figure 4.4). Long terminal branches indicate either a large amount of among individual genetic variation or genotyping error. Controls indicate that genotyping error may account for a large proportion of the length of these branches. Nevertheless, there does appear to be internal structure to the tree and varieties largely appear to cluster together.

Optimal group number— The posterior probabilities of the models given the data asymptoted without indicating a clearly optimal parameterization for number of groups (Figure 4.5 A). High values of *K* demonstrated a variance in posterior probability. This did not appear to be due to non-convergence of the model [based on

plots of ln $P(D)$; data not presented] but was due to convergence at different values. Delta K demonstrated a peak at K equals four, indicating an important level of structure at this value of K. As the value of K was increased, the software returned groups which had no members (Figure 4.6). Beginning at $K = 9$, simulations for a parameterized K returned different numbers of non-empty groups. Despite parameterization of group number to $K = 17$, the software never returned more then approxiametly 10 non-empty groups (Figure 4.6). Here the perspective is taken that this software should not be used to infer a single value for K, but rather a range of values. We use the lower bound of $K = 4$, based on the suggestions of Evanno et al. (2005), and an upper bound of $K = 9$, based on the limit of where the software returns consistent results (Figure 4.6).

Barplots of posterior probability of group membership over a range of values for K were created (Figure 4.7). These plots indicate that some groups are highly stable at varying levels of K. For example, *A. l.* var. *variabilis* (Figure 4.7) largely forms a single cohesive group for all levels of K. Other groups appear more variable. For example, *A. l.* var. *borreganus* (red) includes about three to five groups depending on the value of K (Figure 4.7). These groups appear largely to be exclusive to *A. l.* var. *borreganus*. However, at low frequency there are individuals which are clustered with *A. l.* var. *borreganus* which belong to groups that are largely populated by the varieties *coachellae, fremontii* and *variabilis*.

In order to explore the spatial extent of groups inferred from Structure, the membership coefficients were interpolated using the inverse distance weighting algorithm for K equal to six and nine (Figures 4.8 and 4.9). Some groups are fairly stable at different values of K, some vary, and some are a product of increasing K. Group A in Figure 4.8 and 4.9 is relatively stable among values of K and is composed of varieties *fremontii*, *kennedyi* and *nigricalycis*. Group C for K equals six (Figure 4.8) is analogous to group E for K equals nine (Figure 4.9) and is comprised of the variety *variabilis*. Group E for K equals six (Figure 4.8) is very similar to group H for

K equals nine (Figure 4.9) and is largely comprised of the variety *borreganus*. Group F for *K* equals six (Figure 4.8) is largely comprised of Groups F and G for *K* equals nine (Figure 4.9) and include the federally endangered variety *coachellae*. The remaining groups are relatively unstable and change their among population relationships with varying values of *K*.

4.4 DISCUSSION

Technological advances have allowed contemporary research to steadily increase study complexity in terms of sample size, marker number and analytical methods (including parameter number to be estimated by these methods). As the level of complexity increases, the potential to encounter pitfalls in experimental implementation also increases. This can be manifested in collection or scoring of data error (e.g., genotyping error) or in the implementation and interpretation of complex algorithms (i.e., understanding explicit and implicit assumptions). It is critical that these issues be explored during the interpretation of a particular dataset prior to extrapolating results to the biological system of interest. The present investigation has opened new insights into the management of genotyping error in large datasets and the interpretation of clustering analyses, as well as the inference of groupings within the southern Californian varieties of *A. lentiginosus*.

Managing genotyping error— As datasets increase in size in terms of samples or markers (characters), the practicality of scrutinizing each data point becomes increasingly impractical. Error in amplified fragment length polymorphism (AFLP) datasets may include false positives and false negatives. False positives can be the result of two types: spurious peaks and sizing ambiguity. False negatives can result from an arbitrary choice of minimum relative fluorescence units chosen to call peaks.

Spurious peaks arise due to errors in the peak calling process. Whitlock et al. (2008) reported an average of 25.5 peaks per sample above a threshold of 50 relative fluorescence units resulting from the genotyping of water on an ABI3730. Because

genetic fingerprinting methods lack an expectation (i.e., a microsatellite project has an expectation to find an allele within a range of base pairs) this sort of false positive may be mistakenly scored as a band. Because this band's origin is not from the genome of interest, it would result in misleading data. Whitlock et al. (2008) report that raising the arbitrary threshold for peak calling reduced this problem; however this undoubtedly comes at the cost of excluding potentially informative peaks.

Sizing ambiguity is the product of error associated with the estimate of a fragment's molecular weight. Imagine a set of DNA fragments located at 200 and 201 base pairs. Error in correctly sizing peaks in these bins will lead to a variance around the actual size. As sample size for these peaks increases, this error may become large enough (e.g., standard deviations equal to 0.25 base pairs) that a continuum of peaks is observed between 200 and 201 base pairs. When this continuum spreads over several base pairs, either due to a large error around each bin or several closely spaced bins, it results in a set of ambiguous peaks which cannot be confidently assigned to a single bin. A curious property of this type of error is that it may not be apparent in small sample sizes. It is also important to note that this type of error will only become evident after binning of the dataset has occurred, making the possibility of pre-emptive management low. Here I have chosen to omit bins that span several base pairs.

False negatives may be present in AFLP datasets due to the arbitrary peak height threshold used for determining band presence or absence. Peaks present at lower values than this threshold will be scored as absent. This is not simply a factor or user specification, as a population of low peak heights may span a machine's limit of detectability (a minimum of 50 rfus on ABI machines). If there was an expectation for all peaks within a sample and among samples to fluoresce at equal intensities, this would be a problem of standardizing concentrations. Practical experience suggests great variability in peak intensity among samples and among peaks within samples, even when initial template DNA concentration is standardized. Uniformly low peaks within a particular sample may result from the uniformly poor amplification of a

sample, possibly due to low initial DNA concentration, poor initial DNA quality or over dilution of an AFLP product prior to genotyping. Lack of uniformity in peak height within a sample may be due to PCR competition during the AFLP chemistry which results in non-uniform production of amplicon or possible bias in electrokinetic loading. There does not appear to be a satisfactory manner to manage this type of error. Here I have excluded samples which produce less than three bands for any primer pair as representative of either poor quality or quantity samples.

Some researchers have implemented control samples to assess genotyping error and have reported a pair-wise distance based on the recommendations of Bonin et al. (2004). The reporting of pair-wise distances for AFLP controls is not an appropriate measure of reproducibility (Table 4.2). This is largely due to the binning process, which is usually performed on the entire dataset. As increasingly divergent samples are included, it becomes more likely that bands in samples other than the control will become present. In the control these will be scored as a shared absence and will decrease the pair-wise distance among the controls. It is important to note that Bonin et al. (2004) addressed AFLP and microsatellite data; a pair-wise distance would be appropriate for the latter data type. Distance measures which do not consider shared absences (e.g., Dice, Jaccard or Sørenson distances) are frequently recommended for the analysis of AFLP datasets in the literature (Nei and Li 1979; Lynch 1990; Wolfe and Liston 1998; Koopman and Gort 2004; Bonin et al. 2007). I recommend the use of distances which exclude the comparison of shared absences when reporting the performance of controls. Sometimes a pair-wise distance may be desirable due to ease of interpretation; this should only be reported if shared absences are removed from the controls prior to calculation of this type of distance (and it should be clearly stated that this has been done).

Ultimately error generated in the AFLP process may be due to initial DNA concentration and quality (Bensch and Åkesson 2005; Bonin et al. 2007). These values, particularly DNA quality, may vary largely from taxon to taxon. It is therefore

highly recommended that any study considering the use of AFLP should perform a pilot study employing many controls (e.g., 6 - 10). Furthermore it is recommended that any study implementing AFLPs include internal controls to facilitate the reporting of error rates. Complete randomization of an experiment may introduce error caused by inaccurate plating of DNA and is not advocated here. Randomization or dispersion of samples in units of columns, where each eight sample column contains members of the same population, may provide an acceptable balance between organization and randomization. Columns of populations can then be distributed among several plates and among plate error can be tested for in a hierarchical manner. If among individual comparisons are desired, a column of control individuals from as many populations as possible should also be replicated on each plate of the experiment.

Without knowledge of bias in AFLP error it may be assumed that error is uniformly distributed throughout the dataset (i.e., it is equally likely to affect all samples). Due to their scoring as presence or absence, AFLP datasets are largely empty (i.e., they contain more zeros than ones). Koopman et al. (2008) report a dataset for the genus *Rosa* (Rosaceae) which is composed of 23.7% ones (dataset available at http://treebase.org). The present dataset contains 14.7% ones. It appears reasonable to assume that if error affected either of these datasets (which it likely has), it is more probable that this error has led to the erosion of signal (ones being erroneously called as zeros) rather than to erroneous grouping (i.e., shared false positives). Within this context it seems that the large 'signal to noise ratio' argument for AFLPs is justified. Focus of research questions may also alleviate the issue of genotyping error. Genotyping error will be manifested in datasets as among individual estimates of variance in an ANOVA type context or as long terminal branches in a dendrogram based analysis. While the among individual component of variance may be confounded by error, if the question of interest is at the population or higher level, genotyping error may be tolerated as long as there is adequate signal at these higher levels. The use of clustering algorithms may be of particular use in the case of genotyping error due to their ability to recognize groups despite relatively low levels

of signal. Dionne et al. (2008) used microsatellites (a less error prone data type due to expected allele sizes) to infer structure in Atlantic salmon using the Bayesian clustering software Geneland (Guillot et al. 2005) where 95.44% of the variation in the dataset was among individuals within populations. Similarly, Rosenberg et al. (2002) used microsatellites and the software Structure to infer groups in worldwide human populations whose within population variance ranged from 88.4 to 99.3% (several groupings were employed). These studies indicate that hierarchical Bayesian clustering algorithms are successful at detecting population structure even when a very low proportion of marker variance explains this structure.

Interpretation of clustering results— Determination of the optimal number of groups contained in a sample (i.e., untrained clustering) is a common research question. The software Structure (Pritchard et al. 2000; Falush et al. 2003, 2007), while not specifically intended to address this question, provides an ad hoc assessment of optimal number of groups via optimizing log P($X|K$), the probability of the genotypes given the number of groups. This frequently results in a plot which asymptotes without a clearly 'best' model (Figure 4.5A; Evanno et al. 2005 Figure 2A). That is to say that a 'most likely' model may be chosen, yet it may not be clear how much better this model is than the second or third most likely models. Evanno et al. (2005) make two important observations. First they recommend using the change in slope of this plot to infer 'major' structure in the dataset. This appears to be a formalization of suggestions from other clustering methods (Everitt 2005). Second, they observe that at large values of K (the number of groups) the model tends to converge at different values.

The phenomenon of convergence at different values has been observed elsewhere (Rosenberg et al. 2002). This is likely due to the model encountering local optima, an event that appears to become more prevalent as the complexity of the model increases (i.e., the value of K increases). The result appears to be model convergence; however, convergence may occur at different values (local optima) in different simulations.

Here it is reported that as *K* increases past a critical point, the Structure algorithm begins to return groups which lack members (empty groups; Figure 4.6). Ideally, only data from the most likely simulations should be interpreted, and less likely simulations should be discarded as local optima. However, as variability among simulations increases it may be difficult to determine which are the result of local optima and which is the 'correct' simulation, assuming that at least one simulation has found the correct answer. This emphasizes the importance of running multiple chains per parameterization to ensure one chain has not found a local optimum. In general it is computationally more efficient to run several chains in parallel than several chains in series. Some current implementations of MCMC implement multiple parallel chains to address other sorts of questions (Lunn et al. 2000; Ronquist & Huelsenbeck 2003). An important advancement in hierarchical Bayesian clustering would be the implementation of parallel chains by software developers. Until then the user is encouraged to run multiple chains in series (i.e., one after the other) to ensure convergence may not be at a local optimum.

The present study resulted in groups which lacked members at high values of *K* (Figure 4.6). While P(*X*|*K*) appears to asymptote (Figure 4.5A) the number of non-empty groups appears to reach a clear point of instability (*K* = 9, Figure 4.6). This appears to provide an easy to interpret upper bound on relevant values of *K*. There remains a technical issue of whether higher values of *K* may be appropriate but technically challenging to explore in parameter rich models. It does seem clear that there is a practical upper limit with which to interpret appropriate values of *K* given the present technology. Here it is advocated that several values of *K* should be interpreted (see Rosenberg et al. 2002) using the recommendation of Evanno et al. (2005) as a lower limit and the appearance of empty groups as an upper limit.

***Groupings within the southern California* varieties of *Astragalus lentiginosus*—** Cluster analysis of AFLP phenotypes for southern California varieties of *Astragalus lentiginosus* appear to result in at least six stable groups (Figure 4.7 D). As the

parameterized number of groups increases beyond this value, the groups appear to be subdivisions of groups occurring at *K* equals six, as opposed to the appearance of novel groupings, indicating a level of stability of group assignment above *K* equals six (Figure 4.7). After the number of groups is increased beyond nine, the MCMC chains begin to converge at different values of *K*, indicating instability in the optimal number of groups (Figure 4.6). We interpret values of *K* equal to six (Figs. 4.7D and 4.8) as a lower bound and *K* equal to nine (Figs. 4.7G and 4.9) as an upper bound for group number. These groupings largely track current taxonomy (Barneby 1964) yet some varieties appear to contain multiple genotypic groups, and some varieties are included within the genotypic groups of other varieties. An interpretation of AFLP phenotypic groups (Figure 4.7) within the framework of current taxonomy (Barneby 1964) is provided.

var. *borreganus*— The variety *borreganus* has been described as consisting of a disjunct distribution (Barneby 1945, 1964; Welsh 2007) occurring on dunes first in eastern San Diego and Imperial counties, California into adjacent Arizona and secondly in central San Bernardino county north into extreme southern Nevada. This disjunction is separated by vars. *coachellae* and *variabilis*. Sampled populations of this taxon occur in three (Fig 4.8 A, D, E) to four (Figure 4.9 A, B, H, I) clusters. The northern population, from San Bernardino County (Kelso Dunes), occurs in two stable clusters (Figure 4.8A and E; Fig 4.9 A and H), while the southern population, from San Diego County (Font's Wash), occurs in two (Fig 4.8 D and E) to three clusters (Figure 4.9 B, H, I). This is particularly interesting because the population from Font's Wash was propagated from seed provided by the Desert Legume Program. Consequently, there was no a prior expectation of how many maternal parents this seed represented. Our results indicate it to be more diverse than the population at Kelso Dunes. While this taxon appears to occupy two clusters unique to var. *borreganus*, it also appears to share rare affinities with the geographically proximal var. *coachellae* (Figure 4.8 E; Figure 4.9 H, I), as well as the geographically distant var. *fremontii* (Figure 4.8 E; Figure 4.9 B). While the shared cluster with the

geographically proximal var. *coachellae* may be explained by some level of recent gene-flow, the shared cluster with the geographically distant var. *fremontii* is more difficult to explain, and may be the result of more complex processes such as incomplete lineage sorting (Knaus Chapter 3).

var. *coachellae*—The variety *coachellae* is endemic to the Coachella Valley, Riverside County, CA (Barneby 1964; Welsh 2007). This variety is federally listed as endangered and receives protection under the Endangered Species Act (USFWS 1998). Despite a geographical distribution which is small relative to other taxa within *A. lentiginosus*, the observed molecular diversity within *A. l.* var. *coachellae* is relatively large. This taxon occurs in three (Figure 4.8 B, E, F) to four (Figure 4.9 F, G, H, I) clusters with several of the clusters largely unique to var. *coachellae* (Figure 4.8 F and Figure 4.9 F and G). Other clusters are shared with either geographically distant populations of var. *fremontii* (Figure 4.8 B) at low parameterizations of *K* or with geographically proximal populations of var. *borreganus* (Figure 4.9 H and I) at high parameterizations of *K*. This may be indicative of ancestral relationships among the variety *coachellae* and the varieties *fremontii* and *borreganus* (i.e., incomplete lineage sorting; Knaus Chapter 3; Knaus et al. 2008) or due to recent gene flow. It is notable that the populations of *A. lentiginosus* at eastern Riverside County (Desert Center) and southern San Bernardino County (Joshua Tree National Park), while geographically proximal to the populations within the Coachella Valley, are dramatically different and cluster as such (Figure 4.8 B, C, F; Figure 4.9 E, F, G, H, I). Depending on how *K* is parameterized, var. *coachellae* may have as many as two clusters unique to it, but it also shares clusters with the varieties *borreganus* and *fremontii* (however, these are relatively infrequent). The relatively high degree of genetic diversity observed in var. *coachellae* is also observed by chloroplast simple sequence repeat data (Knaus Chapter 3). A preponderance of data suggests that *A. l.* var. *coachellae* represents an unusually large amount of genetic diversity within the species *A. lentiginosus*, which perhaps justifies protection as a reservoir of genetic diversity for the taxon.

var. *fremontii*— The variety *fremontii*, being relatively common, is among the best sampled varieties in the dataset (Table 4.1). It forms a reasonably good group consisting of three (Figure 4.8 A, B, D) to four (Figure 4.9 A, B, C, and D) clusters. This group contains several genetic clusters, indicating relatively high genetic diversity within this taxon, a finding consistent with chloroplast data (Knaus Chapter 3). Although this may be in part due to its large representation in the sample (Table 4.1).

An important pattern occurs at approximately 36.5° north. This portion of the sample forms an east to west transect through a series of basins and ranges from the Sierra Nevada Mountains eastward into Nevada and the Great Basin. This transect includes three populations of variety *fremontii* at higher elevations and three populations of variety *variabilis* at lower elevations. Here clusters seem to correspond to taxonomy despite a complex geographic arrangement (Figs. 4.8A, 4.8C, 4.9A, 4.9C, and 4.9E). This suggests that observed patterns of genetic diversity do not reflect a simple pattern of isolation by distance, but may be indicative of differential gene flow, which reflects morphological taxonomy. This differential gene flow may be due to barriers to reproduction (allopatry or physiological incompatibility). Another possible explanation is that alleles in the sample which do not reflect ancestry but may reflect selective pressures (Lewontin and Krakauer 1973; Wu 2001; Via 2002; Beaumont and Balding 2004; Beaumont 2005).

The control population from eastern San Bernardino County appears relatively distinct within the dataset due to its clustering predominantly by itself (Figure 4.8 D; Figure 4.9 D). It appears to have slight affinities with var. *borreganus* (Figure 4.8 D) or var. *micans* (Figure 4.9 D), representing two geographically distant populations and varieties. Results from Chloroplast data indicate that the greatest amount of genetic differentiation within *A. lentiginosus* is explained by the among population component

of variance (Knaus Chapter 3). It may be that the control population is simply a relatively unique population.

var. *kennedyi*— The variety *kennedyi* is from the Carson Sink region of Nevada and represents the northernmost populations sampled. This variety occurs in two clusters (Figure 4.8 A, E; Figure 4.9 A, B); however, at low parameterizations of *K* it predominantly belongs to one cluster, while at high parameterization of *K* the two sampled populations appear relatively distinct. Taxonomic affinities of this taxon are with var. *fremontii* (Barneby 1964), a hypothesis supported by either grouping, yet this taxon also appears to harbor its own share of relatively unique genetic diversity (Figure 4.9 B).

var. *micans*— The variety *micans* is a narrow endemic which is morphologically most similar to the varieties *variabilis* and *coachellae*, yet it appears to share AFLP affinities with variety *fremontii* (Figure 4.8 B; Figure 4.9 D). As a population it forms a cohesive group, yet differing parameterizations of *K* groups this population within clusters containing different populations of var. *fremontii*. At low parameterization of *K*, the variety *micans* clusters with a geographically proximal population of var. *fremontii* (Figure 4.8 B), from the nearby Sierra Nevada Mountains. At high parameterization of *K*, the variety *micans* clusters with the control population (see above) of var. *fremontii*, which occurs in San Bernardino County, CA and is geographically distant from the Eureka Valley. Shifting alliances among var. *micans* and the other taxa can also be seen in Figure 4.7 where at very high parameterization of *K* the variety appears in a unique group (Figure 4.7 I and J). It may be best to view the variety *micans* as a relatively unique population among the sampled populations of *A. lentiginosus* (see discussion of the control population above) which possesses tenuous relationships with several other populations.

var. *nigricalycis*— The variety *nigricalycis* is endemic to the San Juaquin Valley of California and is one of the few varieties of *A. lentiginosus* to be distributed beyond

the intermountain deserts (Barneby 1964). This variety is represented by a single population in the sample which falls within one cluster (Figure 4.8 A; Figure 4.9 A). The affinities of this variety are with geographically distant populations of var. *kennedyi* or var. *fremontii*. This result is unexpected but is perhaps best interpreted as the retention of ancestral alleles which have yet to sort themselves into lineages represented by current taxonomy (assuming taxonomy reflects biological phenomena).

var. *variabilis*— The variety *variabilis* forms a cohesive group at all parameterizations of *K* (Figs. 4.7, 4.8C, and 4.9E), with the exception of one population occurring in the Mojave National Preserve in eastern San Bernardino county. This genetic uniformity is particularly interesting because this taxon has been considered so variable in morphology that taxonomic forms have been informally presented (Barneby 1945). The populations of *A. lentiginosus* at Joshua Tree National Park and Desert Center, Riverside County, CA cluster within *A. l.* var. *variabilis*, supporting the hypothesis that they are different from the populations of *A. lentiginosus* in the Coachella Valley which have been recognized taxonomically as var. *coachellae*. One population of var. *variabilis* from the Mojave Preserve appears to not group with the rest of the variety and instead clusters with populations of var. *fremontii* (Figs 4.7, 4.8A and 4.9A). It is notable that this population does not cluster with geographically proximal populations of other varieties (vars. *borreganus* or *fremontii*) either. This pattern is absent in the previously reported chloroplast simple sequence repeat dataset (Knaus Chapter 3, Figure 3.6). This may be due to the relative uniqueness of the geographically proximal population of var. *fremontii* (Figs. 4.8D and 4.9D), as the Mojave Preserve population does appear to cluster with other populations of var. *fremontii*.

Conclusions — Amplified fragment length polymorphisms have a mixed reputation as a molecular tool for population and lineage inference due to issues of genotyping error. This is frequently considered to be overcome through an assumed high signal to noise ratio present in AFLP datasets. Here I report that previous studies which have made an honest attempt to characterize the level of error in their datasets through the

inclusion of controls may not have accurately presented the performance of their controls through the presentation of a pair-wise distance (Bonin et al. 2004). It is recommended that researchers present distances which ignore shared absences to report error, just as they would to analyze their datasets (i.e., Dice, Jaccard, or Sørenson distances). Substantial error appears to be present in the current dataset. However, by making the assumption that this error is uniformly distributed it can be assumed that this error confounds primarily among individual inferences. By focusing attention on hierarchical levels of among population and higher hierarchical levels this error can be circumvented. Regions of relatively high genetic diversity occur in the Coachellae Valley, as well as in the more northern populations in the sample, results which correspond to other molecular data (Knaus Chapter 3) and the isozyme results of Liston (1992) for *Astragalus* species with similar widespread distributions in the Mojave and Sonoran deserts. The mid-Mojave Desert, populated largely by var. *variabilis*, contains a single relatively uniform group which appears consistent with other molecular data (Knaus Chapter 3). This result is curious in that morphological diversity appears so great within this taxon that the recognition of taxonomic forms has been discussed (Barneby 1945). An east-west transect at approximately 36.5° north, demonstrating a geographically and taxonomically complex pattern, appears to support taxonomy over isolation by distance, a phenomenon which may be indicative of selective response to environmental factors within *A. lentiginosus*.

4.5 LITERATURE CITED

Amsellem, L., J. L. Noyer, T. Le Bourgeois, and M. Hossaert-McKey. 2000. Comparison of genetic diversity of the invasive weed *Rubus alceifolius* Poir. (Rosaceae) in its native range and in areas of introduction, using amplified fragment length polymorphism (AFLP) markers. Molecular Ecology 9:443-455.

Barneby, R. C. 1945. *Pugillus Astragalorum* IV: the section *Diplocystium*. Leaflets of Western Botany 4:65-147.

—. 1964. Atlas of North American *Astragalus*. Memoirs of the New York Botanical Garden 13:1-1188.

Beaumont, M. A. 2005. Adaptation and speciation: what can F_{ST} tell us? Trends in Ecology & Evolution 20:435-440.

Beaumont, M. A., and D. J. Balding. 2004. Identifying adaptive genetic divergence among populations from genome scans. Molecular Ecology 13:969-980.
Bensch, S., and M. Åkesson. 2005. Ten years of AFLP in ecology and evolution: why so few animals? Molecular Ecology 14:2899-2914.
Bonin, A., E. Bellemain, P. Bronken Eidesen, F. Pompanon, C. Brochmann, and P. Taberlet. 2004. How to track and assess genotyping errors in population genetics studies. MJolecular Ecology 13:3261-3273.
Bonin, A., D. Ehrich, and S. Manel. 2007. Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionsits. Molecular Ecology 16:3737-3758.
Chessel, D., A.-B. Dufour, and J. Thioulouse. 2004. The ade4 package. R News 4:5-10.
Chung, M., G. Gelembuick, and T. J. Givnish. 2004. Population genetics and phylogeography of endangered *Oxytropis campestris* var. *chartacea* and relatives: artic-alpine disjuncts in eastern North America. Molecular Ecology 13:3657-3673.
Corander, J., and P. Marttinen. 2006. Bayesian identification of admixture events using multilocus molecular markers. Molecular Ecology 15:2833-2843.
Coyne, J. A., and H. A. Orr. 2004, Speciation. Sunderland, Mass., Sinauer Associates.
Darwin, C. 1859, On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. *In* D.M. Porter and P.W. Graham *eds.* The Portable Darwin, 1993, Penguin Books.
Dionne, M., F. Caron, J. J. Dodson, and L. Bernatchez. 2008. Landscape genetics and hierarchical structure in Atlantic salmon: the interaction of gene flow and local adaptation. Molecular Ecology 17:2382-2396.
Dobzhansky, T. 1951, Genetics and the Origin of Species. New York, Columbia University Press.
Erich, D. 2006. AFLPDAT: a collection of R functions for convenient handling of AFLP data. Molecular Ecology Notes 6:603-604.
Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611-2620.
Everitt, B. 2005, An R and S-plus companion to multivariate analysis. London, Springer.
Falush, D., M. Stephens, and J. K. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567-1587.
—. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular Ecology Notes 7:574-578.
Fisher, R. A. 1958, The Genetical Theory of Natural Selection, Dover.
Guillot, G., A. Estoup, F. Mortier, and J. F. Cosson. 2005. A spatial statistical model for landscape genetics. Genetics 170:1261-1280.
Guillot, G., F. Mortier, and A. Estoup. 2005. Geneland: a program for landscape genetics. Molecular Ecology Notes 5:712-715.

Hudson, R. R., and J. A. Coyne. 2002. Mathematical consequences of the genealogical species concept. Evolution 56:1557-1565.

Jones, C. J., K. J. Edwards, S. Castaglione, M. O. Winfield, F. Sala, C. Van de Wiel, G. Bredemeijer et al. 1997. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Molecular Breeding 3:381-390.

Jones, C. J., K. J. Edwards, S. Castaglione, M. O. Winfield, F. Sala, C. Van der Wiel, B. Vosman et al. 1998. Reproducibility testing of AFLPs by a network of European laboratories *in* A. Karp, P. G. Isaac, and D. S. Ingram, eds. Molecular Tools for Screening Biodiversity. London, Chapman and Hall.

Knaus, B. J., R. C. Cronn, and A. Liston. 2005. Genetic characterization of three varieties of *Astragalus lentiginosus* (Fabaceae). Brittonia 57:334-344.

—. 2008. A fistful of *Astragalus*: incipient speciation in the American West? Presentation at Botany 2008.

Koopman, W. J. M., and G. Gort. 2004. Significance tests and weighted values for AFLP similarities, based on *Arabidopsis* in silico AFLP fragment length distributions. Genetics 167:1915-1928.

Koopman, W. J. M., V. Wissemann, K. De Cock, J. Van Huylenbroeck, J. De Riek, G. J. H. Sabatino, D. Visser et al. 2008. AFLP markers as a tool to reconstruct complex relationships: A case study in *Rosa* (Rosaceae). American Journal of Botany 95: 353-366.

Kropf, M., J. W. Kadereit, and H. P. Comes. 2002. Late Quaternary distributional stasis in the submediterranean mountain plant *Anthyllis montana* L. (Fabaceae) inferred from ITS sequences and amplified fragment length polymorphisms markers. Molecular Ecology 11:447-463.

Lewontin, R. C., and J. Krakauer. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. Genetics 74:175-195.

Liston, A. 1992. Isozyme Systematics of *Astragalus* Sect. *Leptocarpi* Subsect. *Californici* Fabaceae. Systematic Botany 17:367-379.

Lock, J. M., and B. D. Schrire. 2005. Galegeae, Pages 475-487 *in* G. Lewis, B. Schrire, B. MacKinder, and M. Lock, eds. Legumes of the World. Kew, Royal Botanical Gardens.

Lunn, D. J., A. Thomas, N. Best, and D. Spiegelhalter. 2000. WinBUGS -- a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing 10:325-337.

Lynch, M. 1990. The similarity index and DNA fingerprinting. Molecular Biology and Evolution 7:478-484.

Mabberley, D. J. 1987, The plant-book : a portable dictionary of the higher plants utilising Cronquist's An integrated system of classification of flowering plants (1981) and current botanical literature arranged largely on the principles of editions 1-6 (1896/97-1931) of Willis's A dictionary of the flowering plants and ferns. New York, Cambridge University Press.

Mayden, R. L. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem, Pages 381-422 *in* M. F. Claridge, H. A. Dawah, and M. R. Wilson, eds. Species: The Units of Biodiversity, Chapman & Hall.

Mayr, E., and P. D. Ashlock. 1991, Principles of Systematic Zoology, McGraw-Hill.
Morjan, C. L., and L. H. Rieseberg. 2004. How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. Molecular Ecology 13:1341-1356.
Nei, M., and W.-H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction nucleases. Proceedings of the National Academy of Science, U.S.A. 76:5269-5273.
Nordborg, M. 2001. Coalescent theory, Pages 179-212 *in* D. J. Balding, M. J. Bishop, and C. Canings, eds. Handbook of Statistical Genetics. Chichester, UK, John Wiley and Sons.
Orr, H. A. 2001. The genetics of species differences. Trends in Ecology & Evolution 16:343-350.
Paradis, E. 2006, Analysis of phylogenetics and evolution with R. New York, Springer.
Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289-290.
Pebesma, E. J. 2004. Multivariable geostatistics in S: the gstat package. Computers & Geosciences 30:383-691.
Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.
R Development Core Team. 2007.R: a language and environment for statistical computing.R foundation for statistical computing, Vienna, Austria.
Rieseberg, L. H., and J. M. Burke. 2001. The biological reality of species: Gene flow, selection, and collective evolution. Taxon 50:47-67.
Rieseberg, L. H., T. E. Wood, and E. J. Baack. 2006. The nature of plant species. Nature 440:524-527.
Ronquist, F., and J. P. Huelsenbeck. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.
Rosenberg, N. A., and M. Nordborg. 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. Nature Reviews Genetics 3:380-390.
Rosenberg, N. A., J. K. Pritchard, J. L. Weber, H. M. Cann, K. K. Kidd, L. A. Zhivotovsky, and M. W. Feldman. 2002. Genetic structure of human populations. Science 298:2381-2385.
Rouppe van der Voort, J. N. A. M., P. van Zandvoort, H. J. van Eck, R. T. Folkertsma, R. C. B. Hutten, J. Draaistra, F. J. Gommers et al. 1997. Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. Molecular and General Genetics 255:438-447.
Rydberg, P. A. 1929. *Astragalanae*. North American Flora 24:251-462.
Schönswetter, P., A. Tribsch, and H. Nikfeld. 2004. Amplified fragment length polymorphism (AFLP) reveals no genetic divergence of the Eastern Alpine endemic *Oxytropis campestris* subsp. *tiroliensis* (Fabaceae) from widespread subsp. *campestris*. Plant Systematics and Evolution 244:245-255.
Syring, J., K. Farrell, R. Businský, R. Cronn, and A. Liston. 2007. Widespread genealogical nonmonophyly in species of *Pinus* subgenus *Strobus*. Systematic Biology 56:163-181.

Travis, S. E., J. Maschinski, and P. Keim. 1996. An analysis of genetic variation in *Astragalus cremnophylax* var. *cremnophylax*, a critically-endangered plant, using AFLP markers. Molecular Ecology 5:735-745.

USFWS. 1998. Endangered and threatened wildlife and plants; determination of endangered or threatened status for five desert milk-vetch taxa from California. Federal Register 63:53596-53615.

Via, S. 2002. The ecological genetics of speciation. The American Naturalist 159:S1-S7.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters et al. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23:4407-4414.

Welsh, S. L. 2007, North American species of *Astragalus* Linnaeus (Leguminosae), a taxonomic revision. Prove, UT, Monte L. Bean Life Science Museum.

Whitlock, R., H. Hipperson, M. Mannarelli, R. K. Butlin, and T. Burke. 2008. An objective, rapid and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. Molecular Ecology Resources 8:725-735.

Wojciechowski, M. F., M. J. Sanderson, B. G. Baldwin, and M. J. Donoghue. 1993. Monophyly of aneuploid *Astragalus* (Fabaceae): evidence from nuclear ribosomal DNA internal transcribed spacer sequences. American Journal of Botany 80:711-722.

Wojciechowski, M. F., M. J. Sanderson, and J.-M. Hu. 1999. Evidence on the monophyly of *Astragalus* (Fabaceae) and its major subgroups based on nuclear ribosomal DNA ITS and chloroplast DNA *trnL* intron data. Systematic Botany 24:409-437.

Wolfe, A. D., and A. Liston. 1998. Contributions of PCR-based methods to plant systematics and evolutionary biology *in* D. E. Soltis, P. S. Soltis, and J. J. Doyle, eds. Molecular Systematics of Plants II: DNA Sequencing. Boston, MA, Kluwer Academic.

Wright, S. 1978, Evolution and the genetics of populations. Volume 4: variability within and among natural populations. Chicago, University of Chicago Press.

Wu, C.-I. 2001. The genic view of the process of speciation. Journal of Evolutionary Biology 14:851-865.

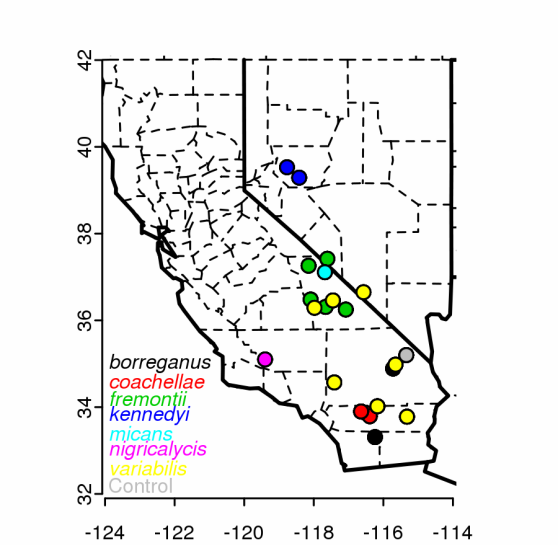


Figure 4.1. Map of the AFLP sample. Spatial extent includes the states of California and Nevada, U.S.A.

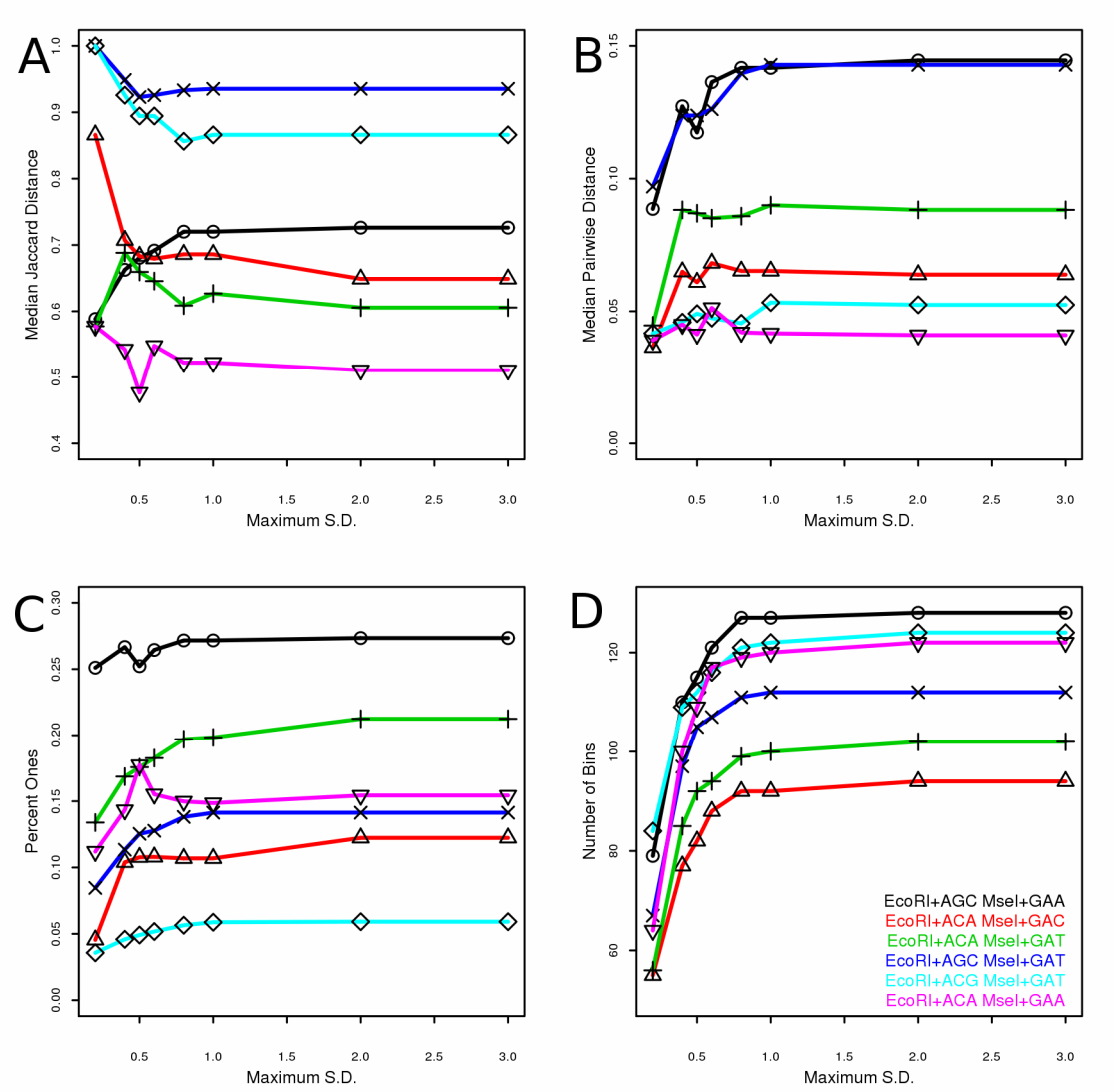


Figure 4.2. Summary of AFLP controls. Independent variable is represented by the maximum width (in standard deviations) allowed for bins.

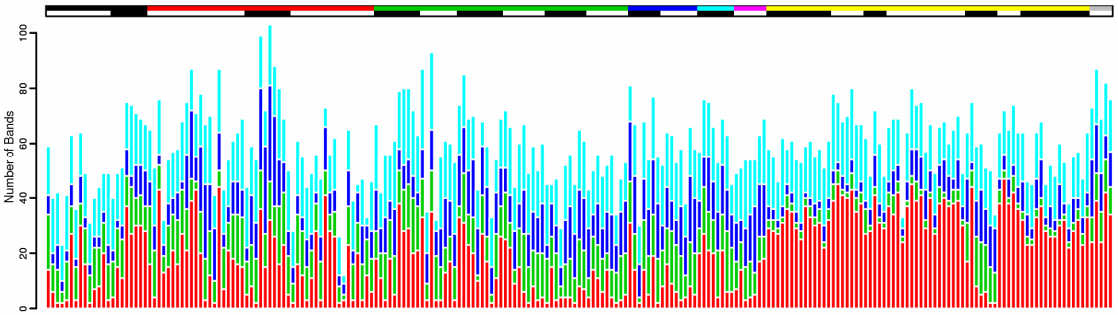


Figure 4.3. Barplot of AFLP band occurance. Samples are represented by vertical bars with color proportional to observed number of bands for each of four primer pairs. The first horizontal row above y=1 represents populations (black and white). The second row above y=1 represents varieties (color follows Figure 4.1).

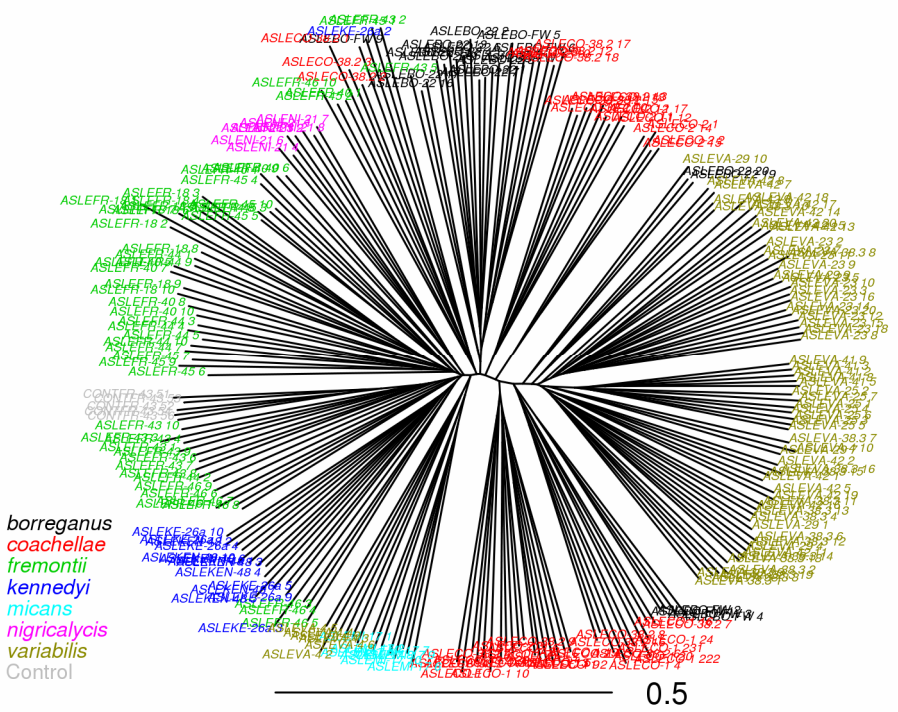


Figure 4.4. Neighbor-joining tree based on AFLP bands. Tree was created using a Jaccard distance based on four primer pairs, 231 samples and 398 bins. Color follows Figure 4.1.

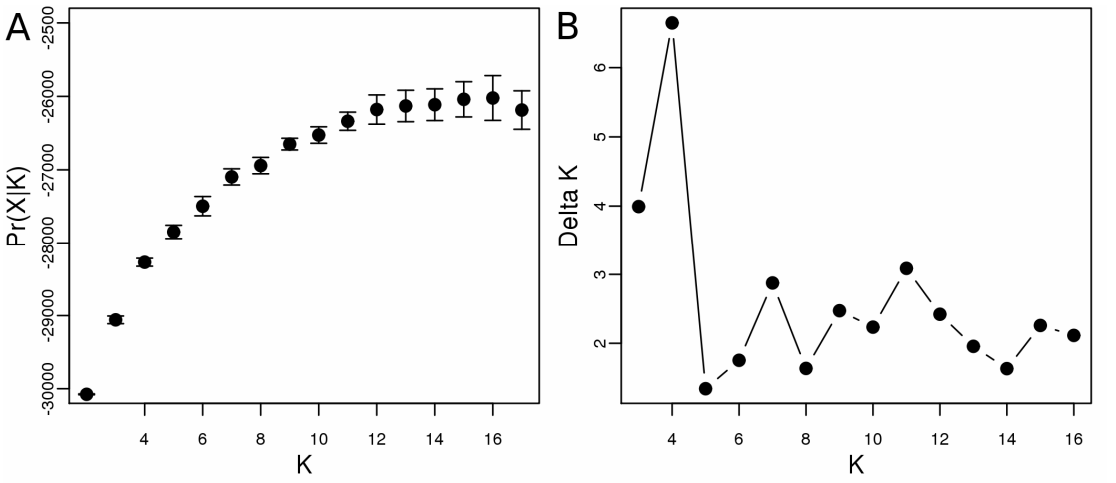


Figure 4.5. Diagnostics for determining optimal group number. Panel A is log probability of the genotypes given the number of groups as a function of the number of groups (*K*). Error bars represent one standard deviation. Panel B is the delta *K* of Evanno et al. (2005).

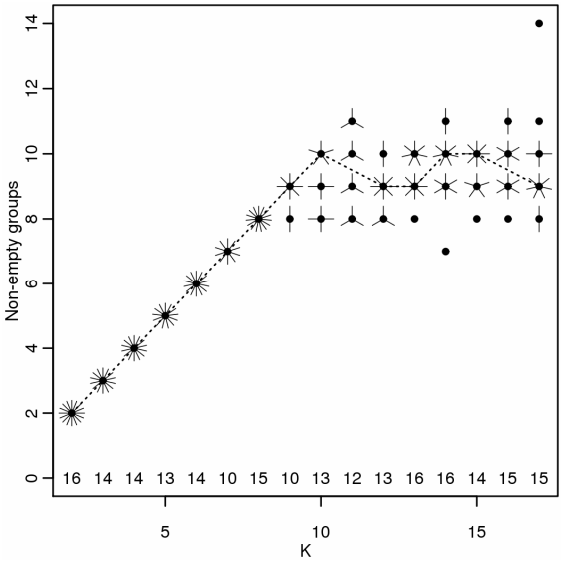
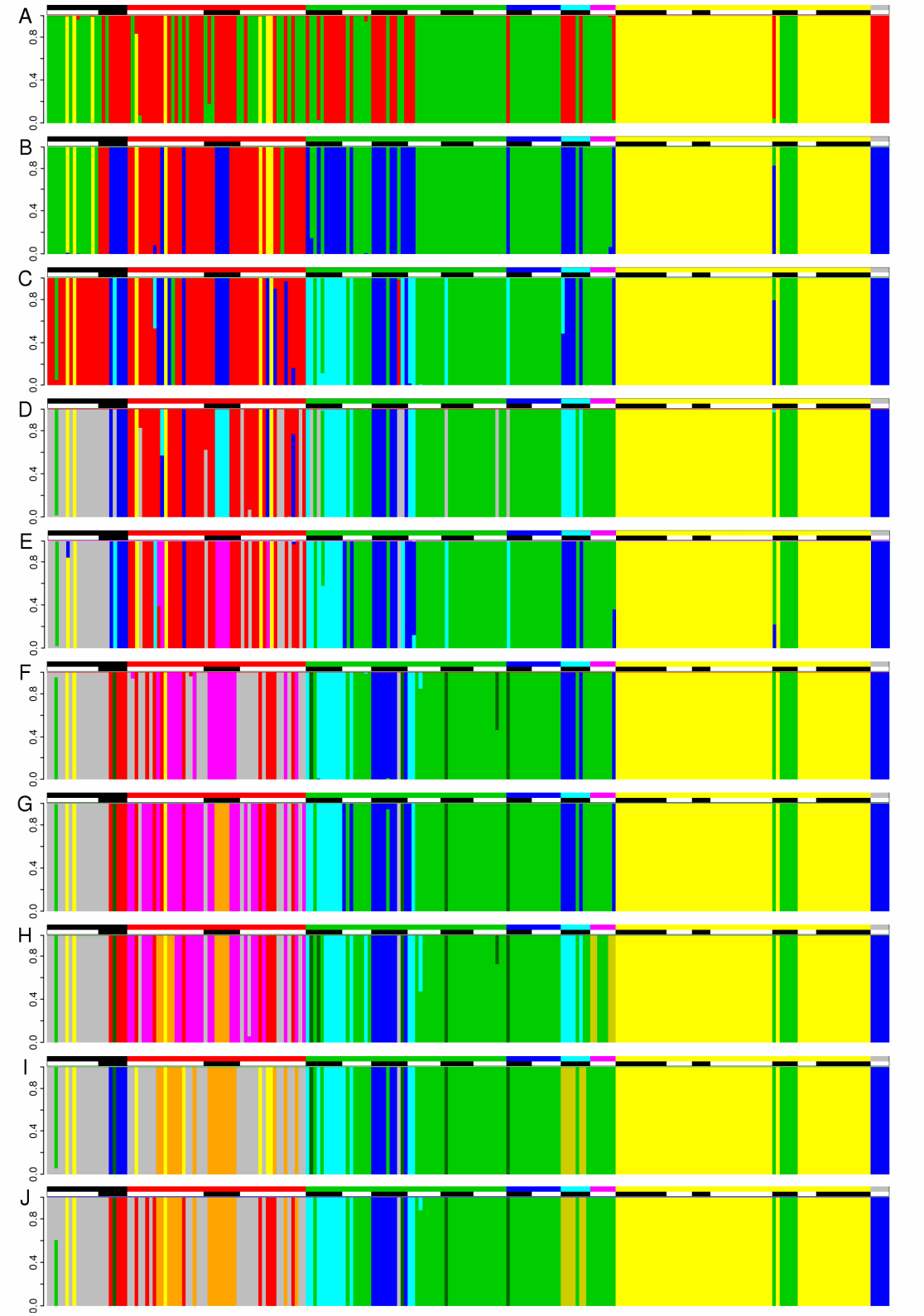


Figure 4.6. Sunflower plot of non-empty groups. 'Rays' represent the number of observations at each point when there are more observations than one. Numbers at y = 0 represent the number of simulations for each value of *K* (simulations which did not appear to converge were omitted).

Figure 4.7. Barplots of probability of group membership for 10 values of *K*. Number of groups (*K*) equals 3 (A) through 12 (J). The first horizontal row above y=1 represents populations (black and white). The second row above y=1 represents varieties (color, follows Figure 4.1).


A
B
C
D
E
F
G
H
I
J
0.0
0.4
0.8

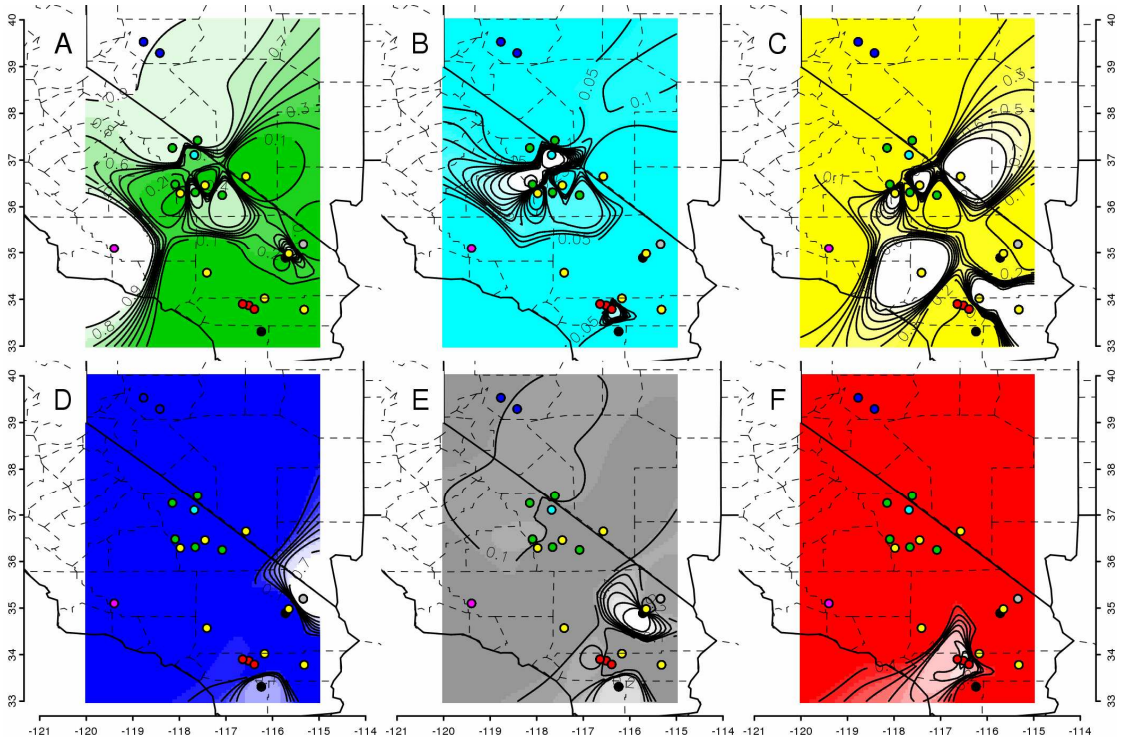


Figure 4.8. IDW interpolations of probability of group membership for $K = 6$. White represents high probability of membership and color represents low probability of membership. Colors correspond to the barplots of Figure 4.7.

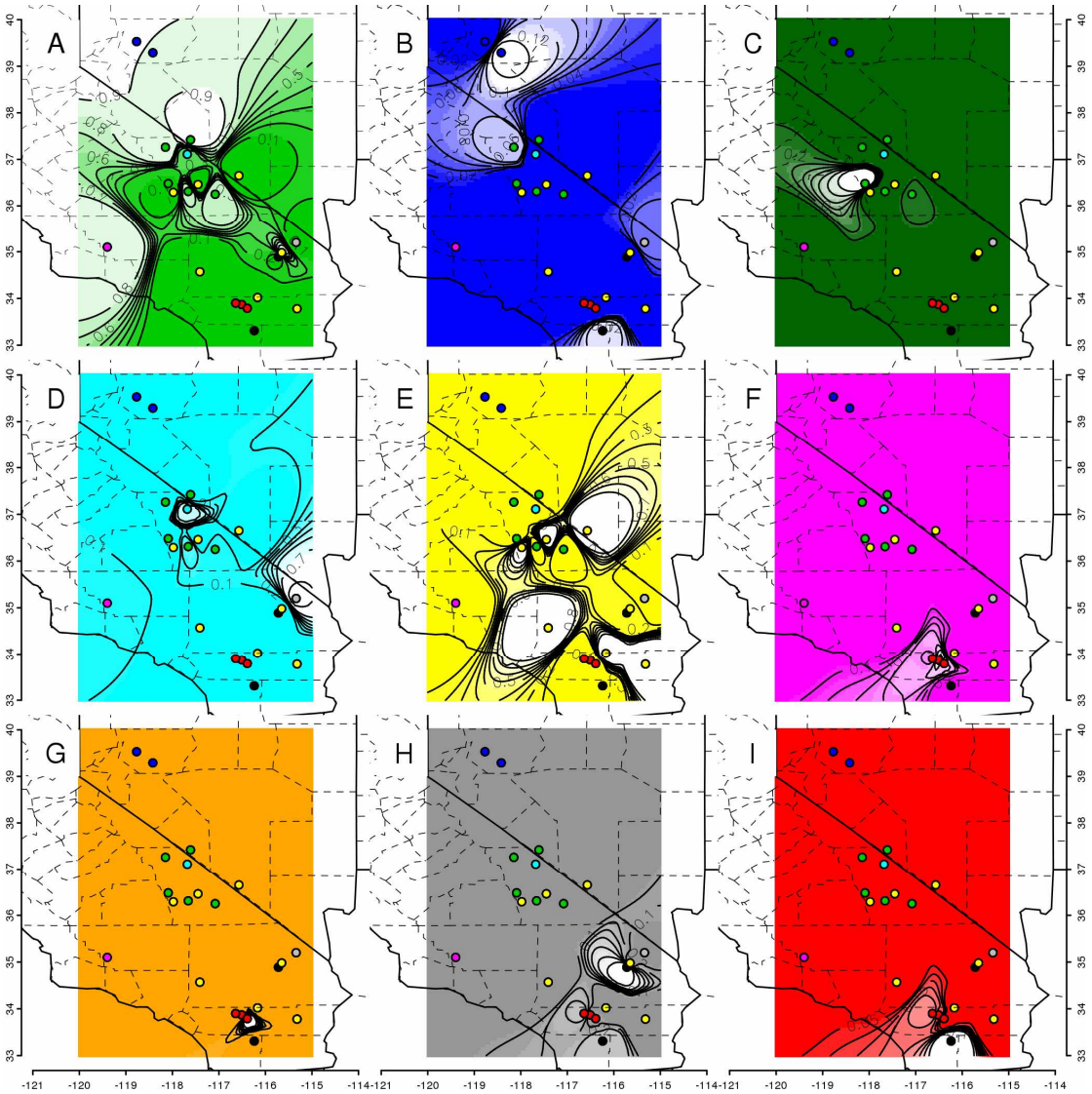


Figure 4.9. IDW interpolations of probability of group membership for $K = 9$. White represents high probability of membership and color represents low probability of membership. Colors correspond to the barplots of Figure 4.7.

Table 4.1. Summary of the AFLP sample.

variety	State	County	n*	population	latitude	longitude
borreganus	CA	San Bernardino	14	ASLEBO-22	34.89	-115.72
borreganus	CA	Imperial	8	ASLEBO-FW	33.31	-116.24
coachellae	CA	Riverside	21	ASLECO-1	33.87	-116.51
coachellae	CA	Riverside	10	ASLECO-2	33.79	-116.39
coachellae	CA	Riverside	18	ASLECO-38.2	33.90	-116.64
fremontii	CA	Inyo	10	ASLEFR-18	36.48	-118.09
fremontii	CA	Inyo	8	ASLEFR-40	36.31	-117.66
fremontii	CA	Inyo	10	ASLEFR-43	35.20	-115.34
fremontii	CA	Inyo	9	ASLEFR-44	36.25	-117.08
fremontii	CA	Inyo	9	ASLEFR-45	37.26	-118.15
fremontii	CA	San Bernardino	5	CONTFR-43	35.20	-115.34
fremontii	NV	Mina	9	ASLEFR-46	37.42	-117.61
kennedyi	NV	Churchill	7	ASLEKE-26a	39.29	-118.42
kennedyi	NV	Churchill	8	ASLEKEN-48	39.53	-118.77
micans	CA	Inyo	8	ASLEMI-17	37.11	-117.68
nigricalycis	CA	Kern	7	ASLENI-21	35.10	-119.40
variabilis	CA	Inyo	7	ASLEVA-25	36.46	-117.45
variabilis	CA	Inyo	5	ASLEVA-41	36.29	-117.98
variabilis	CA	Riverside	17	ASLEVA-38.3	33.78	-115.32
variabilis	CA	Riverside	15	ASLEVA-42	34.02	-116.17
variabilis	CA	San Bernardino	7	ASLEVA-4	34.98	-115.65
variabilis	CA	San Bernardino	14	ASLEVA-23	34.57	-117.41
variabilis	NV	Nye	5	ASLEVA-29	36.65	-116.57

* Sample size represents size after quality control screening.

Table 4.2. Distances for controls. All controls are the product of a single DNA extraction but separate AFLP chemistry. Percent pair-wise distance appears above the diagonal, Jaccard distance appears below the diagonal, percent bands per sample appears along the diagonal.

CONTFR-43_51	**0.1834**	0.0854	0.0653	0.1281	0.0779	0.1357
CONTFR-43_52	0.5920	**0.2186**	0.0854	0.0578	0.0578	0.1658
CONTFR-43_53	0.5563	0.5982	**0.1734**	0.1080	0.0678	0.1106
CONTFR-43_54	0.7037	0.4895	0.6658	**0.2060**	0.0704	0.1432
CONTFR-43_56	0.5869	0.4973	0.5603	0.5487	**0.1910**	0.1382
ASLEFR-43_5	0.8216	0.8424	0.7764	0.8141	0.8190	**0.0829**

5 Conclusions

Investigations into the varieties of *Astragalus lentiginosus* have employed three methodologies at several scales of sampling. Morphometric sampling has been distribution wide; however, due to the criterion of a sample of ten or more, many of the rare or otherwise less collected taxa were not been included in this dataset. This has resulted in a sample of 14 varieties, including 244 individuals (Table 1.2). Analysis of this dataset has resulted in significant correlations with climatic parameters but no finding of regions of discontinuity to define the varieties. Chloroplast simple sequence repeat (CpSSR) analysis represents the most taxonomically comprehensive study to date. The CpSSR sample included three to four individuals per population for 33 varieties, 71 populations and 272 individuals. Analyses of this dataset resulted in a finding of no correspondence of molecular divergence to varietal or sectional circumscription. This lack of correspondence is interpreted as being due to the widespread occurrence of haplotypes which predate the possible divergence among sections and varieties. This phenomenon is also known as incomplete lineage sorting or a lack of coalescence. Amplified fragment length polymorphism (AFLP) analysis focused on the southern California members of *A. lentiginosus*, a grouping consistent with the Rydbergian section *Coulteriana* (Rydberg 1929). This resulted in clustering which largely supports the taxonomy of Barneby (1964). However, molecular diversity is variable within these taxa and genotyping error plays a dramatic role in the interpretation of this data.

5.1 MORPHOMETRICS SUMMARY

Range-wide morphological analysis demonstrated an absence of regions of discontinuity among the varieties of *Astragalus lentiginosus*. This indicates that while the varieties do appear to be cohesive entities, the distinctions among them appear somewhat arbitrary. This is perhaps an obvious progression as initial taxa were described and eventually intermediates among these taxa became evident. Significant correlation of multivariate morphological characters to climatic parameters indicates that climate may have played an important role in the evolution of *A. lentiginosus*.

Because the phenotype is subject to natural selection, it appears intuitive that differentiation among varieties of *A. lentiginosus* may be due to selective pressure. A pattern of isolation by distance would demonstrate a more significant pattern due to geographic distance; instead it is observed that climatic parameters (ecological distance) play a more important role in explaining phenotypic distances. This result suggests that climatic correlation may play the greatest role in phenotypic differentiation (as opposed to geographic distance). This is interpreted as a sign of selection on the phenotype.

5.2 CpSSR SUMMARY

A range-wide survey of *A. lentiginosus* employing five chloroplast simple sequence repeats resulted in almost twice as many observed haplotypes as varieties. Haplotype sharing among varieties indicated a poor correspondence of molecular markers to either varietal (Barneby 1964) or sectional (Rydberg 1929) taxonomy. Hierarchical Bayesian clustering resulted in groups of haplotypes which did not appear to follow either a pattern of varietal or sectional differentiation, or an isolation by distance pattern. Given the significant pattern of phenotypic variation (Knaus Chapter 2), the lack of discernable pattern within *A. lentiginosus* and the large number of haplotypes has been interpreted as an instance where haplotypes may predate lineage divergence. This phenomenon is also known as incomplete lineage sorting or lack of coalescence. The finding of relatively little significance to taxonomic groupings should be expected based on simulations (Knaus et al. 2008). The observed incongruence between morphometric and molecular data has been interpreted as an instance of incomplete coalescence due to a large ancestral population size and a relatively small amount of time since divergence.

5.3 AFLP SUMMARY

A survey of the southern Californian varieties of *A. lentiginosus* (Rydberg's section *Coulteriana*) resulted in genetic structure that largely agrees with taxonomy, with a few notable exceptions. A transect through basins and ranges at 36.5° N shows

support for the variety *variabilis* as occurring in the basins while *fremontii* occurs in the ranges. The federally listed var. *coachellae* appears in multiple groups which are largely exclusive to this taxon, suggesting relatively high amounts of genetic diversity within this taxon as well. This variety also appears to be strongly differentiated from var. *variabilis* to the north but less differentiated from its southern neighbor var. *borreganus*.

5.4 SYNTHESIS

Each dataset (morphology, CpSSR and AFLP) shows a slightly different perspective on the affinities within *A. lentiginosus*. Morphology appears to agree with taxonomy but highlights the somewhat arbitrary nature of the distinctions made by monographers of the group. The CpSSR dataset does not appear to correspond with morphology, taxonomy or geography, indicating that lineage sorting from an ancestrally large population may be ongoing. If this is correct then analyses which assume equilibrium should be expected to return somewhat misleading results. The high number of markers produced in the AFLP analysis resulted in a dataset which appeared to correspond to morphological taxonomy and has revealed complex geographic structure. The strong correlation of the phenotype to possibly selective forces suggests adaptive divergence.

5.5 CONCLUSIONS

Several hypotheses were presented in Chapter one. The results of different marker systems support different hypotheses. Morphology largely tracks taxonomy and is clinal in respect to climate (Knaus Chapter 2). This appears best characterized by the hypothesis of isolation by distance, where morphological characters are correlated climatic distance (as opposed to geographic distance). Chloroplast simple sequence repeats (CpSSRs; Knaus Chapter 3) largely indicate a lack of pattern. The potentially confounding patterns of global panmixia and sudden vicariance appear to be disentangled via hierarchical *F*- statistics, which indicate that the greatest amount of variance is explained by the among population component. This suggests that the

hypothesis of sudden vicariance is supported over that of global panmixia. The lack of pattern to CpSSR data combined with significant morphological pattern is taken as a possible instance of incomplete lineage sorting. Amplified fragment length polymorphism (AFLP) demonstrated a somewhat intermediate perspective between morphology and CpSSRs. Taxonomy was largely supported, and regions of high molecular diversity among CpSSR and AFLPs appeared correlated, yet populations of some varieties were not clustered together, suggesting heterogeneous correspondence of AFLPs to morphology. The use of different marker systems resulted in differing perspectives on divergence within *Astragalus lentiginosus*.

The effect of breeding structure is assumed to be uniform over all loci and all alleles (Lewontin and Krakauer 1973). This is in congruence with the neutral theory (Kimura 1991). Deviation from this pattern has been the foundation for tests of selection such as the Lewontin-Krakauer test (Beaumont 2005) or Q_{ST} vs. F_{ST} comparisons (Merilä and Crnokrak 2001). The present study does not address either of these tests in a formal manner; however, qualitative comparison of each of the three datasets suggests a strict assumption of neutrality over all markers may not adequately explain the observed patterns. Heterogeneity among markers at some level should be expected due to differences in pattern of inheritance (maternally inherited chloroplast versus biparentally inherited nuclear AFLPs) or effective population size of a trait of interest (haploid chloroplast alleles versus diploid nuclear AFLP alleles versus quantitative traits which are likely to be the product of numerous loci). However, the apparent discrepancies among markers in the present study, particularly among morphology and CpSSRs, seem to indicate deviance from a strict pattern of breeding structure. It is hoped the information elucidated in this study will lay the groundwork for more formal implementation of these tests. As the most taxon rich species in the United States flora, *Astragalus lentiginosus* appears not only to be taxonomically diverse but perhaps is also the product of diverse evolutionary phenomena.

5.6 LITERATURE CITED

Barneby, R. C. 1964. Atlas of North American *Astragalus*. Memoirs of the New York Botanical Garden 13:1-1188.

Beaumont, M. A. 2005. Adaptation and speciation: what can F_{ST} tell us? Trends in Ecology & Evolution 20:435-440.

Kimura, M. 1991. The neutral theorey of molecular evolution: a review of recent evidence. Japanese Journal of Genetics 66:367-386.

Knaus, B. J., R. C. Cronn, and A. Liston. 2008. A fistful of *Astragalus*: incipient speciation in the American West? Presentation at Botany 2008.

Lewontin, R. C., and J. Krakauer. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. Genetics 74:175-195.

Merilä, J., and P. Crnokrak. 2001. Comparison of genetic differentiation at marker loci and quantitative traits. Journal of Evolutionary Biology 14:892-903.

Rydberg, P. A. 1929. *Astragalanae*. North American Flora 24:251-462.

6 Literature cited.

Ackerly, D. D., S. A. Dudley, S. E. Sultan, J. Schmitt, J. S. Coleman, C. R. Linder, D. R. Sandquist et al. 2000. The evolution of plant ecophysiological traits: recent advances and future directions. BioScience 50:979-995.

Alexander, J. A. 2005. The taxonomic status of *Astragalus mokiacensis* (Fabaceae), a species endemic to the southwestern United States. Brittonia 57:320-333.

—. 2007. A Taxonomic Revision of *Astragalus mokiacensis* and Allied Taxa within the *Astragalus lentiginosus* Complex of Section *Diphysi*. Ph.D. thesis, Oregon State University.

Alexander, J. A., A. Liston, and S. J. Popovich. 2004. Genetic diversity of the narrow endemic *Astragalus oniciformis* (Fabaceae). American Journal of Botany 91:2004-2012.

Allphin, L., N. Brian, and T. Matheson. 2005. Reproductive success and genetic divergence among varieties of the rare and endangered *Astragalus cremnophylax* (Fabaceae) from Arizona, USA. Conservation Genetics 6:803-821.

Amsellem, L., J. L. Noyer, T. Le Bourgeois, and M. Hossaert-McKey. 2000. Comparison of genetic diversity of the invasive weed *Rubus alceifolius* Poir. (Rosaceae) in its native range and in areas of introduction, using amplified fragment length polymorphism (AFLP) markers. Molecular Ecology 9:443-455.

Austin, M. P. 1985. Continuum concept: ordination methods, and niche theory. Annual Review of Ecology and Systematics 16:39-61.

Avise, J. C. 2004, Molecular markers, natural history, and evolution. Sunderland, Mass., Sinauer Associates.

Barneby, R. C. 1945. *Pugillus Astragalorum* IV: the section diplocystium. Leaflets of Western Botany 4:65-147.

—. 1964. Atlas of North American *Astragalus*. Memoirs of the New York Botanical Garden 13:1-1188.

—. 1977. *Dragma hippomanicum* III: *novitates Californicae*. Brittonia 29: 376-381.

—. 1989, Fabales: Intermountain Flora, v. 3 part B. Bronx, New York Botanical Garden.

Beaumont, M. A. 2005. Adaptation and speciation: what can F_{ST} tell us? Trends in Ecology & Evolution 20:435-440.

Beaumont, M. A., and D. J. Balding. 2004. Identifying adaptive genetic divergence among populations from genome scans. Molecular Ecology 13:969-980.

Becker, R. A., and A. R. Wilks. 1993. Maps in S. AT&T Bell Laboratories Statistics Research Report 95.2.

Bensch, S., and M. Åkesson. 2005. Ten years of AFLP in ecology and evolution: why so few animals? Molecular Ecology 14:2899-2914.

Bonin, A., E. Bellemain, P. Bronken Eidesen, F. Pompanon, C. Brochmann, and P. Taberlet. 2004. How to track and assess genotyping errors in population genetics studies. MJolecular Ecology 13:3261-3273.

Bonin, A., D. Ehrich, and S. Manel. 2007. Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionsits. Molecular Ecology 16:3737-3758.

Brownstein, M. J., J. D. Carpten, and J. R. Smith. 1996. Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. BioTechniques 20:1004-1010.

Chandler, G. T., and M. D. Crisp. 1998. Morphometric and phylogenetic analisis of the *Daviesia ulicifolia* complex (Fabaceae, Mirbelieae). Plant Systematics and Evolution 209:93-122.

Chessel, D., A.-B. Dufour, and J. Thioulouse. 2004. The ade4 package. R News 4:5-10.

Chung, M., G. Gelembuick, and T. J. Givnish. 2004. Population genetics and phylogeography of endangered *Oxytropis campestris* var. *chartacea* and relatives: artic-alpine disjuncts in eastern North America. Molecular Ecology 13:3657-3673.

Clausen, J., D. D. Keck, and W. M. Hiesey. 1939. The concept of species based on experiment. American Journal of Botany 26:103-106.

Clausen, J., D. D. Keck, and W. M. Hiesey. 1940, Experimental Studies on the Nature of Species. Washington D.C., Carnegie Institution of Washington.

Conti, F. 2007. Morphometric study and taxonomy of *Genista pulchella* Vis. *s.l.* (Fabaceae), a south European species. Botanical Journal of the Linneaen Society 153:245-254.

Corander, J., and P. Marttinen. 2006. Bayesian identification of admixture events using multilocus molecular markers. Molecular Ecology 15:2833-2843.

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

Criscione, C. D., and M. S. Blouin. 2007. Parasite phylogeographical congruence with salmon host evolutionary significant units: implications for salmon conservation. Molecular Ecology 16:993-1005.

Darwin, C. 1859, On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. *In* D.M. Porter and P.W. Graham *eds.* The Portable Darwin, 1993, Penguin Books.

Davis, M. B., and R. G. Shaw. 2001. Range shifts and adaptive responses to Quaternary climate change. Science 292:673-679.

de Meeûs, T., and J. Goudet. 2007. A step-by-step tutorial to use HierFstat to analyse populations hierarchically structured at multiple levels. Infection, Genetics and Evolution 7:731-735.

Dieckmann, U., M. Doebeli, J. A. J. Metz, and D. Tautz. 2004, Adaptive speciation: Cambridge studies in adaptive dynamics ; 3. Cambridge, UK ; New York, Cambridge University Press.

Dionne, M., F. Caron, J. J. Dodson, and L. Bernatchez. 2008. Landscape genetics and hierarchical structure in Atlantic salmon: the interaction of gene flow and local adaptation. Molecular Ecology 17:2382-2396.

Dobzhansky, T. 1951, Genetics and the Origin of Species. New York, Columbia University Press.

Doyle, J. J., M. Morgante, S. V. Tingey, and W. Powell. 1998. Size homoplasy in chloroplast microsatellites of wild perennial relatives of soybean (*Glycine* subgenus *Glycine*). Molecular Biology and Evolution 15:215-218.
Dytham, C. 1999, Choosing and Using Statistics: a Biologist's Guide, Blackwell Science.
Ehleringer, J. R. 1981. Leaf absorptances of Mojave and Sonoran Desert plants. Oecologia 49: 366-370.
Endler, J. A. 1977, Geographic Variation, Speciation, and Clines. Princeton, New Jersey, Princeton University Press.
Erich, D. 2006. AFLPDAT: a collection of R functions for convenient handling of AFLP data. Molecular Ecology Notes 6:603-604.
Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611-2620.
Everitt, B. 2005, An R and S-plus companion to multivariate analysis. London, Springer.
Falconer, D. S., and T. F. C. Mackay. 1996, Introduction to Quantitative Genetics, Pearson.
Falush, D., M. Stephens, and J. K. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567-1587.
—. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular Ecology Notes 7:574-578.
Felsenstein, J. 2004, Inferring Phyogenies, Sinauer.
Fisher, R. A. 1958, The Genetical Theory of Natural Selection, Dover.
Gomez-Mestre, I., and M. Tejedo. 2004. Contrasting patterns of quantitative and neutral genetic variation in locally adapted populations of the natterjack toad, *Bufo calamita*. Evolution 58:2343-2352.
Goudet, J. 2005. HIERFSTAT, a package for R to compute and test hierarchical *F*-statistics. Molecular Ecology Notes 5:184-186.
Gray, A. 1856. *Astragalus fremontii in* J. Torrey, ed. Pacific Rail Road Reports, Vol. 4., Report on the Botany of the Expedition. Washington, D.C.
—. 1863. Revision and arrangement (mainly by the fruit) of the North American species of *Astragalus* and *Oxytropis*. Proceedings of the American Academy of Arts and Sciences 6:188-237.
Grayson, D. K. 1993, The desert's past: a natural prehistory of the Great Basin. Washington, Smithsonian Institution Press.
Green, T. W., and G. E. Bohart. 1975. The pollination ecology of *Astragalus utahensis* (Leguminosae). American Journal of Botany 62:379-386.
Greuter, W., J. McNeill, F. R. Barrie, H.-M. Burdet, V. Demoulin, T. S. Filgueiras, D. H. Nicolson et al. 2000, International Code of Botanical Nomenclature (St. Louis Code). Konigstein, Koeltz Scientific Books.
Guillot, G., A. Estoup, F. Mortier, and J. F. Cosson. 2005. A spatial statistical model for landscape genetics. Genetics 170:1261-1280.

Guillot, G., F. Mortier, and A. Estoup. 2005. Geneland: a program for landscape genetics. Molecular Ecology Notes 5:712-715.
Haig, S. M., E. Beever, S. M. Chambers, H. M. Draheim, B. D. Dugger, S. Dunham, E. Elliot-Smith et al. 2006. Taxonomic considerations in listing subspecies under the U.S. endangered species act. Conservation Biology 20: 1584-1594.
Hale, M. L., A. M. Borland, M. H. G. Gustafsson, and K. Wolff. 2004. Causes of size homoplasy among chloroplast microsatellites in closely related *Clusia* species. Journal of Molecular Evolution 58:182-190.
Hedrick, P. W. 2005. A standardized genetic differentiation measure. Evolution 59:1633-1638.
Hooker, W. J. 1833, Flora Boreali-Americana, v. 1. London, Treuttel & Wurtz.
Hudson, R. R. 1991. Gene geneologies and the coalescent process, Pages 1-44 *in* D. J. Futuyma, and J. Antonovics, eds. Oxford Surveys in Evolutionary Biology. Oxford, UK, Oxford UNiversity Press.
Hudson, R. R., and J. A. Coyne. 2002. Mathematical consequences of the genealogical species concept. Evolution 56:1557-1565.
Huxley, J. 1938. Clines: an auxiliary taxonomic principle. Nature 142:219-220.
—. 1939. Clines: an auxiliary method in taxonomy. Bijdragen tot de Dierkunde (Contributions to Zoology) 27:491-520.
Isely, D. 1998, Native and Naturalized Leguminosae (Fabaceae) of the United Stated. Provo, Utah, Monte L. Bean Life Science Museum.
Jones, C. J., K. J. Edwards, S. Castaglione, M. O. Winfield, F. Sala, C. Van de Wiel, G. Bredemeijer et al. 1997. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Molecular Breeding 3:381-390.
Jones, C. J., K. J. Edwards, S. Castaglione, M. O. Winfield, F. Sala, C. Van der Wiel, B. Vosman et al. 1998. Reproducibility testing of AFLPs by a network of European laboratories *in* A. Karp, P. G. Isaac, and D. S. Ingram, eds. Molecular Tools for Screening Biodiversity. London, Chapman and Hall.
Jones, M. E. 1895. Contributions to Western Botany. No. 7. Proceedings of the California Academy of Sciences 5:611-732.
—. 1923, Revision of North American *Astragalus*. Salt Lake City, Utah.
Karron, J. D. 1989. Breeding systems and levels of inbreeding depression in geographically restricted and widespread species of *Astragalus* (Fabaceae). American Journal of Botany 76:331-340.
Karron, J. D., Y. B. Linhart, C. A. Chaulk, and C. A. Robertson. 1988. Genetic structure of populations of geographically restricted and widespread species of *Astragalus* (Fabaceae). American Journal of Botany 75:1114-1119.
Kimura, M. 1991. The neutral theorey of molecular evolution: a review of recent evidence. Japanese Journal of Genetics 66:367-386.
Knaus, B. J., R. C. Cronn, and A. Liston. 2005. Genetic characterization of three varieties of *Astragalus lentiginosus* (Fabaceae). Brittonia 57:334-344.
—. 2008. A fistful of *Astragalus*: incipient speciation in the American West? Presentation at Botany 2008.

Koopman, W. J. M., and G. Gort. 2004. Significance tests and weighted values for AFLP similarities, based on *Arabidopsis* in silico AFLP fragment length distributions. Genetics 167:1915-1928.

Koopman, W. J. M., V. Wissemann, K. De Cock, J. Van Huylenbroeck, J. De Riek, G. J. H. Sabatino, D. Visser et al. 2008. AFLP markers as a tool to reconstruct complex relationships: A case study in *Rosa* (Rosaceae). American Journal of Botany 95: 353-366.

Kropf, M. 2008. Intraspecific patterns of European mountain plants: a morphometric analysis confirms molecular results in the submediterranean oreophyte *Anthyllis montana* L. (Fabaceae). Taxon 57:511-524.

Kropf, M., J. W. Kadereit, and H. P. Comes. 2002. Late Quaternary distributional stasis in the submediterranean mountain plant *Anthyllis montana* L. (Fabaceae) inferred from ITS sequences and amplified fragment length polymorphisms markers. Molecular Ecology 11:447-463.

Lankau, R. A., and S. Y. Strauss. 2007. Mutual feedbacks maintain both genetic and species diversity in a plant community. Science 317:1561-1563.

Lewontin, R. C., and J. Krakauer. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. Genetics 74:175-195.

Linnaeus, C. 1753, Species Plantarum. Facsimile of the first edition with introduction by W.T. Stearn. 1957. London, The Ray Society.

Liston, A. 1990. An evolutionary study of *Astragalus* sect. *Leptocarpi* subsect. *californici* (Fabaceae). Ph.D. Thesis thesis, Claremont Graduate School, Claremont.

—. 1992. Isozyme Systematics of *Astragalus* Sect. *Leptocarpi* Subsect. *Californici* Fabaceae. Systematic Botany 17:367-379.

Lock, J. M., and B. D. Schrire. 2005. Galegeae, Pages 475-487 *in* G. Lewis, B. Schrire, B. MacKinder, and M. Lock, eds. Legumes of the World. Kew, Royal Botanical Gardens.

Lunn, D. J., A. Thomas, N. Best, and D. Spiegelhalter. 2000. WinBUGS -- a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing 10:325-337.

Lynch, M. 1990. The similarity index and DNA fingerprinting. Molecular Biology and Evolution 7:478-484.

Mabberley, D. J. 1987, The plant-book : a portable dictionary of the higher plants utilising Cronquist's An integrated system of classification of flowering plants (1981) and current botanical literature arranged largely on the principles of editions 1-6 (1896/97-1931) of Willis's A dictionary of the flowering plants and ferns. New York, Cambridge University Press.

Mallet, J. 2007. Subspecies, semispecies, superspecies, Pages 523-526 *in* S. Levin, and et al., eds. Encyclopedia of Biodiversity, Academic Press.

Mayden, R. L. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem, Pages 381-422 *in* M. F. Claridge, H. A. Dawah, and M. R. Wilson, eds. Species: The Units of Biodiversity, Chapman & Hall.

Mayr, E. 1982. Of what use are subspecies? Auk 99:593-595.

Mayr, E., and P. D. Ashlock. 1991, Principles of Systematic Zoology, McGraw-Hill.

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

Meagher, T. R., J. Antonovics, and R. Primack. 1978. Experimental ecological genetics in *Plantago*. III. Genetic variation and demography in relation to survival of *Plantago cordata*, a rare species. Biological conservation 14:243-257.

Merilä, J., and P. Crnokrak. 2001. Comparison of genetic differentiation at marker loci and quantitative traits. Journal of Evolutionary Biology 14:892-903.

Morjan, C. L., and L. H. Rieseberg. 2004. How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. Molecular Ecology 13:1341-1356.

Morris, A. B., R. S. Baucom, and M. B. Cruzan. 2002. Stratified analysis of the soil seed bank in the cedar glade endemic *Astragalus bibullatus*: evidence for historical changes in genetic structure. American Journal of Botany 89:29-36.

Navascués, M., and B. C. Emerson. 2005. Chloroplast microsatellites: measures of genetic diversity and the effect of homoplasy. Molecular Ecology 14:1333-1341.

Nei, M., and W.-H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction nucleases. Proceedings of the National Academy of Science, U.S.A. 76:5269-5273.

Nordborg, M. 2001. Coalescent theory, Pages 179-212 *in* D. J. Balding, M. J. Bishop, and C. Canings, eds. Handbook of Statistical Genetics. Chichester, UK, John Wiley and Sons.

Orr, H. A. 2001. The genetics of species differences. Trends in Ecology & Evolution 16:343-350.

Paradis, E. 2006, Analysis of phylogenetics and evolution with R. New York, Springer.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289-290.

Patten, M. A., and P. Unitt. 2002. Diagnodability versus mean differences of Sage Sparrow subspecies. The Auk 119:26-35.

Pebesma, E. J. 2004. Multivariable geostatistics in S: the gstat package. Computers & Geosciences 30:383-691.

Podolsky, R. H., and T. P. Holtsford. 1995. Population structure of morphological traits in *Clarkia dudleyana* I. Comparison of F_{ST} between allozymes and morphological traits. Genetics 140:733-744.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

R Development Core Team. 2007.R: a language and environment for statistical computing.R foundation for statistical computing, Vienna, Austria.

Raven, P. H., and G. B. Johnson. 2002, Biology, McGraw-Hill.

Reig, S., M. J. Daniels, and D. W. Macdonald. 2001. Craniometric differentiation within wild-living cats in Scotland using 3D morphometrics. Journal of Zoology 253:121-132.

Rieseberg, L. H., and J. M. Burke. 2001. The biological reality of species: Gene flow, selection, and collective evolution. Taxon 50:47-67.
Rieseberg, L. H., T. E. Wood, and E. J. Baack. 2006. The nature of plant species. Nature 440:524-527.
Roff, D. A., and T. Mousseau. 2005. The evolution of the phenotypic covariance matrix: evidence for selection and drift in *Melanoplus*. Journal of Evolutionary Biology 18:1104-1114.
Ronquist, F., and J. P. Huelsenbeck. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.
Rosenberg, N. A., and M. Nordborg. 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. Nature Reviews Genetics 3:380-390.
Rosenburg, N. A., J. K. Pritchard, J. L. Weber, H. M. Cann, K. K. Kidd, L. A. Zhivotovsky, and M. W. Feldman. 2002. Genetic structure of human populations. Science 298:2381-2385.
Rouppe van der Voort, J. N. A. M., P. van Zandvoort, H. J. van Eck, R. T. Folkertsma, R. C. B. Hutten, J. Draaistra, F. J. Gommers et al. 1997. Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. Molecular and General Genetics 255:438-447.
Rydberg, P. A. 1929a. Astragalanae. North American Flora 24:251-462.
—. 1929b. Scylla and Charybdis. Proceedings of the International congress of plant sciences, Ithica, New York 2:1539-1551.
Sandquist, D. R., and J. R. Ehleringer. 1997. Intraspecific variation in leaf pubescence and drought response in *Encelia farinosa* associated with contrasting desert environments. New Phytologist 135:635-644.
Scherson, R. A., R. Vidal, and M. J. Sanderson. 2008. Phylogeny, biogeography, and rates of diversification of New World *Astragalus* (Leguminosae) with an emphasis on South American radiations. American Journal of Botany 95:1030-1039.
Schönswetter, P., A. Tribsch, and H. Nikfeld. 2004. Amplified fragment length polymorphism (AFLP) reveals no genetic divergence of the Eastern Alpine endemic *Oxytropis campestris* subsp. *tiroliensis* (Fabaceae) from widespread subsp. *campestris*. Plant Systematics and Evolution 244:245-255.
Sheidai, M., Z. Yazdanbakhsh, M. Assadi, and M. Moussavi. 2001. Cytology and morphometric study of *Alhagi* (Leguminosae) species in Iran. Nordic Journal of Botany 21:83-91.
Sheldon, E. P. 1894. A preliminary list of the North American species of *Astragalus*. Minnesota Botanical Studies 1:116-175.
Slatkin, M. 1976. The rate of spread of an advantageous allele in a subdivided population *in* S. Karlin, and E. Nevo, eds. Population Genetics and Ecology. New York, NY, Academic Press, Inc.
Small, E., and B. Brookes. 1990. A numerical taxonomic analysis of the *Medicago littoralis - M. truncatula* complex. Canadian Journal of Botany 68:1667-1674.
Small, E., W. F. Grant, and C. W. Crompton. 1984. A taxonomic study of the *Lotus corniculatus* complex in Turkey. Canadian Journal of Botany 62:1044-1053.

Spellenberg, R. 1976. Chromosome numbers and their cytotaxonomic significance for North American *Astragalus* (Fabaceae). Taxon 25:463-476.
—. 1993. *Astragalus*, Pages 583-605 *in* J. C. Hickman, ed. The Jepson Manual: Higher Plants of California. Berkeley, CA, University of California Press.
Spitze, K. 1993. Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. Genetics 135:367-374.
St Clair, J. B., N. L. Mandel, and K. W. Vance-Borland. 2005. Genecology of douglas fir in western Oregon and Washington. Annals of Botany 96:119-1214.
Stearn, W. T. 1957. Introduction Species Plantarum: a Facsimile of the First Edition, 1753. London, The Ray Society.
Stebbins, G. L. 1993. Concepts of Species and Genera *in* Flora of North America Editorial Committee, ed. Flora of North America. New York and Oxford, Oxford University Press.
Stebbins, G. L. J. 1950, Variation and Evolution in Plants. New York, Columbia University Press.
Storfer, A. 1996. Quantitative genetics: a promising approach for the assessment of genetic variation in endangered species. Trends in Ecology and Evolution 11:343-348.
Storz, J. F. 2002. Contrasting patterns of divergence in quantitative traits and neutral DNA markers: analysis of clinal variation. Molecular Ecology 11:2537-2551.
Syring, J., K. Farrell, R. Businský, R. Cronn, and A. Liston. 2007. Widespread genealogical nonmonophyly in species of *Pinus* subgenus *Strobus*. Systematic Biology 56:163-181.
Tabachnick, B. G., and L. S. Fidell. 2007, Using Multivariate Statistics, Pearson.
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 17:1105-1109.
Thompson, R. S., and J. I. Mead. 1982. Late quaternary environments and biogeography in the great basin. Quaternary Research 17:39-55.
Travis, S. E., J. Maschinski, and P. Keim. 1996. An analysis of genetic variation in *Astragalus cremnophylax* var. *cremnophylax*, a critically-endangered plant, using AFLP markers. Molecular Ecology 5:735-745.
USDA NRCS. 2006. The PLANTS database, National Plant Data Center, Baton Rogue, LA 70874-4490 USA.
USFWS. 1978. Determination that 11 plant taxa are endangered and 2 plant taxa are threatened species. Federal Register 43:17910-17916.
—. 1998. Endangered and threatened wildlife and plants; determination of endangered or threatened status for five desert milk-vetch taxa from California. Federal Register 63:53596-53615.
Via, S. 2002. The ecological genetics of speciation. The American Naturalist 159:S1-S7.
Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters et al. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23:4407-4414.

Waitt, D. E., and D. A. Levin. 1997. Genetic and phenotypic correlations in plants: a botanical test of Cheverud's conjecture. Heredity 80:310-319.
Wallace, A. R. 1858. On the tendency of varieties to depart indefinitely from the original type *in* D. M. Porter, and P. W. Graham, eds. The Portable Darwin, 1993, Penguin Books.
Walsh, B. 2001. Quantitative genetics in the age of genomics. Theoretical Population Biology 59:175-184.
Weising, K., and R. C. Gardner. 1999. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. Genome 42:9-19.
Welsh, S. L. 1981. New taxa of western plants - in tribute. Brittonia 33:294-303.
—. 2007, North American species of *Astragalus* Linnaeus (Leguminosae), a taxonomic revision. Prove, UT, Monte L. Bean Life Science Museum.
Welsh, S. L., and N. D. Atwood. 2001. New taxa and nomenclatural proposals in miscellaneous families - Utah and Arizona. Rhodora 103:71-95.
Welsh, S. L., N. D. Atwood, S. Goodrich, and L. C. Higgins. 2003, A Utah Flora. Provo, Brigham Young University.
Welsh, S. L., and R. C. Barneby. 1981. *Astragalus lentiginosus* (Fabaceae) revisited - a unique new variety. Iselya 2:1-2.
Whitlock, R., H. Hipperson, M. Mannarelli, R. K. Butlin, and T. Burke. 2008. An objective, rapid and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. Molecular Ecology Resources 8:725-735.
Wilding, C. S., R. K. Butlin, and J. Grahame. 2001. Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. Journal of Evolutionary Biology 14:611-619.
Wilson, E. O., and W. L. Brown. 1953. The subspecies concept and its taxonomic application. Systematic Zoology 2:97-111.
Wojciechowski, M. F., M. J. Sanderson, B. G. Baldwin, and M. J. Donoghue. 1993. Monophyly of aneuploid *Astragalus* (Fabaceae): evidence from nuclear ribosomal DNA internal transcribed spacer sequences. American Journal of Botany 80:711-722.
Wojciechowski, M. F., M. J. Sanderson, and J.-M. Hu. 1999. Evidence on the monophyly of *Astragalus* (Fabaceae) and its major subgroups based on nuclear ribosomal DNA ITS and chloroplast DNA *trnL* intron data. Systematic Botany 24:409-437.
Wolfe, A. D., and A. Liston. 1998. Contributions of PCR-based methods to plant systematics and evolutionary biology *in* D. E. Soltis, P. S. Soltis, and J. J. Doyle, eds. Molecular Systematics of Plants II: DNA Sequencing. Boston, MA, Kluwer Academic.
Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proceedings of the Sixth International Congress of Genetics 1:356-366.
—. 1978, Evolution and the genetics of populations. Volume 4: variability within and among natural populations. Chicago, University of Chicago Press.

Wu, C.-I. 2001. The genic view of the process of speciation. Journal of Evolutionary Biology 14:851-865.
Zink, R. M. 2004. The role of subspecies in obscuring avian biological diversity and misleading conservation policy. Proceedings of the Royal Society of London B 271:561-564.

APPENDICES

Appendix A. Morphology specimens. State, county, collection and number, herbarium acronym and accession number and inferred geographic position. Herbaria acronyms follow *Index Herbariorum* (http://sciweb.nybg.org/science2/IndexHerbariorum.asp).

Astragalus lentiginosus Douglas ex Hooker **var. *araneosus*** (Sheldon) M.E.Jones– NV, Lincoln Co.: *Tiehm, A. & M. Williams #6620* (NY), 38.35˚ N, 114.20˚ W; White Pine Co.: *Holmgren, N.H. & J.L. Reveal #977* (NY), 38.90˚ N, 114.35˚ W; *Ripley, H.D. & R.C. Barneby #3535* (RSA #143745), 38.85˚ N, 114.41˚ W; *Williams, M.J. #77-10-6* (NY), 39.20˚ N, 115.88˚ W; *Williams, M.J. & A. Tiehm #81-16-12* (RSA #301284), 39.12˚ N, 114.28˚ W. UT, Beaver Co.: *Holmgren, N.H. & P.K. #10391* (NY), 38.32˚ N, 113.49˚ W; *Welsh, S.L. #20484* (NY), 38.40˚ N, 113.29˚ W; *Welsh, S.L. & M. Chatterley #19568* (NY), 38.47˚ N, 113.46˚ W. Iron Co.: *Franklin, B. & L. Armstrong #7658* (NY), 37.86˚ N, 113.04˚ W; *Welsh, S.L. #5224* (NY), 37.62, ˚ N, 113.39˚ W; *Welsh, S.L. #23785* (NY), 37.93˚ N, 112.85˚ W. Juab Co.: *Schoener, C.S. & M. Wright #67* (NY), 39.71˚ N, 111.84˚ W. Millard Co.: *Atwood, D. #24701* (NY), 38.79˚ N, 113.77˚ W. Piute Co.: *Atwood, D. #7204* (NY) 38.47˚ N, 111.90˚ W. San Pete Co.: *McNeal, Frey, Gray & Smookler #1708* (NY), 39.52˚ N, 111.49˚ W. Sevier Co.: *Cronquist, A. #11558* (NY) 38.73˚ N, 112.50˚ W; *Welsh, S.L. & J.R. Murdock #12382* (NY), 38.84˚ N, 111.90˚ W; *Zupan, C. & K. Thorne #163* (NY), 38.55˚ N, 112.22˚ W. Wayne Co.: *Holmgren, N.H., J.L. Reveal & C. LaFrance #2089* (NY), 38.26˚ N, 111.62˚ W; *Kass, R. #5057* (NY), 38.21˚ N, 110.63˚ W.

A. l.* var. *borreganus M.E.Jones— CA, Imperial Co.: *Balls, E.K. & P.C. Everett #22896* (RSA #124378), 32.82˚ N, 114.84˚ W; *Davidson, C., A. Romspert & H. Suprenant #7749* (RSA #415578), 32.76˚ N, 114.83˚ W; *Johansen, D. A., & Ewan, J. A. #7142* (POM #186866), 32.76˚ N, 114.84˚ W; *Kline, E., s.n.* (RSA #498239), 32.85˚ N, 115.57˚ W; *Rich, B.M. #79002* (RSA #291587), 32.75˚ N, 114.87˚ W; *Roos, J. & L. #4166* (RSA #45524), 32.71˚ N, 115.06˚ W. San Bernardino Co.: *Benson, L. #8254* (POM #275872), 35.10˚ N, 116.27˚ W; *Cooper, N.C. #3341* (RSA #415580), 34.08˚ N, 114.85˚ W; *Martens, S. & B. Baldwin #125* (RSA #295870), 34.91˚ N, 115.73˚ W; *Peirson, F. W. #7759* (RSA #90171), 35.37˚ N, 116.12˚ W; *Ripley, H. D., Barneby, R. C. #3360* (RSA #112046), 35.47˚ N, 115.27˚ W; *Stone, R.D., S. Castagnoli & G. de Nevers #77* (RSA #296935), 35.64˚ N, 115.96˚ W; *Thorne, R.F., C.W. Tilforth & R.K. Benjamin #49392* (RSA #275194), 34.91˚ N, 115.73˚ W; *Thorne, R.F., C.W. Tilforth, A. Scmida, et al #51239* (RSA #275585), 34.91˚ N, 115.73˚ W; *Woglum, R.S. s.n.* (RSA #595223), 35.01˚ N, 115.65˚ W; *Wolf, C.B. #10236* (RSA #24666), 34.91˚ N, 115.73˚ W. San Diego Co.: *Jones, M.E. s.n.* (POM #27036), 33.23˚ N, 116.27˚ W; *Jones, M.E. s.n.* (POM #27037), 33.23˚ N, 116.27˚ W; *Moran, R. #6538* (RSA #139585), 35.29˚ N, 116.23˚ W.

A. l.* var. *chartaceus M.E.Jones— CA, Siskiyou Co.: *Brandegeee, K. s.n.* (UC #83718), 41.73˚ N, 122.53˚ W. ID, Bear Lake Co.: *Ripley, H.D. & R.C. Barneby #8788* (RSA #109618), 42.48˚ N, 111.37˚ W. NV, Humboldt Co.: *Tiehm, A. & L. Birdsey #5000* (RENO #2522), 41.31˚ N, 118.43˚ W; *Tiehm, A. & B. Rogers*

#4433 (RENO #2521), 41.73˚ N, 119.16˚ W; *Williams, M.J. & A. Tiehm #82-59-1* (RENO #2520), 41.53˚ N, 118.83˚ W. Washoe Co.: *Pinzl, A. #3915* (NESH #7096), 40.92˚ N, 119.56˚ W; *Tiehm, A. #12523* (RENO #2525), 40.85˚ N, 119.56˚ W. OR, Harney Co.: *Hitchcock, C.L. & C.V. Muhlick #21127* (RM #256739), 42.81˚ N, 118.41˚ W. Lake Co.: *Colwell, A. & J. Myers #JM234* (RENO #2527), 42.60˚ N, 119.57˚ W; *Ripley, H.D. & R.C. Barneby #6049* (RSA #109626), 42.29˚ N, 119.82˚ W.

A. l.* var. *coachellae Barneby ap. Shreve & Wiggins— CA, Riverside Co.: *Clokey, I.W. & B.C. Templeton #4710* (POM #250336), 33.83˚ N, 116.54˚ W; *Davidson, C. #7815* (RSA #415588), 33.72˚ N, 116.22˚ W; *Epling, C. s.n.* (OSC #30495), 33.82˚ N, 116.39˚ W; *Grant, G.R. #6714* (POM #27038), 33.83˚ N, 116.54˚ W; *Hall, H.M. #5757* (ORE #50304), 33.93˚ N, 116.64˚ W; *Hall, H.M. #5757* (POM #24940), 33.93˚ N, 116.64˚ W; *Hawks, D. s.n.* (RSA #507459), 33.90˚ N, 116.56˚ W; *Johnston, I. #1065* (POM #1694), 33.93˚ N, 116.64˚ W; *Jones, M.E. s.n.* (POM #24941), 33.83˚ N, 116.54˚ W; *Minthorn, T. s.n.* (RSA #498238), 33.72˚ N, 116.22˚ W; *Mitchell, D. s.n.* (RSA #627041), 33.79˚ N, 116.39˚ W; *Munz, P.A. & D. Keck #4963* (POM #130), 33.72˚ N, 116.31˚ W; *Parish, S.B. #6119* (ORE #50299), 33.93˚ N, 116.64˚ W; *Parish, S.B. & W.F. #25* (ORE #50301), 33.93˚ N, 116.64˚ W; *Peirson, F.W. #585* (RSA #90167), 33.93˚ N, 116.64˚ W*; Pierson, F.W. s.n.* (RSA #498235), 33.93˚ N, 116.64˚ W; *Ripley, H.D. & R.C. Barneby #4271* (RSA #112049), 33.83˚ N, 116.54˚ W*; Spencer, M.F. #1501* (POM #9368), 33.93˚ N, 116.64˚ W; *Templeton, B.C. & I.W. Clokey #1072* (RSA #415586), 33.91˚ N, 116.65˚ W; *Wolf, C.B. #3685* (RSA #6117), 33.91˚ N, 116.65 ˚ W.

A. l.* var. *diphysus (A.Gray) M.E.Jones— AZ, Apache Co.: *Higgins, L.C. #5420* (NY), 34.49˚ N, 109.62˚ W. Coconino Co.: *Batchelder, G. & E. Lehto s.n.* (NY), 35.11˚ N, 111.05˚ W; *Demaree, D. #43977* (NY), 35.56˚ N, 111.35˚ W; *Demaree, D. #44052* (NY), 35.56˚ N, 111.35˚ W*; Demaree, D. #44194* (NY), 35.20˚ N, 111.65˚ W; *Higgins, L.C. #5415* (NY), 35.20˚ N, 111.44˚ W; *MacDougal, D.T. #438* (US #47694), 35.20˚ N, 111.42˚ W. Navajo Co.: *Spellenberg, R., R. Delson, J. Syvertsen #3286* (NY), 34.88˚ N, 110.13˚ W. Yavapai Co*.: Demaree, D. #43937* (NY), 34.90˚ N, 112.19˚ W; *Demaree, D. #43958* (NY), 34.69˚ N, 112.13˚ W. NM, Chaves Co.: *Higgins, L.C. #7031* (NY), 33.60˚ N, 104.32˚ W. San Juan Co.: *Welsh, S., & M. Porter #24422* (NY), 36.79˚ N, 108.69˚ W. Santa Fe Co.: *Fendler, A. #146* (K, H2005/02196 18), 35.69˚ N, 105.94˚ W; *Heller, A.A. & E.G. #3541* (NY), 35.69˚ N, 105.94˚ W. Socorrow Co.: *Webber, W.A. & P. Salamun #12765* (NY), 34.12˚ N, 107.24˚ W. TX, Hudspeth Co.: *Correll, D.S. & H.B. #38532* (NY), 31.30˚ N, 105.85˚ W.

A. l.* var. *floribundus A.Gray— CA, Lassen Co.: Ripley, *H.D. & R.C. Barneby #5678* (RSA #112071), 40.03˚ N, 120.10˚ W; *Tiehm, A. #5751* (RENO #2545), 40.36˚ N, 120.24˚ W. Plumas Co.: *Harnach, W & N #1059* (RENO #2536), 39.76˚ N, 120.23˚ W. NV, Carson City Co.: *Anderson, C.L. s.n.* (K #H2005/02196 26), 39.16˚ N, 119.77˚ W; Jones, *M.E. s.n.* (POM #25704), 39.16˚ N, 119.77˚ W; *Pinzl, A. #11564* (NESH #12764), 39.19˚ N, 119.73˚ W; *Pinzl, A. #11575* (NESH #12763), 39.19˚ N, 119.76˚ W. Eureka Co.: *Pinzl, A. #11949* (NESH #13302),

39.86˚ N, 116.20˚ W. Lyon Co.: *Tiehm, A. #5882* (RENO #2549), 38.96˚ N, 119.28˚ W; *Tiehm, A. #10337* (NESH #5352G), 39.05˚ N, 119.33˚ W. Storey Co.: *Tiehm, A. #5698* (RENO #2546), 39.51˚ N, 119.67˚ W. Washoe Co.: *Kennedy, P.B. #3000* (RENO #2550), 39.64˚ N, 119.84˚ W; *Petersen, M.F. #215* (RENO #2547), 39.39˚ N, 119.74˚ W; *Tiehm, A. #6339* (RENO #2543), 39.54˚ N, 119.62˚ W.

A. l.* var. *fremontii (A.Gray) S.Watson— CA, San Bernardino Co.: *Alexander, A.M. & L. Kellogg #1329* (POM #299198), 35.23˚ N, 115.35˚ W*; Brandegee, K. & J.S. Brandegee s.n.* (POM #24932), 35.24˚ N, 115.50˚ W; *Everett, P.C. & E.K. Balls #23150* (RSA #124292), 35.27˚ N, 115.28˚ W; *Knaus, B.J. #126* (NA), 35.19˚ N, 115.34˚ W; *Martens, S. & B. Baldwin #105* (RSA #295907), 34.98˚ N, 115.73˚ W; *Ripley, H.D. & R.C. Barneby #3298* (POM #267419), 35.40˚ N, 115.63˚ W; *Ripley, H.D. & R.C. Barneby #4286* (POM #265860), 35.38˚ N, 115.89˚ W; *Ripley, H.D. & R.C. Barneby #4286* (RSA #112081), 35.84˚ N, 115.89˚ W; *Thorne, R.F. & C.W. Tilforth #43705* (RSA #251967), 35.21˚ N, 115.29˚ W; *Woglum, R.S. #2292* (RSA #605676), 35.34˚ N, 115.31˚ W; *Wolf, C.B. #3334* (RSA #5121), 35.37˚ N, 115.50˚ W. Inyo Co.: *Hall, H. M. and H. P. Chandler #7349* (OSC #144341), 36.80˚ N, 118.20˚ W; *Knaus, B.J. #44* (NA), 36.47˚ N, 118.09˚ W; *Knaus, B.J. #63* (NA), 36.35˚ N, 117.57˚ W; *Knaus, B.J. #111* (NA), 36.31˚ N, 117.65˚ W; *Knaus, B.J. #125b* (NA), 36.24˚ N, 117.07˚ W; *Knaus, B.J. #128* (NA) 37.26˚ N, 118.15˚ W; *Munz, P. A. #18073* (OSC #134057), 36.57˚ N, 118.09˚ W. NV, Esmeralda Co.: *Knaus, B.J. #129* (NA), 37.42˚ N, 117.61˚ W. Mineral Co.: *Vreeland, P.H. #11-23* (RENO #2634), 38.60˚ N, 118.11˚ W. Nye Co*.: Holmgren, N.H. #12321* (OSC #190883), 38.20˚ N, 116.18˚ W.

A. l.* var. *ineptus (A.Gray) M.E.Jones— CA, Alpine Co.: *Gifford, A.D. #750* (UC #572476), 38.40˚ N, 119.62˚ W; *Hoover, R.F. #4438* (UC #764115), 38.34˚ N, 119.63˚ W; *Wiggins, I.L. #9312* (UC #652566), 38.35˚ N, 119.63˚ W. Inyo Co.: *Peirson, F.W. s.n.* (JEPS #27459), 37.45˚ N, 118.72˚ W; *Peirson, F.W. s.n.* (UC #511813), 37.45˚ N, 118.72˚ W; *Raven, P. & G.L. Stebbins, Jr. #237* (UC #914619), 37.21˚ N, 118.54˚ W; *Taylor, D.W. #6650* (JEPS #091411), 37.17˚ N, 118.54˚ W. Mono Co.: *Alexander, A.M. #3930* (JEPS #710), 38.47˚ N, 119.27˚ W; *Alexander, A. & L. Kellogg s.n.* (UC #1368641), 38.24˚ N, 119.44˚ W; *Alexander, A.M. & L. Kellogg #4036* (UC #702111), 38.43˚ N, 119.32˚ W; *Alexander, A.M. & L. Kellogg #4211* (JEPS #22590), 32.33˚ N, 119.64˚ W; *Eastwood, A. & J.T. Howell #7533* (UC #863008), 38.33˚ N, 119.64˚ W; *Hendrix, T.M. #395* (UC M 128815), 38.12˚ N, 119.33˚ W; *Honer, M. #506* (UC #1786858), 37.81˚ N, 118.82˚ W; *Hoover, R.F. #5537* (UC #764112), 38.42˚ N, 119.37˚ W; *Sharsmith, C.W. #2563* (UC #712304), 37.89˚ N, 119.21˚ W; *Sharsmith, C.W. #2957* (UC #712294), 38.08˚ N, 119.28˚ W; *Wiggins, I.L. & R.C. Rollins #572* (UC #727389), 38.24˚ N, 119.08˚ W. Tuolumne Co.: *Colwell, A. et al .#AC05-133* (UC #1861862), 38.07˚ N, 119.33˚ W; *Sharsmith, C.W. #2715* (UC #712308), 37.84˚ N, 119.22˚ W; *Sharsmith, C.W. #2901* (UC #712293), 38.30˚ N, 119.66˚ W.

A. l.* var. *kennedyi (Rydberg) Barneby**—** NV, Churchill Co.: *Billings, W.D. #1523* (RENO #6290), 39.28˚ N, 118.42˚ W; *Headley, F.B. #7* (RENO #26542), 39.46˚ N, 118.75˚ W; *Kennedy, P.B. #1691* (RENO #9461), 39.53˚ N, 118.77˚ W*; Knaus, B.J. #70* (NA), 39.29˚ N, 118.42˚ W; *Knaus, B.J. #131* (NA), 39.53˚ N, 118.77˚ W; *Lehenbauer, P.A. s.n.* (RENO #2650), 39.93˚ N, 118.77˚ W; *Rust, R. & L. Hanks s.n.* (NESH #6173G), 39.30˚ N, 118.40˚ W; *Tiehm, A. #1038* (RENO #2637), 39.56˚ N, 118.73˚ W; *Williams, M.J. #78-44-1* (RENO #53935), 39.30˚ N, 118.41˚ W; *Williams, M.J. #78-111-5* (RENO #3453), 39.29˚ N, 118.40˚ W. Lyon Co.: *Schmidt, W.H. s.n.* (RENO #15531), 38.98˚ N, 119.10˚ W; *Tiehm, A #678* (RENO #34789), 39.68˚ N, 119.12˚ W. Mineral Co.: *Mozingo, H. #77-66* (RENO #39777), 38.45˚ N, 118.77˚ W; *Tiehm, A. #8551* (NESH #3117G), 38.96˚ N, 118.73˚ W. Nye Co.: *Tiehm, A. #5909* (RENO #67386), 38.92˚ N, 118.13˚ W; *Williams, M.J. #80-28-1* (RENO #49999), 38.86˚ N, 117.52˚ W. Washoe Co.: *Billings, W.D. #1603* (RENO #11848), 39.95˚ N, 119.51˚ W; *Frandsen & Brown #148* (RENO #2641), 39.95˚ N, 119.60˚ W; *Tiehm, A. #2099* (RENO #37218), 39.92˚ N, 119.56˚ W.

***A. l.* var. *lentiginosus*—** CA, Siskiyou Co.: *Heller, A.A. #8062* (UC #144159), 41.65˚ N, 122.52˚ W. ID, Ada Co.: *Ertter, B. & L. Smithman #4231* (RENO #2659), 43.55 ˚ N, 116.16 ˚ W. Gem Co.: *MacBride, J.F. #896* (RM #71671), 43.87˚ N, 116.50˚ W. OR, Baker Co.*: Jones, M.E. s.n.*(POM #25716), 44.78˚ N, 117.83˚ W; Klamath Co.: *Hitchcock, C. L. #25674* (OSC #137558), 42.44˚ N, 121.27˚ W; *Rittenhouse, B. #288* (OSC #173939), 42.28˚ N, 121.20˚ W. Morrow Co.: *Halse, R.R. #3431* (OSC #169682), 45.71˚ N, 119.57˚ W; *Kagan, J. s.n.* (OSC #202116), 45.50˚ N, 119.82˚ W. Sherman Co.: *Ewing, F. C. s.n.* (OSC #5765), 45.59˚ N, 120.70˚ W; *Gorman, M. W. s.n.* (ORE #112857), 45.48˚ N, 120.73˚ W. WA, Adams Co.: *Cotton, J. S. #975* (RM #114549), 46.76˚ N, 118.31˚ W. Grant Co.: *Rogers, H. T. #448* (POM 263311), 47.94˚ N, 119.00˚ W. Kittitas Co.: *Thompson, J.W. #11440* (POM #224349), 46.95˚ N, 119.99˚ W.

A. l.* var. *palans (M.E.Jones) M.E.Jones**—** AZ, Coconino Co.: *Christy, C.M. #493* (NY), 36.90˚ N, 111.48˚ W; *Gierisch, R.K. #4183* (NY), 35.87˚ N, 111.49˚ W; *Higgins, L.C. #5195* (NY), 36.20˚ N, 111.39˚ W; *LaDoux, D. & M. #781* (NY), 36.79˚ N, 111.61˚ W. Navajo Co.: *Welsh, S.L. #20381* (NY), 36.73˚ N, 110.25˚ W. CO, Mesa Co.: *Atwood, D. #9262* (NY), 38.52˚ N, 108.98˚ W; *Siplivinsky, V. #3379* (RM #344459), 39.06˚ N, 108.66˚ W. Montrose Co.: *Barneby, R.C. #13046* (RSA #143932), 38.23˚ N, 108.77˚ W; *Payson, E. #335* (RM #80898), 38.22˚ N, 108.57˚ W. UT, Emery Co.: *Jones, M.E. s.n.* (POM #25835), 38.82˚ N, 110.68˚ W. Garfield Co.: *Tuhy, J.S. & J.S. Holland #3127* (NY), 37.66˚ N, 111.09˚ W. Grand Co.: *Barneby, R.C. #12753* (RSA #106343), 38.69˚ N, 109.67˚ W; *Cronquist, A. #8974* (NY), 38.67˚ N, 109.50˚ W; *Thorne, K., J. Chandler & B. Franklin #4614* (RM #390257), 38.68˚ N, 109.39˚ W. Kane Co.: *Shultz, L.M. & J.S. #9931* (NY), 37.25˚ N, 111.95˚ W; *Welsh, S.L. #1687* (NY), 37.30˚ N, 111.03˚ W. San Juan Co.: *Atwood, N.D. & D. Trotter #23411* (NY), 37.85˚ N, 109.16˚ W; *Eastwood, A. s.n.* (POM #25852), 37.53˚ N, 109.23˚ W; *Higgins, L.C. & B. Welsh #13216* (NY), 37.45˚ N, 110.57˚ W; *Tuhy, J.S. #1581* (RM #359891), 38.16˚ N, 109.45˚ W.

A. l.* var. *salinus (Howell) Barneby— ID, Blaine Co.: *Hitchcock, C. L., and C. V. Muhlick #22703* (OSC #116235), 43.31˚ N, 114.07˚ W. Clark Co.: *Hitchcock, C. L., C. V. Muhlick #22786* (OSC #50314), 44.18˚ N, 112.23˚ W. OR, Harney Co.: *Acker, S. #135* (ORE #118479), 43.36˚ N, 118.97˚ W; *Jones, M. E. #25440* (OSC #181366), 43.59˚ N, 119.05˚ W; *Reveal, J.L. #2414* (OSC #206437), 43.36˚ N, 118.97˚ W. Lake Co.: *Cusick, W. M. #2618* (ORE #9616) 42.47˚ N, 120.60˚ W; *Shelly, J. S. #556* (OSC #162755), 42.45˚ N, 119.28˚ W. Malheur Co.: *Joyal, E. #446* (OSC #161837), 42.43˚ N, 118.02˚ W. NV, Elko Co.: *Tiehm, A. & L. Birdsey #5179* (RENO #2703), 41.33˚ N, 115.97˚ W. Eureka Co.: *Pinzl, A. #6563* (NESH #8216), 40.56˚ N, 116.59˚ W; *Sage, R. s.n.* (RENO #2696), 39.92˚ N, 116.55˚ W. Humboldt Co.: *Pinzl, A. #7941* (NESH #9519), 41.11˚ N, 117.70˚ W; *Pinzl, A. #7953* (NESH #9521), 41.12˚ N, 117.78˚ W. Lander Co.: *Williams, M.J. & A. Tiehm #80-46-3* (RENO #2687), 39.44˚ N, 117.23˚ W. Lincoln Co.: *Pinzl, A. #11615* (NESH #12768), 37.93˚ N, 114.14˚ W. Pershing Co.: *Pinzl, A. #9025* (NESH #11564), 40.85˚ N, 119.22˚ W; *Pinzl, A. #10870* (NESH #12043), 40.47˚ N, 119.09˚ W. Washoe Co.: *Kaye, T.N. #1257* (OSC #174506), 41.38˚ N, 119.41˚ W. UT, Beaver Co.: *Atwood, D. #24734* (NY), 38.28˚ N, 113.52˚ W; *Higgins, L., D. Atwood & S. Welsh #20500* (NY), 38.22˚ N, 113.86˚ W.

A. l.* var. *scorpionis M.E.Jones— NV, Elko Co.: *Atwood, D. & R. Burraychak #13481* (NESH #6749G), 40.63˚ N, 115.37˚ W. Lander Co.: *Williams, M.J. et al. #73-D-9A* (RENO #1517), 39.48˚ N, 117.04˚ W; *Williams, M.J et al. #76-38-10* (RENO #1477), 39.21˚ N, 117.11˚ W. Lincoln Co.: *Thorne, K.H. #1141* (RENO #1149), 38.36˚ N, 114.35˚ W. Nye Co.: *Tiehm, A. #5337* (RENO #1154), 38.35˚ N, 115.50˚ W; *Tiehm, A. & J. Nachlinger #13903* (RENO #1155), 38.67˚ N, 116.28˚ W; *Tiehm, A. & M. Williams #2767* (RENO #1519), 38.67˚ N, 116.96˚ W; *Williams, M.J. et al. #80-176-12* (RENO #1479), 39.10˚ N, 117.54˚ W. White Pine Co.: *Williams, M.J. & A. Tiehm #84-46-1* (NESH #3422G), 39.93˚ N, 114.92˚ W; *Williams, M.J & A. Tiehm #84-46-1* (RSA #338648), 39.93˚ N, 114.92˚ W.

A. l.* var. *variabilis Barneby— CA, San Bernardino Co.: *Clokey, I.W. & B.C. Templeton #5756* (POM #250339), 34.58˚ N, 117.41˚ W; *Cooper, N.C. #3379* (RSA #415572), 34.54˚ N, 117.29˚ W; *Davidson, C., B. Gustafson & R.F. Thorne #8124* (RSA #415573), 34.49˚ N, 117.17˚ W; *Ferris, R.S. & R.P. Rossbach #9497* (RSA #20229), 34.44˚ N, 116.97˚ W; *Feudge, J.B. #76* (POM #48096), 34.54˚ N, 117.29˚ W; *Helmkamp, G.K. #3020* (RSA #614561), 34.14˚ N, 115.69˚ W; *Johnson, E.R. #2958* (RSA #4867), 34.54˚ N, 117.26˚ W; *Johnston, I.M. s.n.* (POM #9125), 34.50˚ N, 117.31˚ W; *Knaus, B.J. #58* (NA) 34.57˚ N, 117.41˚ W; *McNeal, D. W. #3124* (OSC #171902), 34.44˚ N, 116.97˚ W; *Munz, P.A. #4445* (POM #8924), 34.43˚ N, 117.30˚ W; *Munz, P.A. & F. Youngberg #15179* (POM #229128), 34.53˚ N, 117.23˚ W; *Murphy, D. G. s.n.* (OSC #98704), 34.14˚ N, 116.05˚ W; *Myers, S. & J. Hirshberg s.n.* (RSA #489186), 34.11˚ N, 116.43˚ W; *Parish, S.B. #9225* (POM #24931), 34.54˚ N, 117.29˚ W; *Ripley, H.D. & R.C. Barneby #3264* (RSA #112141), 34.47˚ N, 117.29˚ W; *Ripley, H.D. & R.C. Barneby #4282* (RSA #112138), 34.54˚ N, 117.29˚ W; *Roos, J. #487* (POM #263167), 34.54˚ N, 117.29˚ W; *Spencer, M.F.*

#368 (POM #47259), 34.43˚ N, 117.30˚ W. Riverside Co.: *Hitchcock, C. L. #5958* (OSC #43360), 33.93˚ N, 115.70˚ W; *Hitchcock, C.L. #5958* (RSA #27928), 33.93˚ N, 115.70˚ W.

Appendix B. CpSSR specimens. State, county, population identifier, collection number (when available), latitude and longitude for the chloroplast simple sequence repeat sample.

Astragalus amphioxys A.Gray var. ***amphioxys***— AZ, Mohave Co.: ASAMAM 36.89° N 113.93° W.

A. iodanthus S.Watson var. ***iodanthus***— NV, Pershing Co.: ASIO, 40.72° N 119.31° W.

A. platytropis A.Gray— NV, Clark Co.: ASPL, 36.26° N 115.69° W.

A. pseudiodanthus Barneby— NV, Nye Co.: ASPS, 38.24° N 117.32° W.

A. purshii Douglas ex Hooker var. ***tinctus*** M.E.Jones— CA, Mono Co.: ASPU, 37.57° N 118.74° W.

A. utahensis (Torrey) Torrey & A.Gray— ID, Bingham Co.: ASUT_1, 43.27° N 111.98° W. NV, White Pine Co.: ASUT_2, 39.19° N 114.3° W.

Astragalus lentiginosus Douglas ex Hooker var. ***albifolius*** M.E.Jones— CA, Inyo Co.: ASLEAL-8, 36.6° N 118.1° W.

A. l. var. ***antonius*** Barneby— CA, San Bernardino Co.: ASLEAN-10, 34.35° N 117.63° W.

A. l. var. ***araneosus*** (E.Sheldon) Barneby— UT, Beaver Co.: ASLEAR_1, 38.37° N 112.95° W. Juab Co.: ASLEAR_2, 39.57° N 112.77° W. Piute Co.: ASLEAR_3, 38.32° N 112.22° W.

A. l. var. ***australis*** Barneby— AZ, Pima Co.: ASLEAUS, 32.34° N 111.11° W.

A. l. var. ***borreganus*** M.E.Jones— CA, San Bernardino Co.: ASLEBO-22, *B.J.Knaus #53*, 34.89° N 115.72° W. San Diego Co.: ASLEBO_FW, 33.31° N 116.24° W.

A. l. var. ***chartaceus*** M.E.Jones— OR, Jefferson Co.: ASLECH-35, *B.J.Knaus #97*, 44.78° N 120.9° W. Deschutes Co.: ASLECH-36, *B.J.Knaus #103*, 43.9° N 120.98° W.

A. l. var. ***coachellae*** Barneby— CA, Riverside Co.: ASLECO-1, 33.87° N 116.51° W; ASLECO-2 33.79° N 116.39° W; ASLECO-38.2 33.9° N 116.64° W.

A. l. var. ***floribundus*** A.Gray— CA, Plumas Co.: ASLEFL_DY 39.76° N 120.22° W. OR, Lake Co.: ASLEFL-26b, *B.J.Knaus #66*, 42.77° N 120.37° W.

A. l. var. ***fremontii*** (A.Gray) S.Watson— CA, Inyo Co.: ASLEFR-18, 36.48° N 118.09° W; ASLEFR-44, *B.J.Knaus #125*, 36.25° N 117.08° W; ASLEFR-45, *B.J.Knaus #128*, 37.26° N 118.15° W. San Bernardino Co.: ASLEFR-43, 35.2° N 115.34° W. NV, Esmeralda Co.: ASLEFR-46, *B.J.Knaus #129*, 37.42° N 117.61° W. Eureka Co.: ASLEFRE_1, 39.45° N 116.31° W. Mineral Co.: ASLEFRE_2, 38.2° N 117.35° W. UT, Washington Co.: ASLEFRE_3, 37.31° N 113.97° W.

A. l. var. ***idriensis*** M.E.Jones— CA, Kern Co.: ASLEID-20, *B.J.Knaus #50*, 35.35° N 119.83° W.

A. l. var. ***ineptus*** (A.Gray) M.E.Jones— CA, Mono Co.: ASLEIN-11, 38.34° N 119.63° W.

A. l. var. ***kennedyi*** (Rydberg) M.E.Jones— NV, Churchill Co.: ASLEKEN-26a, *B.J.Knaus #65*, 39.29° N 118.42° W; ASLEKEN-48, *B.J.Knaus #131*, 39.53° N 118.77° W.

A. l. var. ***kernensis*** (Jepson) Barneby— CA, Tulare Co.: ASLEKER-6, 36.23° N 118.2° W.

A. l. var. ***lentiginosus***— CA, Lassen Co.: ASLELE-12, *B.J.Knaus #136*, 40.67° N 121.19° W; ASLELE-51, *B.J.Knaus #136*, 40.41° N 120.76° W. OR, Harney Co.: ASLELE-34, 42.76° N 118.74° W. Klamath Co.: ASLELE-38, *B.J.Knaus #109*, 42.27° N 121.3° W.

A. l. var. ***maricopae*** Barneby— AZ, Maricopa Co.: ASLEMAR, 33.84° N 111.82° W.

A. l. var. ***micans*** Barneby— CA, Inyo Co.: ASLEMI-17, *B.J.Knaus #109*, 37.11° N 117.68° W.

A. l. var. ***mokiacensis*** (A.Gray) M.E.Jones— AZ, Mohave Co.: ASLEMO-1, 36.53° N 113.73° W. NV, Clark Co.: ASLEMO-2, 36.24° N 114.18° W.

A. l. var. ***nigricalycis*** M.E.Jones— CA, Kern Co.: ASLENI-21, *B.J.Knaus #51*, 35.1° N 119.4° W.

A. l. var. ***palans*** (M.E.Jones) M.E.Jones— AZ, Coconino Co.: ASLEPAL, 36.93° N 111.47° W.

A. l. var. ***piscinensis*** Barneby— CA, Inyo Co.: ASLEPIS, 37.46° N 118.4° W.

A. l. var. ***salinus*** (Howell) Barneby— NV, Humboldt Co.: ASLESA-31, *B.J.Knaus #79*, 41.12° N 117.76° W. Lincoln Co.: ASLESA_1, 38.14° N 114.72° W; ASLESA_2, 38.41° N 114.37° W. Lyon Co.: ASLESA-47, *B.J.Knaus #130*, 39.04° N 119.22° W. Pershing Co.: ASLESA-49, *B.J.Knaus #132*, 40.44° N 119° W. Washoe Co.: ASLESA-50, *B.J.Knaus #135*, 40.76° N 119.49° W. OR, Harney Co.: ASLESA-33, *B.J.Knaus #90*, 42.29° N 118.71° W. Lake Co.: ASLESA-37, *B.J.Knaus #108*, 43.32° N 121.05° W. Malheur Co.: ASLESA-32, *B.J.Knaus #84*, 42.43° N 118.08° W. UT, Iron Co.: ASLESA_3, 37.84° N 113.66° W.

A. l. var. ***scorpionis*** M.E.Jones— NV, White Pine Co.: ASLESC-1, 39.26° N 115.52° W; ASLESC-2, 39.14° N 114.96° W.

A. l. var. ***semotus*** Jepson— CA, Mono Co.: ASLESEM-13, 37.5° N 118.19° W; ASLESEM-5, 37.89° N 118.32° W.

A. l. var. ***sesquimetralis*** (Rydberg) Barneby— CA, Inyo Co.: ASLESES-16, 37.19° N 117.55° W. NV, Mineral Co.: ASLESES-27, *B.J.Knaus #72*, 38.35° N 118.11° W.

A. l. var. ***sierrae*** M.E.Jones— CA, San Bernardino Co.: ASLESI-9, 34.26° N 116.92° W.

A. l. var. ***stramineus*** (Rydberg) Barneby— NV, Clark Co.: ASLESTR, 36.79° N 114.19° W.

A. l. var. ***trumbullensis*** S.L.Welsh & N.D.Atwood— AZ, Mojave County, ASLETR, 36.26° N 113.5° W.

A. l. var. ***ursinus*** (A.Gray) Barneby— UT, Washington Co.: ASLEUR, 37.02° N 113.85° W.

A. l. var. ***variabilis*** Barneby— CA, Inyo Co.: ASLEVA-25, *B.J.Knaus #61*, 36.46° N 117.45° W; ASLEVA-41, *B.J.Knaus #117*, 36.29° N 117.98° W. Riverside Co.: ASLEVA-38.3, 33.78° N 115.32° W; ASLEVA-42, *B.J.Knaus #121*, 34.02° N 116.17° W. San Bernardino Co.: ASLEVA-3 34.61° N 116.97° W; ASLEVA-4,

34.98° N 115.65° W; ASLEVA-23, *B.J.Knaus #58*, 34.57° N 117.41° W. NV, Nye Co.: ASLEVA-29, 36.65° N 116.57° W.

A. l. var. ***vitreus*** Barneby— UT, Washington Co.: ASLEVIT, 37.18° N 113.4° W.

A. l. var. ***wilsonii*** (Greene) Barneby— AZ, Yavapai Co.: ASLEWIL, 34.69° N 111.75° W.

A. l. var. ***yuccanus*** M.E.Jones— AZ, Mohave Co.: ASLEYUC_1, 34.87° N 114.15° W; ASLEYUC_2, 35.22° N 114.16° W.

Appendix C. AFLP specimens. State, county, population identifier, collection number (when available), latitude and longitude for the amplified fragment length polymorphism sample.

Astragalus lentiginosus Douglas ex Hooker var. ***borreganus*** M.E.Jones— CA, San Bernardino Co.: ASLEBO-22, *B.J.Knaus #53*, 34.89° N 115.72° W. San Diego Co.: ASLEBO_FW, 33.31° N 116.24° W.

A. l. var. ***coachellae*** Barneby— CA, Riverside Co.: ASLECO-1, 33.87° N 116.51° W; ASLECO-2 33.79° N 116.39° W; ASLECO-38.2 33.9° N 116.64° W.

A. l. var. ***fremontii*** (A.Gray) S.Watson— CA, Inyo Co.: ASLEFR-18, 36.48° N 118.09° W; ASLEFR-40, *B.J.Knaus #125*, 36.31° N 117.66° W; ASLEFR-44, *B.J.Knaus #125*, 36.25° N 117.08° W; ASLEFR-45, *B.J.Knaus #128*, 37.26° N 118.15° W. San Bernardino Co.: ASLEFR-43, 35.2° N 115.34° W. NV, Esmeralda Co.: ASLEFR-46, *B.J.Knaus #129*, 37.42° N 117.61° W.

A. l. var. ***kennedyi*** (Rydberg) M.E.Jones— NV, Churchill Co.: ASLEKEN-26a, *B.J.Knaus #65*, 39.29° N 118.42° W; ASLEKEN-48, *B.J.Knaus #131*, 39.53° N 118.77° W.

A. l. var. ***micans*** Barneby— CA, Inyo Co.: ASLEMI-17, *B.J.Knaus #109*, 37.11° N 117.68° W.

A. l. var. ***nigricalycis*** M.E.Jones— CA, Kern Co.: ASLENI-21, *B.J.Knaus #51*, 35.1° N 119.4° W.

A. l. var. ***variabilis*** Barneby— CA, Inyo Co.: ASLEVA-25, *B.J.Knaus #61*, 36.46° N 117.45° W; ASLEVA-41, *B.J.Knaus #117*, 36.29° N 117.98° W. Riverside Co.: ASLEVA-38.3, 33.78° N 115.32° W; ASLEVA-42, *B.J.Knaus #121*, 34.02° N 116.17° W. San Bernardino Co.: ASLEVA-4, 34.98° N 115.65° W; ASLEVA-23, *B.J.Knaus #58*, 34.57° N 117.41° W. NV, Nye Co.: ASLEVA-29, 36.65° N 116.57° W.

Appendix D. *Astragalus lentiginosus* Synonymy

Astragalus lentiginosus Douglas ex Hook. 1831.
A. l.* var. *albifolius M.E.Jones 1923.
A. l.* var. *ambiguus Barneby 1964.
A. l.* var. *antonius Barneby 1945.
A. l.* var. *araneosus (E.Sheldon) Barneby 1945.
A. l.* var. *australis Barneby 1945.
A. l.* var. *borreganus M.E.Jones 1898.
A. l. var. *caesariatus* Barneby 1944 = *A. l.* var. *idriensis* M.E.Jones.
A. l. var. *carinatus* M.E.Jones 1923 = *A. l.* var. *lentiginosus* Barneby.
A. l. var. *charlestonensis* (Clokey) Barneby 1945 = ***A. l.*** var. *kernensis* (Jepson) Barneby.
A. l.* var. *chartaceus M.E.Jones 1895.
A. l.* var. *coachellae Barneby 1964.
A. l. var. *coulteri* (Bentham) M.E.Jones 1898 = *A. l.* var. *coachellae* Barneby.
A. l. var. *cuspidocarpus* M.E.Jones 1895 = *A. l.* var. *chartaceus* M.E.Jones.
A. l.* var. *diphysus (A.Gray) M.E.Jones 1895.
A. l.* var. *floribundus A.Gray 1865.
A. l.* var. *fremontii (A.Gray 1857) S.Watson 1871.
A. l.* var. *higginsii S.L.Welsh 1981.
A. l.* var. *idriensis M.E.Jones 1902.
A. l.* var. *ineptus (A.Gray) M.E.Jones 1923.
A. l.* var. *kennedyi (Rydberg) Barneby 1945.
A. l.* var. *kernensis (Jepson) Barneby 1945.
A. l.* var. *latus (M.E.Jones) M.E.Jones 1923.
A. l.* var. *lentiginosus Barneby 1964.
A. l. var. *macdougali* (E.Sheldon) M.E.Jones 1895 = *A. l.* var. *diphysus* (A.Gray) M.E.Jones.
A. l. var. *macrolobus* (Rydberg) Barneby 1945 = *A. l.* var. *salinus* (Howell) Barneby.
A. l.* var. *maricopae Barneby 1945.
A. l.* var. *micans Barneby 1956.
A. l.* var. *mokiacensis (A.Gray) M.E.Jones 1923.
A. l. var. *mokiacensis* fma. β Barneby 1945 = *A. l.* var. *ambiguus* Barneby.
A. l.* var. *multiracemosus S.L.Welsh & N.D.Atwood 2007.
A. l.* var. *negundo S.L.Welsh & N.D.Atwood 2007.
A. l.* var *nigricalycis M.E.Jones 1895.
A. l.* var. *oropedii Barneby 1945.
A. l.* var. *palans (M.E.Jones) M.E.Jones 1898.
A. l.* var. *piscinensis Barneby 1977.
A. l. var. *platyphyllidius* (Rydberg) M.Peck. 1940 = *A. l.* var. *chartaceus* M.E.Jones.
A. l.* var. *pohlii S.L.Welsh & Barneby 1981.
A. l.* var. *salinus (Howell) Barneby 1945.
A. l.* var. *scorpionis M.E.Jones 1923.
A. l.* var. *semotus Jepson 1936.

A. l.* var. *sesquimetralis (Rydberg) Barneby 1945.
A. l.* var. *sierrae M.E.Jones 1923.
A. l.* var. *stramineus (Rydberg) Barneby 1945.
A. l. var. *tehatchapiensis* (Rydberg) Barneby 1945 = *A. l.* var. *idriensis* M.E.Jones.
A. l. var. *toyabensis* Barneby 1945 = *A. l.* var. *scorpionis* M.E.Jones.
A. l. var. *tremuletorum* Barneby 1945 = *A. l.* var. *scorpionis* M.E.Jones.
A. l.* var. *trumbullensis S.L.Welsh & N.D.Atwood 1981.
A. l. var. *typicus* Barneby 1945 = *A. l.* var. *lentiginosus* Barneby.
A. l.* var. *ursinus (A.Gray) Barneby 1945.
A. l.* var. *variabilis Barneby 1945.
A. l.* var. *vitreus Barneby 1945.
A. l. var. *wahweapensis* S.L.Welsh 1978 = *A. l.* var. *diphysus* (A. Gray) M.E. Jones.
A. l.* var. *wilsonii (Greene) Barneby 1945.
A. l.* var. *yuccanus M.E.Jones 1898.

***Astragalus* L.**
Astragalus agninus Jepson 1943 = *A. l.* var. *borreganus* M.E.Jones.
A. albifolius (M.E.Jones) Abrams 1944 = *A. l.* var. *albifolius* M.E. Jones.
A. amplexus Payson 1915 = *A. l.* var. *palans* (M.E.Jones) M.E. Jones.
A. araneosus E.Sheldon 1894 = *A. l.* var. *araneosus* (E.Sheldon) Barneby.
A. arthu-schottii A.Gray 1863 = *A. l.* var. *borreganus* M.E.Jones (in part).
A. bryantii Barneby 1944 = *A. l.* var. *palans* (M.E.Jones) M.E.Jones.
A. coulteri Benth. 1848 = *A. l.* var. *borreganus* M.E.Jones.
A. coulteri var. *fremontii* (A.Gray) M.E.Jones 1895 = *A. l.* var. *fremontii* (A.Gray) S.Watson.
A. diphysus A.Gray 1849. = *A. l.* var. *diphysus* (A.Gray) M.E.Jones.
A. d. var. *albiflorus* A.Gray 1849 = *A. l.* var. *diphysus* (A.Gray) M.E.Jones.
A. d. var. *albiflorus* (A.Gray) Schoener 1974. = *A. l.* var. *diphysus* (A.Gray) M.E.Jones.
A. d. var. *latus* M.E.Jones 1893 = *A. l.* var. *latus* (M.E.Jones) M.E.Jones.
A. eremicus Sheldon 1893 = *A. l.* var. *fremontii* (A.Gray) S.Watson.
A. fremontii A.Gray 1856 = *A. l.* var. *fremontii* (A.Gray) S.Watson.
A. f. subsp. *eremicus* (Sheldon) Abrams 1944 = *A. l.* var. *fremontii* (A.Gray) S.Watson.
A. f. var. *yuccanus* (M.E.Jones) Tidestom 1941 = *A. l.* var. *yuccanus* M.E.Jones.
A. heliophilus (Rydberg) Tidestrom 1925 = *A. l.* var. *salinus* (Howell) Barneby.
A. idriensis (M.E.Jones) Abrams 1944 = *A. l.* var. *idriensis* M.E.Jones.
A. ineptus A.Gray 1864 = *A. l.* var. *ineptus* (A.Gray) M.E.Jones.
A. kernensis Jepson 1923 = *A. l.* var. *kernensis* (Jeps.) Barneby.
A. k. ssp. *charlestonensis* Clokey 1942 = *A. l.* var. *kernensis* (Jepson) Barneby.
A. latus (M.E.Jones 1893) M.E. Jones 1894 = *A. l.* var. *latus* (M.E.Jones) M.E.Jones.
A. macdougali E.Sheldon 1894 = *A. l.* var. *diphysus* (A.Gray) M.E.Jones.
A. merrillii (Rydberg) Tidestrom 1937 = *A. l.* var. *chartaceus* M.E.Jones.
A. mokiacensis A.Gray 1878 = *A. l.* var. *mokiacensis* (A.Gray) M.E.Jones.
A. nigricalycis (M.E.Jones) Abrams. 1944 = *A. l.* var *nigricalycis* M.E.Jones.
A. palans M.E.Jones 1893 = *A. l.* var. *palans* (M.E.Jones) M.E.Jones

A. p. var. *araneosus* (E.Sheldon) M.E.Jones 1895 = *A. l.* var. *araneosus* (E.Sheldon) Barneby.
A. salinus Howell 1893 = *A. l.* var. *salinus* (Howell) Barneby.
A. sierrae (M.E.Jones) Tidestrom 1937 = *A. l.* var. *sierrae* M.E.Jones 1923.
A. tehatchapiensis (Rydberg) Tidestrom 1937 = *A. l.* var. *idriensis* M.E.Jones.
A. ursinus A.Gray 1878 = *A. l.* var. *ursinus* (A.Gray) Barneby.

A. wilsonii Greene 1897 = *A. l.* var. *wilsonii* (Greene) Barneby.
A. yuccanus (M.E.Jones) Tidestrom 1935 = *A. l.* var. *yuccanus* M.E.Jones

Cystium Steven
Cystium agninum (Jepson) Rydberg = *A. l.* var. *borreganus* M.E.Jones.
C. albifolium (M.E.Jones) Rydberg 1929 = *A. l.* var. *albifolius* M.E.Jones
C. araneosum (E.Sheldon) Rydberg 1913. = *A. l.* var. *araneosus* (E.Sheldon) Barneby.
C. arthu-schottii (A.Gray) Rydberg 1929 = *A. l.* var. *borreganus* M.E.Jones.
C. cornutum Rydberg 1929 = *A. l.* var. *chartaceus* M.E.Jones.
C. eremicum (Sheldon) Rydberg 1929. = *A. l.* var. *fremontii* (A.Gray) S.Watson.
C. floribundum (A.Gray) Rydberg 1929. = *A. l.* var. *floribundus* A.Gray.
C. fremontii (A.Gray) Rydberg 1929 = *A. l.* var. *fremontii* (A.Gray) S.Watson.
C. griseolum Rydberg 1929. = *A. l.* var. *fremontii* (A.Gray) S.Watson.
C. heliophilum Rydberg 1917 = *A. l.* var. *salinus* (Howell) Barneby.
C. idriense (M.E.Jones) Rydberg 1929 = *A. l.* var. *idriensis* M.E.Jones.
C. ineptum (A.Gray) Rydberg 1905 = *A. l.* var. *ineptus* (A.Gray) M.E.Jones.
C. kennedyi Rydberg 1929 = *A. l.* var. *kennedyi* (Rydberg) Barneby.
C. kernense (Jepson) Rydberg 1929 = *A. l.* var. *kernensis* (Jepson) Barneby.
C. latum (M.E.Jones) Rydberg = *A. l.* var. *latus* (M.E.Jones) M.E.Jones.
C. lentiginosum (Douglas) Rydberg 1913 = *A. l.* var. *lentiginosus*
C. macdougali (E.Sheldon) Rydberg 1929. = *A. l.* var. *diphysus* (A.Gray)
C. macrolobum Rydberg 1929 = *A. l.* var. *salinus* (Howell) Barneby.
C. merrillii Rydberg 1929 = *A. l.* var. *chartaceus* M.E.Jones.
C. nigricalyce (M.E.Jones) Rydberg = *A. l.* var *nigricalycis* M.E.Jones
C. ormsbyense Rydberg 1929 = *A. l.* var. *floribundus* A.Gray.
C. pardalotum Rydberg 1929 = *A. l.* var. *variabilis* Barneby 1945.
C. platyphyllidium Rydberg 1929 = *A. l.* var. *chartaceus* M.E.Jones.
C. salinum (Howell) Rydberg 1917 = *A. l.* var. *salinus* (Howell) Barneby.
C. scorpionis (M.E. Jones) Rydberg 1929 = *A. l.* var. *scorpionis* M.E.Jones 1923.
C. sesquimetrale Rydberg 1929 = *A. l.* var. *sesquimetralis* (Rydberg) Barneby 1945.
C. sierrae (M.E.Jones) Rydberg = *A. l.* var. *sierrae* M.E.Jones 1923.
C. stramineum Rydberg 1929 = *A. l.* var. *stramineus* (Rydberg) Barneby 1945.
C. tehatchapiense Rydberg 1929 = *A. l.* var. *idriensis* M.E.Jones.
C. vulpinum Rydberg 1929 = *A. l.* var. *salinus* (Howell) Barneby.
C. yuccanum (M.E.Jones) Rydberg 1929 = *A. l.* var. *yuccanus* M.E.Jones

Hamosa Medikus
Hamosa amplexus (Payson) Rydberg 1917 = *A. l.* var. *palans* (M.E.Jones) M.E.Jones

Phaca Linnaeus
Phaca inepta (A.Gray) Rydberg 1900 = *A. l.* var. *ineptus* (A.Gray) M.E.Jones.
P. lentiginosa (Douglas) Piper 1906 = *A. l.* var. *lentiginosus*

Tium Medikus
Tium amplexum (Payson) Rydberg 1929 = *A. l.* var. *palans* (M.E.Jones) M.E.Jones.
T. mokiacense (A.Gray) Rydberg 1929 = *A. l.* var. *mokiacensis* (A.Gray) M.E.Jones.
T. palans (M.E. Jones) Rydberg 1929 = *A. l.* var. *palans* (M.E.Jones) M.E.Jones.
T. ursinum (A. Gray) Rydberg 1929 = *A. l.* var. *ursinus* (A. Gray) Barneby.
T. wilsonii (Greene) Rydberg 1929 = *A. l.* var. *wilsonii* (Greene) Barneby.

Tragacantha Tournefort
Tragacantha diphysa (A.Gray) Kuntze 1891. = *A. l.* var. *diphysus* (A.Gray) M.E.Jones 1895.
T. lentiginosa (Douglas) Kuntze 1891 = *A. l.* var. *lentiginosus*

Appendix E. Phylogeny of the New World *Astragalus*.

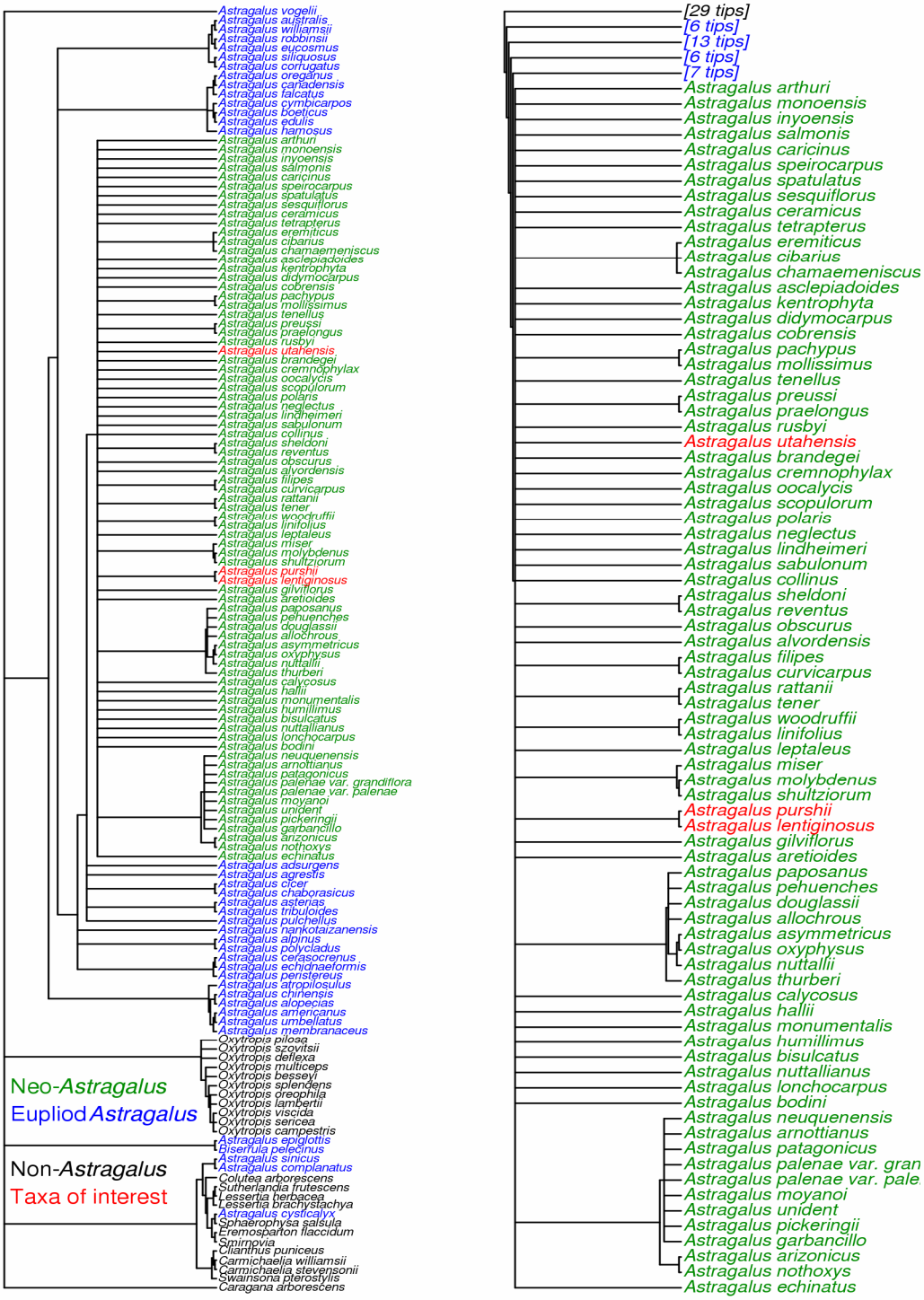


Figure E1. Reconstruction of the phylogeny of New World *Astragalus* (Wojciechowski et al. 1999) based on nuclear ITS and chloroplast *trnL* intron data. Left panel includes the entire dataset, right panel focuses on only the New World taxa.