DIVERSITY OF A DISEASE RESISTANCE GENE HOMOLOG IN *ANDROPOGON GERARDII* (POACEAE) IS CORRELATED WITH PRECIPITATION

by

MATTHEW ROUSE

B.S., Oklahoma State University, 2005

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Plant Pathology
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2007

Approved by:

Major Professor
Karen Garrett

Abstract

Ecological clines often result in gradients of disease pressure in natural plant communities, imposing a gradient of selection on disease resistance genes. We describe the diversity of a resistance gene homolog in natural populations of the dominant tallgrass prairie grass, *Andropogon gerardii*, across a precipitation gradient ranging from 47.63 cm/year in western Kansas to 104.7 cm/year in central Missouri. Since moisture facilitates infection by foliar bacterial pathogens, plants along this precipitation gradient will tend to experience heavier bacterial disease pressure to the east. In maize, the gene *Rxo1* confers resistance to the pathogenic bacterium *Burkholderia andropogonis. Rxo1* homologs have been identified in *A. gerardii* and *B. andropogonis* is known to infect natural populations of *A. gerardii*. The spatial genetic structure of *A. gerardii* was assessed from central Missouri to western Kansas by genotyping with AFLP markers. Samples were also genotyped for *Rxo1* homologs by amplifying an 810 base pair region of the leucine-rich repeat and digesting with restriction enzymes. We compared *Rxo1* homolog diversity to AFLP diversity across different spatial scales. Genetic dissimilarity based on AFLP markers was lower than would have occurred by chance at distances up to 30 m, and different prairies were more dissimilar than would have occurred by chance, but there was not a longitudinal trend in within-prairie dissimilarity as measured by AFLP markers. Dissimilarity of the *Rxo1* homologs was higher in the east suggesting the presence of diversifying selection in the more disease-conducive eastern environments.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my co-advisors Dr. Karen Garrett and Dr. Scot Hulbert for supporting me in completing this research. I would also like to thank Dr. Amgad Saleh for providing expertise and training. The research was supported by The Land Institute and The Ecological Genomics Institute at Kansas State University. The Department of Plant Pathology provided a Tillman Fellowship to support my studies and living expenses. Several excellent teachers and mentors contributed significantly to my development as a scientist during the last two years including Dr. William Bockus, Dr. Robert Bowden, Dr. John Leslie, Dr. Timothy Todd, and Dr. Barbara Valent. The Department of Plant Pathology facilitated collaborations that made this research possible resulting in thankfully-received help from individuals including Dr. Paul St. Amand, Casey Anderson, Dr. Guihua Bai, Dr. Jianfa Bai, Dr. Amy Bernardo, Shauna Dendy, Sarah Farlee, Joshua Farr, Dr. John Fellers, Erin Frank, Peter Nyori, Bruce Ramundo, Amadou Seck, Adam Sparks, Zhongwen Tang, Dr. Steve Travers, Dr. Frank White, and Dr. Chulee Yaege. I acknowledge the help of Dr. James Jurgenson, University of North Iowa, with DNA extractions. Dr. Kathleen Keeler, University of Nebraska, provided much help regarding *Andropogon gerardii* expertise. Dr. Mark Ungerer, Division of Biology, served as a member of the thesis defense committee. I thank Nick Rasmussen for help with laboratory work.

Dedication

I dedicate this thesis to my immediate family: John, Kay, Caleb, and Jesse Rouse.

CHAPTER 1 - Diversity of *Rxo1* in *Andropogon gerardii*

Introduction

Plant disease resistance varies across disease gradients (Burdon and Thrall, 1999; Dinoor, 1970; Nevo et al., 1984), but the population genetic structure of disease resistance genes across disease gradients has not been addressed in natural plant populations. Higher disease pressure may select for higher resistance gene diversity in the presence of rapid pathogen evolution. Conversely, diversity may be lower under higher disease pressure if purifying selection for useful, conserved resistance genes is operating where there is more disease pressure. In the absence of disease, diversity may increase as selection is relaxed. Or, diversity may not vary with disease pressure if the particular locus is not important for disease resistance.

Roughly one percent of the protein coding genes in plant genomes are disease resistance genes, with 207 putative resistance genes in *Arabidopsis*, 398 in *Populus*, and 535 in *Oryza* (Tuskan et al., 2006). The importance of this class of genes in plant genomes is also demonstrated by the unusually high level of selection found at disease resistance gene loci (Jiang et al., 2007). The molecular evolution of resistance genes in natural plant populations has not been described, though Burdon and Thrall (1999) have made impressive progress in understanding the spatial dynamics of phenotypic resistance and pathogenicity in the flax-flax rust pathosystem. Within-population dynamics for host-pathogen systems have been addressed through modeling (Bergelson et al., 2001; Leonard, 1994; Leonard and Czochor, 1980; Leonard, 1977; Leonard, 1969). However, assumptions in models, such as costs and benefits of resistance for plants and costs of virulence for pathogens, developed for plant-pathogen coevolution are largely not supported in real populations (Bergelson et al., 2001). This led Bergelson et al. (2001) to suggest that modeling studies should adapt to specific host-pathogen systems. Field studies of resistance gene diversity in natural populations have been helpful in identifying the various types of selection occurring at these loci (Rose et al., 2007) and indicating that different regions of resistance gene sequences are under different types of selection (Caicedo and Schaal, 2004). However, such studies focus on one or a few individuals from several populations and examine overall trends independent of ecological context. Levels of within-population resistance gene diversity for different levels of disease pressure have not been described.

In order to study resistance gene evolution within a natural population, accurate measures of genetic diversity over space are needed. Therefore, the spatial genetic structure of the population of interest must be characterized. The scale of selection for resistance can be relatively small. For example, differences in selection for herbivore resistance were found within 4 ha for oak seedlings (Sork et al., 1993). Therefore, it is important to accurately assess how genetic diversity varies over space, both for resistance genes and neutral markers. Genetic isolation by distance (Wright, 1943) may influence spatial genetic structure due to clonal growth, limited dispersal, genetic drift, selection, and density. The study of the distribution of genetic diversity over space has recently received much attention (see Manel et al. (2003) for review), with the development of models and measures of the spatial distribution of genetic diversity (Rousset, 2000; Vekemans and Hardy, 2004).

Andropogon gerardii (big bluestem), the dominant plant species in the North American tallgrass prairie ecoregion, presents a unique opportunity to address questions about the distribution of resistance gene variation because of its relatively continuous spatial distribution across a precipitation gradient in the Great Plains and the identification of a maize resistance gene homolog in *A. gerardii*.

Many plant species exist in populations with multiple cytotypes, *A. gerardii* being an important example (see Keeler (1990) for review). The mechanisms maintaining variation in ploidy and its adaptive significance are unknown. Populations of mixed ploidy may complicate estimates of genetic diversity if the cytotypes of individuals are unknown and genetic diversity varies with ploidy. *A. gerardii* populations consist predominantly of hexaploids and enneaploids with some individuals of intermediate cytotypes and a base number of chromosomes of x = 10 (Keeler, 1990; Keeler, 1992). In general, hexaploids are more common in eastern prairies whereas western prairies (west of the Missouri River in eastern Kansas and Nebraska) have populations with mixed ploidy (Keeler, 1990). *A. gerardii* ssp. *hallii* was found to display a similar distribution of hexaploids and enneaploids, in three Nebraska populations (Keller, 1992). Hexaploid *A. gerardii* tend to behave as allopolyploids in meiosis, though some secondary associations of bivalents have been reported (Norrmann et al., 1997). Meiosis in enneaploids varies greatly with univalents, bivalents, trivalents, quadrivalents, pentavalents, and hexavalents reported (Norrmann et al., 1997). This suggests that inheritance of traits in *A. gerardii* is complex. Early attempts to correlate ploidy with environmental effects such as moisture

availability and burning regime in *A. gerardii* yielded nonsignificant results (Keeler, 1992). Variation in ploidy seems to be randomly mixed within populations, often at fine scales (Keeler, 1992). Limited data from an allozyme analysis suggested that enneaploids and hexaploids cross in the field as allozyme phenotypes were found to be shared more often across cytotypes within plots, than among plots (Keeler et al., 2002). Keeler and Davis (1999) suggested that strong selection against aneuploids may occur as well as selection maintaining both enneaploids and hexaploids due to greater seed set in hexaploids and greater vegetative vigor in enneaploids in a common garden experiment. An extensive study of fitness of *A. gerardii* clones in a natural population revealed that enneaploids did contribute many viable seeds (though these seeds were rarely enneaploid), and hexaploid fitness was much greater than enneaploids or aneuploids based on frequency of good seeds per individual (Keeler, 2004). *A. gerardii* clones were spatially monitored over four years and were found to change very little (Keeler, 2004). In addition, Keeler (2004) found a turnover rate of 1.8% plant/year suggesting very low recruitment and death. The unusual presence of enneaploid and aneuploid cytotypes in the praire despite their great fitness disadvantage in this prairie remains a mystery, though Keeler (2004) suggests that when populations are reduced to a few individuals after a severe population purge, conditions may favor recruitment of the unusual cytotypes which then remain in the population for long periods of time.

Because of its importance in conservation and restoration ecology, the spatial genetic structure of *A. gerardii* has been addressed previously (Gustafson et al. 2004; Gustafson et al., 1999; Keeler, 2004). Gustafson et al. (1999) did not find a statistically significant relationship between random amplified polymorphic DNA (RAPD) similarity and geographic distance overall, as some prairies exhibited this trend whereas others did not. Most of the variation (89%) was found within populations and 11% among populations. Gustafson et al. (2004) found that genetic diversity did not differ among remnant prairies, restored prairies or cultivars. However, they did find that local remnant and restored populations were genetically different than non-local remnants and cultivars, suggesting that location of seed selected for restoration purposes may be important even though genetic diversity of restored prairies was not reduced compared to natural prairies.

Burkholderia andropogonis has been observed to cause low levels of bacterial stripe disease on *A. gerardii* at Konza Prairie Biological Station (KPBS) in Kansas (Morgan, 2003). *B.*

andropogonis also infects sorghum and maize in warm and humid areas (Muriithi and Claflin, 1997), and tends to be more common in relatively humid eastern Kansas than relatively dry western Kansas (L. E. Claflin, personal communication). Zhao et al. (2004) identified a maize locus, *Rxo1*/*Rba1*, that confers resistance to pathogenic *B. andropogonis* and non-pathogenic *Xanthomonas oryzae* pv. *oryzicola*. A single gene, *Rxo1*, conferred resistance to both pathogens (Zhao et al., 2005). Zhao et al. (2005) also transformed *Rxo1* into rice and found that resistance to the rice pathogen, *X. oryzae* pv. *oryzicola*, was maintained. Experiments failed to find a direct protein-protein interaction between *Rxo1* and *AvrRxo1* (type III bacterial effector), suggesting *Rxo1* may act as a guard (see Dangl and Jones (2006) for review of resistance genes acting as guards). The conserved function of *Rxo1* in recognizing two different pathogens supports this hypothesis. A homolog of *Rxo1* is transcribed in *A. gerardii* (S. Hulbert, data not shown).

It is well established that the probability of foliar infection tends to increase with increasing precipitation and humidity (Huber and Gillespie, 1992). Net primary productivity also increases by a factor of 5.5 from the Shortgrass Steppe Long Term Ecological Research Site in northeastern Colorado to KPBS in Kansas (Lane et al., 2000). Alexander et al. (2007) found decreased disease (smut and rust) on *Carex blanda* field and herbarium specimens from western Kansas compared to eastern Kansas. This trend was attributed to the drier conditions in western Kansas providing an environment less suitable for disease, the fact that western Kansas was at the edge of the range for *C. blanda*, and that western populations were smaller and more isolated, decreasing the potential for successful dispersal.

A. gerardii exists across a precipitation gradient from east to west, resulting in a gradient of disease pressure for many pathogens. This provides an opportunity to study the effect of environment on resistance gene diversity. Measuring the changes in *Rxo1* homolog diversity relative to amplified fragment length polymorphism (AFLP) diversity across this gradient will reveal how selection at the *Rxo1* homolog varied with disease pressure. The spatial genetic structure of *A. gerardii* based on AFLP markers provides a measure of genome-wide diversity as a context for *Rxo1* homolog diversity at different scales. In this study, our first objective was to determine how cytotype influenced genetic diversity within a population of *A. gerardii* of known cytotype in Colorado. Our second objective was to determine genome-wide diversity in *A. gerardii*, using AFLP markers, at multiple scales within a prairie and between five prairies across the precipitation gradient. This provides perspective for tallgrass prairie conservation and

restoration, and a reference for evaluation of resistance gene diversity across the same gradient. Our final objective was to determine diversity in *Rxo1* homologs in *A. gerardii* for the same individuals.

Materials and Methods

Tissue Collection

One hundred tissue samples were collected from each of five prairies along a precipitation gradient in the central United States (Table 1.1). Each plant sample consisted of a 15 cm basal cutting made from two leaves (with no visible disease) for each tissue sampled. These two cuttings were then placed in a sealed plastic bag with 15 ml of silica gel (Demis Products, Lithonia, GA and Miracle Coatings, Anaheim, CA) for storage in bags with silica gel for up to one year. Percent cover of *A. gerardii* was similar at KPBS, The Land Institute (TLI), and Wilson Lake (about 50%), whereas percent cover was very high at Tucker prairie (about 75%). At Smoky Valley Ranch, *A. gerardii* occurred in dense patches only in relatively moister areas such as in drainages and at the base of mesas. Within each prairie, four 40 m transects were established with five sampling points at 10 m intervals (Fig. 1.1). At each sampling point, two leaves were collected at the point, itself, 1.3 m in each direction along the transect, and 1.3 m in each direction perpendicular to the transect, for a total of 5 samples per point and 25 samples per transect. This hierarchical sampling method was employed to allow comparison of genetic diversity at multiple spatial scales.

Tissues were also collected from 65 clones of known ploidy in Boulder, Colorado, that had been mapped and characterized by Keeler (2004). Plants sampled were in plots 28, 52, 61, and 102 established by Jane and Carl Bock in the City of Boulder Open Space and Mountain Parks. One sample of *Schizachyrium scoparium* and two of *A. gerardii* ssp. *hallii* were collected from the United States Department of Agriculture (USDA) Natural Resources Conservation Service (NRCS) Plant Materials Center (Manhattan, Kansas). *S. scoparium* was used as the outgroup in clustering analyses.

DNA Extraction

DNA was extracted according to Doyle and Doyle (1987) with several modifications. For each tissue sample, several 2 cm cuttings were placed in a 1.5 ml micro-centrifuge tube and

ground to a fine powder with a plastic peg and liquid nitrogen. Immediately after grinding, samples were placed in a –20 C freezer for up to 3 hours. Eight-hundred µl of 65 C 2X CTAB:β-mercapto-ethanol buffer (99:1 v/v) were added to each tube and mixed with a pipette tip. To mix the extraction buffer with ground plant material, the tubes were then inverted 10 times, placed in a 65 C water bath for 5-10 min., inverted 10 times again, and placed in a 65 C water bath for 20-25 min. (30 min. total). Four-hundred µl of chloroform:iso-amyl alcohol (24:1 v/v) were added to each tube and tubes were gently inverted for 2-3 min.. Tubes were then placed in a micro-centrifuge and spun at 13,400 rcf for 5 min. Five-hundred µl of the aqueous phase were then transferred to a clean micro-centrifuge tube. To precipitate the DNA, 500 µl of iso-propanol were added to each tube. Each tube was then inverted 10 times and set aside for 5-10 min. to allow nucleic acids to precipitate. Tubes were then placed in a micro-centrifuge and spun at 9,300 rcf for 5 min. The tubes were inverted to decant the aqueous/alcohol mixture and placed on a clean paper towel for 5 min. Then, 600 µl of TE buffer (100mM Tris HCl pH 8.0 and 1mM EDTA pH 8.0) were added to each tube. Tubes were placed in a 4 C refrigerator overnight. Tubes were flicked and briefly micro-centrifuged to re-suspend the nucleic acids. Three-hundred µl of phenol:choloroform:iso-amyl alcohol (25:24:1 v/v/v) were added to each tube. Tubes were shaken up and down by hand for a few seconds. Tubes were then placed in a micro-centrifuge and spun at 13,400 rcf for 5 min. Four-hundred-fifty µl of the aqueous phase were then transferred to a clean micro-centrifuge tube. Four-hundred-fifty µl of chloroform:iso-amyl alcohol (24:1 v/v) were added to each tube. Tubes were shaken up and down by hand for a few seconds. Tubes were then placed in a micro-centrifuge and spun at 13,400 rcf for 5 min. Two-hundred-fifty µl of the aqueous phase were then transferred to a clean micro-centrifuge tube. One µl of RNAse A (Sigma-Aldrich, 2639 kunits/ ml) was added to each tube. Tubes were flicked and briefly centrifuged before being placed in a 37 C incubator for 30 min. To precipitate DNA, 250 µl of iso-propanol were added to each tube. Each tube was then inverted 10 times and set aside for 5-10 min. at 25 C to allow DNAs to precipitate. Tubes were then centrifuged at 9,300 rcf for 10 min. The tubes were inverted to decant the aqueous/alcohol mixture and placed on a clean paper towel for a few min. 1 ml of 70% ethanol at -20 C was added to each tube to wash DNA pellets. Tubes were then centrifuged at 9,300 rcf for 5 min. The tubes were inverted to decant the ethanol and placed on a clean paper towel for a few min. to dry the DNA pellets. Any remaining ethanol was removed with a pipette. Then, each DNA pellet was dissolved in 50 µl of

TE buffer was added to each tube. Tubes were placed in a 4 C refrigerator overnight. Tubes were flicked and briefly micro-centrifuged to re-suspend the DNAs. One µl of each DNA solution was assayed using a 1% agarose gel submerged in 0.5X TBE buffer to determine DNA concentrations in comparison to known DNA concentrations of λDNA (New England Biolabs) digested with *Hind*III (New England Biolabs) using a Gel-Doc EQ gel-reading system (Biorad Laboratories, Inc.). DNA solutions were then stored at –20 C.

Amplified Fragment Length Polymorphism (AFLP) Genotyping

AFLPs have become a standard molecular marker with diverse applications (see Meudt and Clarke, 2007 for review). AFLP fingerprinting was performed according to Vos et al. (1995) with many modifications. All primers were ordered from Integrated DNA Technologies. The digestion and ligation reactions were combined with the following components in a 21µl reaction: 2 µl genomic DNA, 0.25 µl *Eco*R1 (Promega, 12 units/ µl), 0.16 µl *Mse*1 (New England Biolabs, 10 units/ µl), 0.27 µl T4 DNA ligase (Promega, 3 units/ µl), 2 µl 10X H buffer (Promega), 2 µl ligase buffer (Promega), 0.4 µl *Eco*R I adapter mix (5 ρm/µl CTCGTAGACTGCGTACC and 5 ρm/µl AATTGGTACGCAGTCTAC), 0.4 µl *Mse*1 adapter mix (50 ρm/µl GACGATGAGTCCTGAG and 50 ρm/µl TACTCAGGACTCAT), and 13.52 µl sterilized distilled water. Genomic DNA amount was not kept constant across samples as Trybush et al. (2006) found that variation in DNA template did not affect AFLP profiles in the range of 12.5-500 ng. The digestion-ligation reactions were left overnight at 25 C. The preamplification reaction was performed with the following components in a 40 µl reaction: 10 µl digestion-ligation template:distilled water (1:9 v/v), 8 µl 5X PCR buffer (Promega), 4 µl 25mM $MgCl_2$, 1.6 µl 2mM dNTPs, 0.76 µl *Eco*R1-A primer (100ng/µl AGACTGCGTACCAATTCA), 0.76 µl *Mse*1-C primer (100ng/µl GATGAGTCCTGAGTAAC), 0.15 µl Gotaq Flexi DNA Polymerase (Promega, 5 units/ µl), and 14.73 µl sterilized distilled water. The preamplification reactions were performed on a MJ Research PTC-200 thermocycler with 1 min. at 94 C followed by 30 cycles of 30 seconds at 94 C, 1 min. at 56 C, and 1 min. at 72 C. Sixteen primer pairs were tested for feasibility for selective amplification. We selected two primer pairs: *Eco*R1-AAA/*Mse*1-CTG and *Eco*R1-ACC/*Mse*1-CTG based upon number of fragments and their distribution. *Eco*R1-AAA was labeled with the fluorescent dye 6FAM and *Eco*R1-ACC was labeled with the fluorescent dye HEX. The selective amplification reactions

were performed with the following components in a 20.5 µl reaction: 1.5 µl preamplification reaction template:distilled water (1:19 v/v), 4 µl 5X PCR buffer (Promega), 2 µl 25mM $MgCl_2$, 2 µl 2mM dNTPs, 2 µl 50 ng/µl *Eco*R1-selective primer, 3 µl 50 ng/µl *Mse*1-selective primer, 0.2 µl Gotaq Flexi DNA Polymerase (Promega, 5 units/µl), and 5.3 µl distilled water. The selective amplification reactions were performed on a MJ Research PTC-200 thermocycler with 2 min. at 95 C followed by 13 cycles of 30 seconds at 65 C (-0.7 C per cycle), 90 seconds at 72 C, and 30 seconds at 94 C, followed by 30 seconds at 94 C, followed by 23 cycles of 30 seconds at 56 C, 90 seconds at 72 C, and 30 seconds at 94 C, followed by 30 seconds at 56 C and 5 min. at 72 C. The two selective amplification products (for 6FAM and HEX primer pairs) were combined as follows per 10 µl dilution: 1 µl *Eco*R1-AAA/*Mse*1-CTG selective amplification template:1 µl *Eco*R1-ACC/*Mse*1-CTG selective amplification template:distilled water (1:1:4 v/v/v), 8.8 µl formamide, 0.2 µl GeneScan 500 Liz size standard (Applied Biosystems). This dilution was then incubated in a thermocycler for 5 minutes at 95 C. Plates were submitted to the USDA Small Grain Genotyping Laboratory (Manhattan, Kansas) for AFLP analyses (ABI 3100 DNA Analyzer, Applied Biosystems). AFLP data was analyzed with GeneMarker version 1.6 with the manufacturer's suggested settings except that the "smooth" option was selected and "reject" and "check" for peak evaluation were set to zero.

Sixty-three AFLP polymorphic fragments were scored for 6FAM-labeled *Eco*R1-AAA/*Mse*1-CTG and 32 polymorphic fragments were scored for HEX-labeled *Eco*R1-ACC/*Mse*1-CTG. Sample sizes for the various prairies can be found in table 1.1. Additionally, we genotyped one *Schizachyrium scoparium* individual and two *A. gerardii* ssp. *hallii* individuals. Statistical analyses were conducted using R programming software (The R Development Core Team, 2007).

Seven replicates of 15 DNA extracts were run through the AFLP protocol, separately, in order to estimate repeatability of the AFLP fragments. Vos et al. (1995) suggested that different banding patterns from the same individual could be due to incomplete digestion of separate samples, not necessarily variability in DNA extract. AFLP peaks with low repeatability were removed from final analyses. We found AFLP peak repeatability to be 90.6%. Much of the variation in repeatability of peaks was due to one or two replicates of each DNA extract being disproportionately different compared to the rest of the replicates (data not shown). Overall, replicates for the repeatability test, run on the same plate as random arrangements of other

samples, resulted in poorer quality peaks than other samples (data not shown). This may be due to more frequent handling and thawing for the samples that were repeatedly analyzed. Therefore, the 90.6% band repeatability average is likely an underestimate of the band repeatability for most samples.

Rxo1 *Homolog Identification*

Homologs of the *Zea mays* disease resistance gene *Rxo1* (GI:60615303) were identified in *Andropogon gerardii* by using the following conserved primer pairs. The following primer pairs were used: CTCCTGAGTTACGTCAGTGTG and CAGTGTCTTCAAAGCTGCACGC, GGCCATGCAGCTTAGAAGAC and ATCGAGGCACAAAAGCCTAA, GCAGAGAGGAACAGCTTTGG and CCCCTGTGGGAACTTCACTA, GGAAACAATGAGGCAATGCT and AGGAACCAGTCTGCTTGGAA, and TTCCTGCAAACCGAAGTACC and TTCCCTTTTGAATGCTGCTT. The PCR amplifications were composed of the following components per 25 µl reaction: 1 µl genomic DNA, 4 µl 5X PCR buffer (Promega), 2 µl 25mM $MgCl_2$, 2.5 µl 2mM dNTPs, 1 µl 25 nM/ml forward primer, 1 µl 25 nM/ml reverse primer, 0.25 µl Gotaq Flexi DNA Polymerase (Promega, 5 units/ µl), and 13.25 µl distilled water. The amplifications were performed on a MJ Research PTC-200 thermocycler with 5 min. at 95 C followed by 30 cycles of 30 seconds at 95 C, 30 seconds at 55 C, and 90 seconds at 72 C, followed by 1 min. at 72 C. Presence of amplification products were checked by gel electophoresis. PCR amplicons were purified using the QIAquick PCR purification kit and protocol (Qiagen Sciences). Concentrations of purified PCR products were checked with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies). Since *Andropogon gerardii* is polyploid it was necessary to clone amplified fragments of the *Rxo1* homolog in order to avoid simultaneous sequencing of different alleles (data not shown). Amplified PCR products were cloned with the Qiagen PCR CloningPlus kit and protocol (Qiagen Sciences) and plasmids were isolated with the QIAprep spin minprep kit and protocol (Qiagen Sciences, Maryland). Plasmid concentrations were checked with a Nanodrop ND-1000 spectrophotometer (NanoDrop Technologies). If concentrations were not greater than 200 ng/µl, tubes with purified product were set in a hood in order to evaporate sufficient amounts of water in order to increase the concentration to greater than 200 ng/µl. 10 µl of 200-300 ng/µl plasmid DNA were submitted per sample to the USDA Small Grain Genotyping Laboratory (Manhattan,

Kansas) for sequencing using SP6 and T7 primers. Sequences were aligned with GeneMapper Software version 3.5 (Applied Biosystems).

Rxo1 ***Homolog Amplification and Restriction Enzyme Digestion***

Based upon sequences of *Rxo1* homolog clones from a wide geographic extent, we picked the conserved primer pair AGATTCTCGACGAGTTGCTGTGCT and AGCCTAAGAAGCCCATTTCCGTGA to amplify a 810 base-pair fragment towards the 3' end of *Rxo1* (homologous to the maize leucine-rich-repeat region of the gene). This primer pair successfully amplified fragments from all DNA extracts following the same protocol for *Rxo1* homolog identification section with the exception that the annealing temperature was set to 60 C.

Four restriction enzymes were used to perform a restriction fragment length polymorphism (RFLP) assay for the amplified fragments: *Bsa*M1 (Promega), *Mse*1 (New England Biolabs), *Rsa*1 (Promega), and *Taq*1 (Promega). The enzymes were used to digest the amplified fragments individually in four different reactions. For *Mse*1, 5 µl of PCR products were combined with 0.25 µl (10u/µl) enzyme, 1.5 µl 10X acetylated BSA (Promega), 1.5 buffer 2 (New England Biolabs), and 7 µl distilled water. The mixture was incubated at 37 C for 12 hours. For *Rsa*1, buffer C (Promega) was used instead of buffer 2. For *Taq*1, buffer E (Promega) was used and the mixture was incubated at 65 C for 2 hours. For *Bsa*M1, buffer D (Promega) was used, 0.125 µl enzyme (20u/µl) was used, and the mixture was inculated at 65 C for 2 hours. All of the product for each digestion was loaded into a 2% agarose gel submerged in 0.5X TBE buffer which ran for 2 hours at 95 volts. RFLP data were collected by scoring for presence or absence of fragments for each individual.

Five fragments were scored for the *Bsa*M1 digestion of the amplified homolog of *Rxo1*, nine for *Mse*1, nine for *Rsa*1, and 14 for *Taq*1. This resulted in 37 markers for 100 samples in each of the five prairies and 63 samples from Boulder, Colorado. We found the RFLP markers to be greater than 99% repeatable. RFLP data from Boulder, Colorado were included in the dataset for calculating number and frequency of haplotypes, though fewer plants were collected from Boulder.

Relatedness of Individuals of Different Ploidy

In order to determine any subdivision among different ploidy levels, we constructed a UPGMA derived bootstrap consensus tree based on AFLP data for the plants of known ploidy

from Boulder, Colorado. We conducted the bootstrap analysis with 1000 iterations. The tree was constructed with PAUP 4.0 (Sinauer Associates) and viewed with Treeview 1.6.6 (Page, 1996).

Dissimilarity and Geographic Distance

Scale of sampling is important in determining spatial autocorrelation in clonal species (Hammerli and Reusch, 2003). To test for spatial structure at different scales using the AFLP and *Rxo1*-RFLP datasets, a similarity matrix with an entry for each pairwise comparison of ramets from all prairies was calculated for each of the two datasets. We used the simple match coefficient to calculate similarity (the number of matches, in terms of shared absence or presence of a peak, divided by the total number of possible matches). Dissimilarity was calculated simply by subtracting similarity from one. Kosman and Leonard (2005) argue that the simple match coefficient is the most appropriate index for polyploid species. Plants from Colorado were not included in this analysis as they were collected at different spatial scales, making comparison to other prairies difficult. The mean dissimilarities of all ramets at a particular distance from each other were used to calculate the overall mean dissimilarity for that distance. For example, since there are five plants in the first transect subgroup of a transect and five in the second transect subgroup, the mean of all of the comparisons among these two transect subgroups was used as one datum in calculating the overall 10 meter mean dissimilarity (see Figure 1.1). Mean dissimilarity was calculated for 1.77539 m (average within-subgroup distance), 10 m, 20 m, 30 m, 40 m, 2139 m (average distance between transects), and 473025 m (average distance between prairies).

We calculated 95% confidence intervals for each distance to test whether the mean dissimilarity at that distance was significantly different than expected by chance. The first null hypothesis used for the randomization tests was that each pairwise measure of dissimilarity between ramets was equally likely to be observed for ramets any distance apart. Following this null hypothesis, the plant identification numbers were randomized along the axis of the similarity matrix and means were calculated again for each distance for each randomization. The plant identification numbers were randomized instead of the similarity values in the matrix to preserve the relevant structure of the dataset. 1000 permutations yielded 1000 means under this null hypothesis for each distance. We used the 25th and 9975th ordered means as the 95% confidence intervals for dissimilarity at each distance under this null hypothesis of no difference in

dissimilarity at any distance. This method was used for the RFLP similarity matrix and the AFLP similarity matrix. If the observed mean for a particular distance was above or below the confidence interval, that mean was considered significantly greater or lesser than expected by chance under the null hypothesis, indicating nonrandom population structure.

In addition to overall AFLP and RFLP dissimilarity across distances, we calculated this relationship for each prairie with separate similarity matrices. Mean dissimilarity values and confidence intervals were calculated for each prairie as above, for each distance measured within a prairie. The average among-transect distance is unique to each prairie: 300 m for Tucker prairie, 908 m for Konza prairie, 2886 m for TLI, 2480 m for Wilson Lake, and 710 m for Smoky Valley Ranch.

To address our third objective, we examined *Rxo1* and AFLP dissimilarity across longitude and precipitation. We used R to perform linear regression analyses of both *Rxo1* and AFLP among-transect means, with longitude and precipitation as predictors in separate analyses. Mean dissimilarity for two random, independent pairs of transects was calculated. *Rxo1* dissimilarity was compared to AFLP dissimilarity across both longitude and precipitation as a control.

Results

Samples collected in Boulder, Colorado, were from clones of known ploidy. Significant clustering of *A. gerardii* clones in the bootstrap consensus tree was rare with only one cluster of two individuals with greater than 95% bootstrap support in addition to the well supported cluster separating *A. gerardii* clones from *S. scoparium* (Fig. 1.2). In fact, this significant cluster included individuals of differing ploidy. The remaining Boulder clones were not segregated by ploidy. The two *A. gerardii* ssp. *hallii* clones did not cluster together. There was no difference in average fragment number among cytotypes. Average number of fragments was 59.37 (n=38, 58.76-59.97 95% confidence interval) for hexaploids and 59.33 (n=21, 58.25-60.42 95% confidence interval) for enneaploids. Similarly, the average number of *Rxo1*-RFLP fragments was 10.71 (n=38, 10.59-10.83 95% confidence interval) for hexaploids and 10.62 (n=21, 10.41-10.83 95% confidence interval) for enneaploids.

Mean pairwise dissimilarity based on AFLP markers increased with increasing distance between individuals (Fig. 1.3). Samples were significantly more dissimilar in AFLP pattern than

expected by chance up to 30 m, and marginally significantly more dissimilar than expected by chance at 2139 m. The among-prairie mean dissimilarity was marginally higher than expected by chance. The confidence intervals fluctuate extensively across the pairwise sampling distances we evaluated, with an especially wide interval at 40 m. This was due to the sample size of means used to calculate the mean dissimilarity, since there are relatively few 40 m comparisons (one set of comparisons per transect).

To tease apart the components of the AFLP dissimilarity by distance relationship, mean dissimilarity was examined at four scales for each prairie. Figure 1.4 outlines these four scales by prairie, across longitude. Smoky Valley Ranch and KPBS exhibited large differences in mean AFLP dissimilarity at different scales, whereas Tucker prairie exhibited similar levels of mean AFLP dissimilarity at all scales. Four of the prairies were similar in mean AFLP dissimilarity, while AFLP dissimilarity at KPBS was lower.

Mean *Rxo1* dissimilarity was significantly less than expected up to 30 m, not significant at 40 m, significantly less dissimilar than expected at 2139 m, and significantly more dissimilar than expected at 473025 m (Fig. 1.5). This pattern across distance was not as uniform as the AFLP pattern, but was roughly similar. Mean *Rxo1* dissimilarity by prairie, at different scales, across longitude, increased strikingly with increasing longitude (Fig. 1.6). As for mean AFLP dissimilarity, prairies varied in their mean RFLP dissimilarity at different distances.

The relationship between mean *Rxo1* dissimilarity and longitude was clear (Fig. 1.6). However, mean *Rxo1* dissimilarity must be examined in the context of mean AFLP dissimilarity to determine whether the pattern in dissimilarity across longitude for *Rxo1* is different from the pattern for the genome as a whole. At Tucker prairie, AFLP mean dissimilarity remained relatively low but not significant at all scales (Fig. 1.7), whereas *Rxo1* mean dissimilarity was less dissimilar than expected by chance up to 10 m, and marginally more dissimilar than expected by chance at 40 m (Fig. 1.8). At KPBS, AFLP mean dissimilarity was less than expected up to 20 m and was greater than expected at 908 m (Fig. 1.9). *Rxo1* mean dissimilarity remained less dissimilar than expected only within subgroups and was more dissimilar than expected at 40 m (Fig. 1.10). At TLI, AFLP mean dissimilarity remained less dissimilar than expected up to 20 m (Fig. 1.11) whereas *Rxo1* mean dissimilarity remained less dissimilar than expected up to 10 m (Fig. 1.12). At Wilson Lake, AFLP mean dissimilarity was less dissimilar than expected only at the subgroup scale (Fig. 1.13) whereas *Rxo1* mean dissimilarity was less

dissimilar than expected at 1.77539 m, 30 m, and 40 m and was more dissimilar than expected at 2480 m (Fig. 1.14). At Smoky Valley Ranch, AFLP mean dissimilarity was less dissimilar than expected up to 10 m and more dissimilar than expected at 710 m (Fig. 1.15) whereas *Rxo1* mean dissimilarity remained less dissimilar than expected up to 20 m and more dissimilar than expected at 710 m (Fig. 1.16).

In general, *Rxo1* mean dissimilarity in the western prairies (Wilson Lake and Smoky Valley Ranch) was lower relative to AFLP dissimilarity, *Rxo1* mean dissimilarity closely resembled AFLP mean dissimilarity at TLI (intermediate longitude), and *Rxo1* mean dissimilarity was more dissimilar relative to AFLP dissimilarity in the eastern prairies (KPBS and Tucker prairie; though strikingly not so at small scales for Tucker prairie; Figs. 1.7 and 1.8).

At the among-transect scale, *Rxo1* dissimilarity was positively correlated with both longitude (Fig. 1.17) and precipitation (Fig. 1.19). There was no correlation between AFLP dissimilarity and either longitude (Fig. 1.18) or precipitation (Fig. 1.20), indicating that the change in *Rxo1* dissimilarity across longitude and precipitation is distinct from genome-wide variation.

Discussion

A. gerardii is often found in populations of mixed ploidy (Keeler, 1990), and our results indicate that there is gene flow among cytotypes. Our clustering analysis revealed that individuals of a particular ploidy level were not more closely related to each other than to individuals of other ploidy levels. This is consistent with allozyme data suggesting that individuals of dissimilar cytotype from the same plot were more similar than individuals of the same cytotype from different plots (Keeler et al., 2002). We also found that *A. gerardii* ssp. *hallii* (sand bluestem) individuals did not cluster separately from *A. gerardii* individuals. Though *A. gerardii* ssp. *hallii* is often considered a different species than *A. gerardii* (*Andropogon hallii*), our data suggest that there is gene flow between the two. We anticipated that the average number of AFLP fragments in enneaploids would be greater than the average number of fragments in hexaploids, since enneaploids have 50% more genomic DNA and AFLPs are dominant markers. However, we found no difference. *Rxo1*-RFLP band number was also the same across cytotypes. Since we did not know the cytotypes of samples outside the Boulder population, we could not take proportion cytotype into account in population comparisons. The equal number of bands in

enneaploids and hexaploids also suggests that our comparisons of population similarity were more straightforward than they might otherwise have been.

The origin and mechanism for maintenance of enneaploids is not known since enneaploids have not been found to have a fitness advantage in the field (Keeler, 2004) and crosses within or among any cytotype very rarely yield enneaploid individuals (Norrmann et al., 1997; Norrmann et al., 2003; Keeler, 2004). Since enneaploids were not found to have an increase in heterozygosity, our data suggest production of enneaploids through first division restitution of a hexaploid gamete. This process may be mediated by temperature, as chromosome doubling has been observed at high temperatures for maize and other species of the tribe Triticeae (Dorsey, 1937; Randolph, 1932). Such temperature-dependent polyploidy may explain the origin and maintenance of enneaploids and explain the adaptive role of populations with mixed cytotypes, since enneaploids seem to be more vegetatively fit in marginal environments where heat stress is more common (Keeler, 2004). Experiments testing the hypothesis of temperature-dependent cytotypes during meiosis are needed.

A previous study of *A. gerardii* identified decreasing RAPD similarity with increasing distance within a subset of populations, though this trend was not consistent among prairies (Gustafson et al., 1999). Our results demonstrate spatial genetic structuring among prairies; individuals were more diverse than expected by chance at the among-prairies level (Fig. 1.3). We did not find spatial genetic structure in Tucker prairie (Fig. 1.7) based on AFLP data. Tucker prairie is a 65 hectare remnant amidst cropland. As the other prairies are connected to large contiguous tracts of prairie, recent gene flow at Tucker prairie may be different. Or, alternatively, clone size may be smaller at Tucker prairie. There was a general trend of increasing AFLP dissimilarity with increasing distance for all other prairies.

Plant density was a major factor influencing spatial genetic structure in several species according to Vekemans and Hardy (2004). Chung and Epperson (2000) found clonal structure in the tree *Eurya emarginata* at up to 19 m. Calderon et al. (2007) found spatial genetic structure due to both clonality and limited dispersal of larvae at 50-60 cm for ramets and 30-40 cm for genets in a marine sponge, *Crambe crambe*. Genets 1 m apart were as different, on average, as genets hundreds of km apart. The data presented in this study exhibit similar spatial genetic structure at relatively small scales.

Limits to repeatability of AFLP markers prevented us from accurately determining whether sampled individuals were members of the same clone. Though microsatellite markers would likely resolve clonal structure, these have not yet been developed for *A. gerardii.* AFLP markers do not require the resources needed for microsatellites, and are suitable for our objectives. As average clone size for *A. gerardii* has been estimated at 3.20 m^2 at KPBS (Keeler et al., 2002) and 0.20 m^2 at Boulder (Keeler and Davis, 1999), we expect that many of the samples within the same subgroup were genets of the same clone. This complicates estimation of among-clone genetic diversity at small scales since more than one sample may have been taken from the same clone. As a result, our estimates of spatial genetic structure incorporate both clone size and genetic diversity among clones. Inclusion of clones in the estimate of spatial genetic structure gives a more accurate representation of the genetic mosaic of *A. gerardii* over space. For *A. gerardii* in our experiment, multiple sampling of the same clone was likely to be rare beyond the subgroup scale, and unlikely to be a factor in our comparisons among transects.

Clone size contributes to the spatial genetic structure of clonal organisms and is thought to be determined by interclonal competition, frequency of disturbance, time since establishment, time since disturbance, and site quality (Suvanto and Latva-Karjanmaa, 2005). Hammerli and Reusch (2003) found that three components of clone structure contribute to spatial autocorrelation in clonal species: ramets, clone fragments, and entire clones. Much of the literature on spatial genetic structure involving clonal species describes aspen systems (Kemperman and Barnes, 1976; Suvanto and Latva-Karjanmaa, 2005) or aquatic grass systems (Alberto et al., 2005; Hammerli and Reusch, 2003; Ruggiero et al., 2005). The contribution of clones to spatial genetic structure may vary across populations. Organisms more commonly reproduce asexually at higher altitudes and in resource-poor environments, possible due to a fitness advantage of asexual reproduction in well-adapted individuals (Peck et al., 1998). Dense stands of *Acropora palmata*, a coral, were shown by Baums et al. (2006) to have higher genotypic richness and diversity due to increased asexual recruitment in dense stands. Most of the spatial genetic structure in *Populus tremula*, European aspen, was found to be due to the spread of clones as analyses considering one ramet for each clone displayed only marginal spatial genetic structure (Suvanto and Latva-Karjanmaa, 2005). Subterranean disturbance of clones seems to reduce spatial genetic structure in the aquatic grass, *Cymodocea nodosa* (Alberto et al., 2005). Since clone size in *A. gerardii* is relatively small, Keeler et al. (2002) propose a

model of *A. gerardii* colonization after death due to drought or other natural disasters, followed by seed recruitment and clonal spread. The limited clone sizes at KPBS seem to support this model. Keeler et al. (2002) suggest that clone size in Eastern prairies may be much greater than in Western prairies due to greater clonal reproduction in the less variable environmental conditions of the east. Variance in clone size in different populations can confound comparisons of genetic diversity among populations. Balloux et al. (2003) developed models which demonstrate that increasing clonal reproduction increases allelic diversity and decreases genotypic diversity. Clones can now be more readily identified using molecular techniques such as microsatellites, compared to morphological measures (Suvanto and Latva-Karjanmaa, 2005). Separation of clones based on morphological assumptions is particularly difficult in organisms where clones intermingle (Ruggiero et al., 2005). Aquatic grass clones were found to be up to 60 m wide based on microsatellite identity (Ruggiero et al., 2005). A thorough analysis of the structure of clones in *A. gerardii* across this environmental gradient would elucidate the adaptive role of clonal growth.

Many of the AFLP fragments were found at relatively small scales, with among transect comparisons being only marginally less dissimilar than expected by chance (Fig. 1.3). This means that one prairie is likely to contain a large portion of the neutral variation in *A. gerardii* across Kansas and Missouri. This finding is consistent with data from Arkansas populations where 11% of the variation was found among prairies, and 89% was found within prairies (Gustafson et al., 1999). For evaluating levels of diversity required for effective conservation strategies for *A. gerardii* and tallgrass prairie, it will be important to also consider variation for genes under selection.

Resistance gene evolution has previously been evaluated based on a small number of individuals from each population, leaving within-population dynamics largely unknown, though modeling studies have attempted to assess population-level dynamics (Bergelson et al., 2001). Though analysis of resistance gene evolution at large spatial scales elucidates overarching patterns of diversity, the mechanisms maintaining this diversity are only inferred. We present the first study of diversity among individuals within populations, providing insight into the mechanism of resistance gene evolution at the population level.

Traditionally, resistance genes were thought to evolve in a gene-for-gene context where resistance genes directly interact with avirulence genes (Flor, 1971). The coevolutionary

dynamics in such a system can be described by the Red Queen Hypothesis (Lythgoe and Read, 1998), where both pathogen and host are rapidly adapting to one another, resulting in selective sweeps of successful resistance and avirulence genes. Resistance genes evolve in tandem arrays providing novel specificities through recombination (see Hulbert et al., 2001 for review), though in some systems orthologs are more similar than paralogs, suggesting divergent selection independent of recombination (Michelmore and Meyers, 1998). Recent studies of resistance gene evolution show that rapid evolution does not always occur and ancient polymorphisms can be maintained (Bergelson et al., 2001). Ancient polymorphisms are present in *Rpm1* and *Rps5* in *Arabidopsis thaliana* where resistant and susceptible alleles are maintained across broad geographic regions (Stahl et al., 1999; Tian et al., 2002). This polymorphism is due in part to a fitness cost of resistance in plants with *Rpm1* (Tian et al., 2003).

Dodds et al. (2006) identified a direct molecular interaction between resistance genes in the *L* resistance locus in flax (*Linum usitatissimum*) and avirulence genes in flax rust (*Melampsora lini*). Diversifying selection was observed to operate in avirulence loci in flax rust consistent with the selective sweeps model. Rapid diversifying selection in the leucine rich-repeat (LRR) region of resistance genes is thought to be a plant response to the changing nature of pathogen avirulence genes (Michelmore and Meyers, 1998). Caicedo and Schaal (2004) demonstrated selection against amino acid substitution in the 5' ends of Cf-2 homologs and positive selection in the 3' region in wild populations of *Solanum pimpinellifolium*. Diversity in Cf-2 homologs was attributed to single nucleotide polymorphisms and indels in the region coding for the LRR.

In contrast, long-term maintenance of variation has been observed, suggesting balancing selection for *RPS2* resistance in *A. thaliana,* dependent on the presence of *AvrRpt2* in *Pseudomonas syringae* (Bent et al., 1994). Most of the sequence variation in *RPS2* occurs in the LRR (Mauricio et al., 2003). Much of the variation in reaction to *Pseudomonas syringae* pv. *tomato* was attributed to variation in the sequence of *Pto* in wild tomato relatives (Rose et al., 2005). Rose et al. (2007) found both purifying selection and maintenance of polymorphism to be operating in the *Pto* gene in natural populations of tomato relatives, where many populations were represented by a single individual. The conserved function of recognizing avirulence proteins seems to drive purifying selection, whereas the function of recognizing different

avirulence proteins and the pathogen fitness cost of virulence may operate to maintain polymorphisms.

The general trend of increasing resistance gene locus diversity with increasing longitude and precipitation in our study suggest that, within populations, *Rxo1* homolog dissimilarity is positively correlated with disease pressure. Information on which pathogen, if any, interacts with this resistance gene locus in *A. gerardii* would be needed in order to directly assess disease pressure. However, *Rxo1* recognizes avirulent *Burkholderia andropogonis* in maize, and *B. andropogonis* is found to infect *A. gerardii*. It is possible that the *A. gerardii Rxo1* homolog recognizes avirulent *B. andropogonis*. The fact that *Rxo1* recognizes not only avirulent *B. andropogonis*, but also *avrRxo1* in *Xanthomonas oryzae* pv. *oryzicola* (Zhao et al., 2005), suggests that *Rxo1* recognizes a very conserved bacterial effector, a specific site of a bacterial effector, or that *Rxo1* acts as a guard of a target of a bacterial effector.

The data presented do not provide evidence that the *A. gerardii Rxo1* homolog provides a conserved function. Since *Rxo1* homolog diversity was higher in environments more conducive to disease, it is likely that the *Rxo1* homolog was under diversifying selection for the recognition of a rapidly evolving pathogen. To more directly assess evolution at the *Rxo1* homolog locus, it would be desirable to identify the components of intra-individual variation in amplified products of the *Rxo1* homolog (alleles, tandem genes, or genes in different regions or chromosomes) and to sequence these components. Cloning amplified products is necessary in *A. gerardii* because individuals harbor various amplified regions, and multiple individuals from each population would need to be sequenced. Functional analysis of the *Rxo1* homolog locus is needed to more fully understand the molecular evolution of this locus.

Similarity in the *Rxo1* homolog was disproportionately high relative to AFLP similarity in the western prairies (Smoky Valley Ranch and Wilson Lake). We doubt that this is due to disease-mediated selection because disease is not expected to be a strong selective force in these drier and more isolated populations (Huber and Gillespie, 1992; Alexander et al., 2007). Purifying selection at this locus could be due to selection for a gene linked to *Rxo1* that is important under western environmental conditions. Purifying selection has been found at low variability QTL caused by selective sweeps in drought and saline adapted *Helianthus annuus*

populations in Utah (Kane and Rieseberg, 2007). The low level of *Rxo1* diversity in the west may also be explained by selection against resistant alleles in the absence of disease, which may take place if *Rxo1*-mediated resistance has a cost similar to *Rpm1* in *Arabidopsis* (Tian et al., 2003).

Mixtures of host genotypes can be used to manage disease (Garrett and Mundt, 1999). Though we found increased *Rxo1* homolog diversity in populations with increased disease pressure, we do not know if *Rxo1* homolog diversity provides a population-level advantage to plant health. Disease monitoring of populations with various levels of *Rxo1* homolog diversity in a common garden experiment would be one way to test this hypothesis.

Diversity of important genes in native populations may prove to be valuable as selection pressures shift due to changing climate (Garrett et al., 2006). Unfortunately, the diversity within native populations may not be adequate to allow adaptive change to keep up with the rate of climate change in the twenty-first century, as has been shown for *Chamaecrista fasciculata* populations in the Great Plains (Etterson and Shaw, 2001). This may prove to be particularly important for resistance gene diversity, as disease may impose a strong selective pressure and is highly dependent on climate (Garrett et al., 2006). Populations that have not experienced disease pressure in recent evolutionary history may prove to be poorly adapted in the near future, and some species may suffer from inbreeding depression as novel selective forces shape population genetics.

We have demonstrated that resistance gene diversity varies with precipitation, one of the most important environmental drivers of plant disease. Further studies integrating population-level dynamics with evolutionary history will elucidate the complexities and patterns of resistance gene evolution. Such studies are also needed to inform plant pathology and plant breeding for more effective deployment of resistance genes in agriculture to achieve durable resistance.

Figures and Tables

Figure 1.1 Sampling methodology demonstrated at different scales. Each prairie (A) contained four transects. Each 40 m transect (B) contained a series of five subgroups. Five ramets were sampled at each subgroup (C).

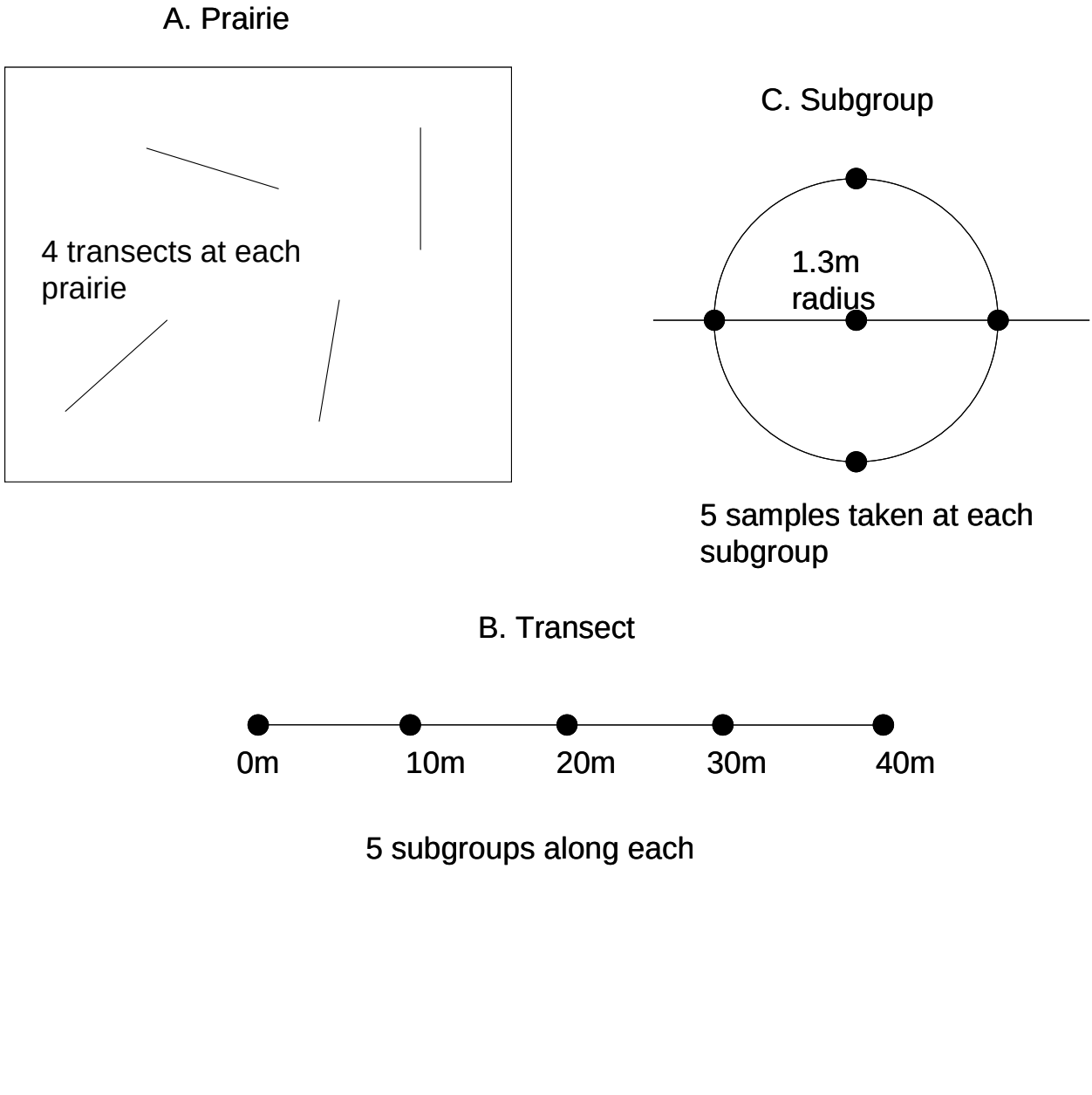

Figure 1.2 Bootstrap consensus tree of 65 *Andropogon gerardii* clones of known ploidy, one *Schizachyrium scoparium* clone, and two *Andropogon gerardii hallii* clones. The *S. scoparium* clone was used as an outgroup.The second set of digits in the sample name indicates ploidy level. Aneuploids are indicated by 75, though their exact number of chromosomes is not known. Bootstrap support based on 1000 iterations is indicated for branches with greater than 50% support.

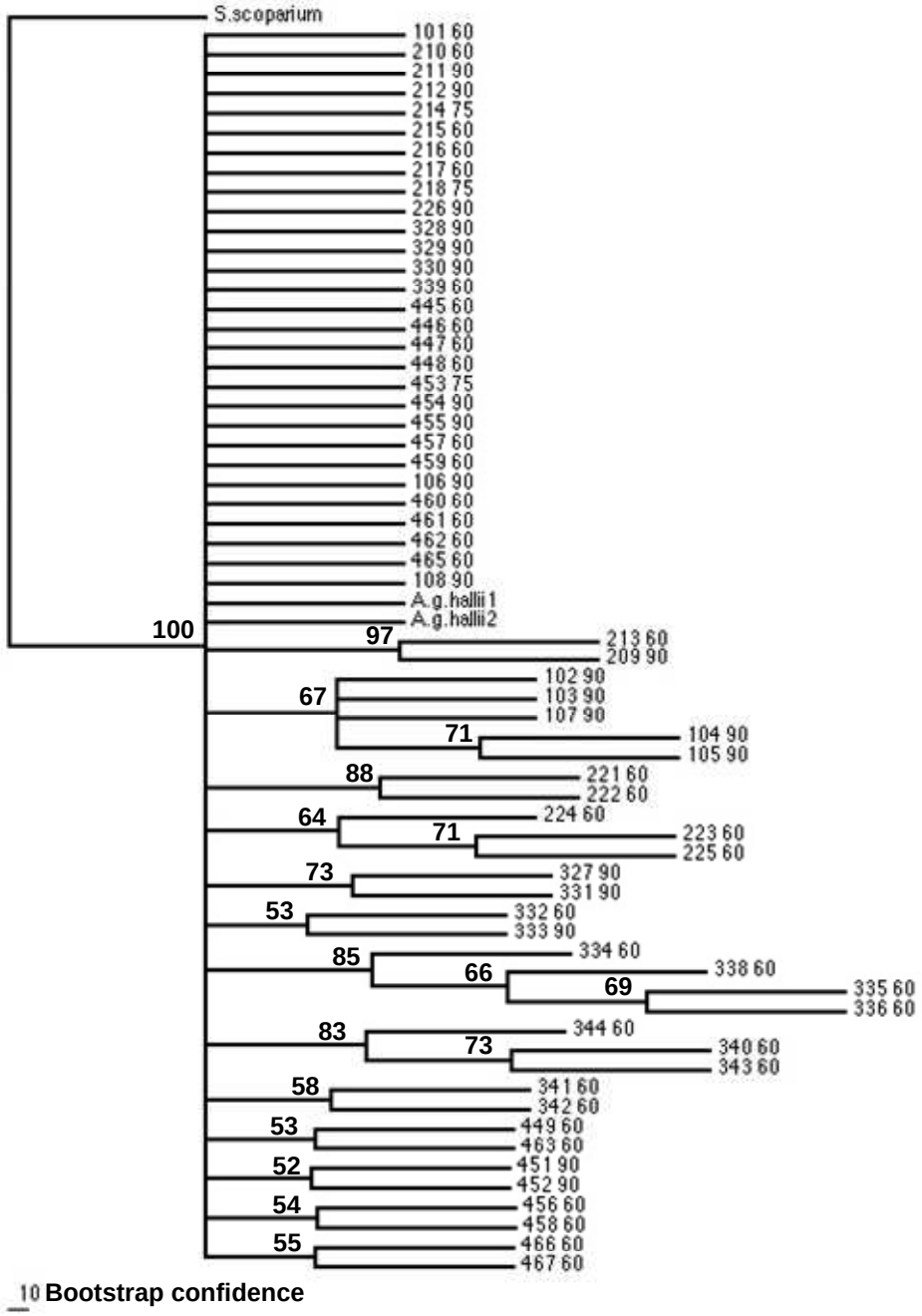

Figure 1.3 AFLP dissimilarity by geographic distance. The data points from left to right are plotted at 1.8 m (within subgroup), 10 m (between adjacent subgroups), 20 m, 30 m, 40 m, 2139 m (between transects), and 473,025 m (between prairies). The solid line indicates observed AFLP dissimilarity and the dotted line indicates a 95% confidence interval based on a randomization test under the null hypothesis of no relationship between dissimilarity and distance.

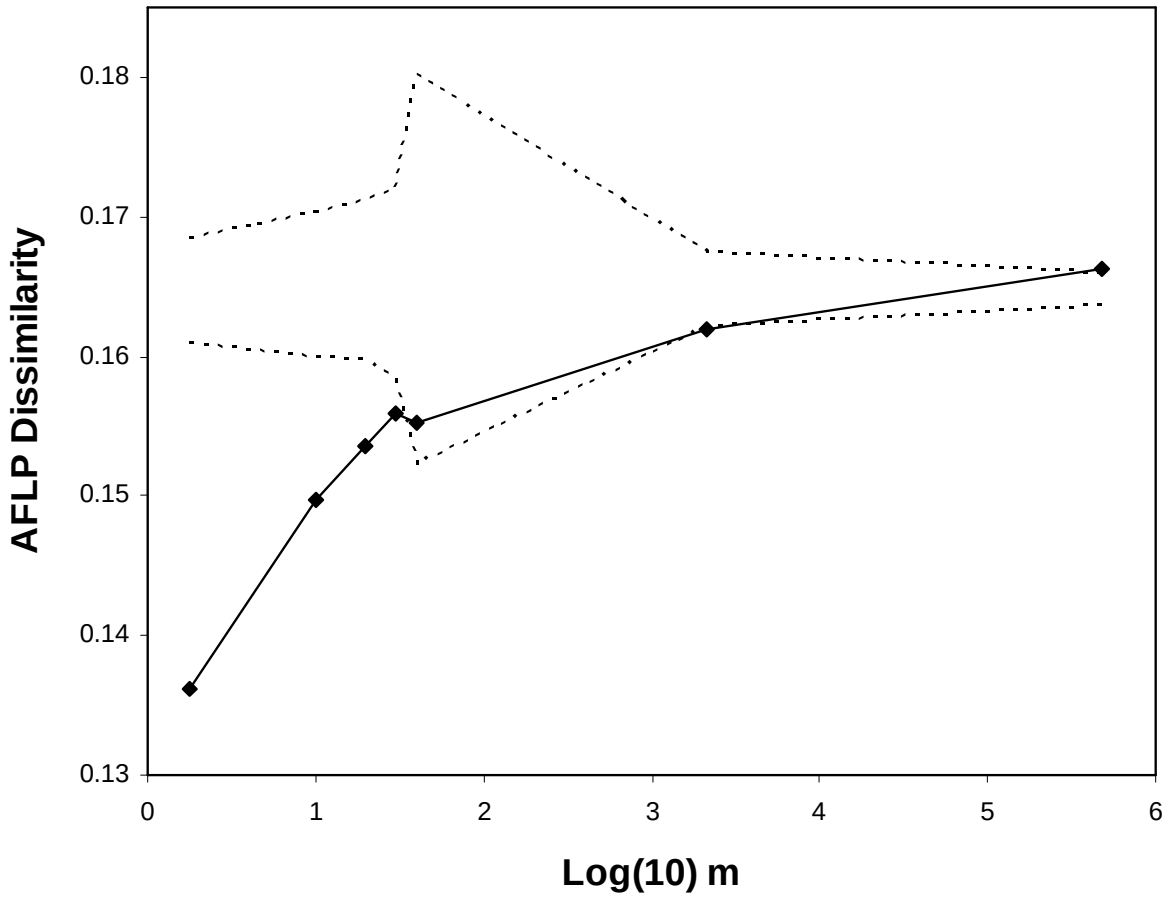

Figure 1.4 AFLP dissimilarity across longitude for different scales. Names of prairies west to east (left to right) are: Smoky Valley Ranch, Wilson Lake, The Land Institute, KPBS, and Tucker prairie. Diamonds represent within-subgroup dissimilarities, squares represent within-transect, among-subgroup dissimilarities, triangles represent among-transect dissimilarities, and "-"s represent within-prairie dissimilarities.

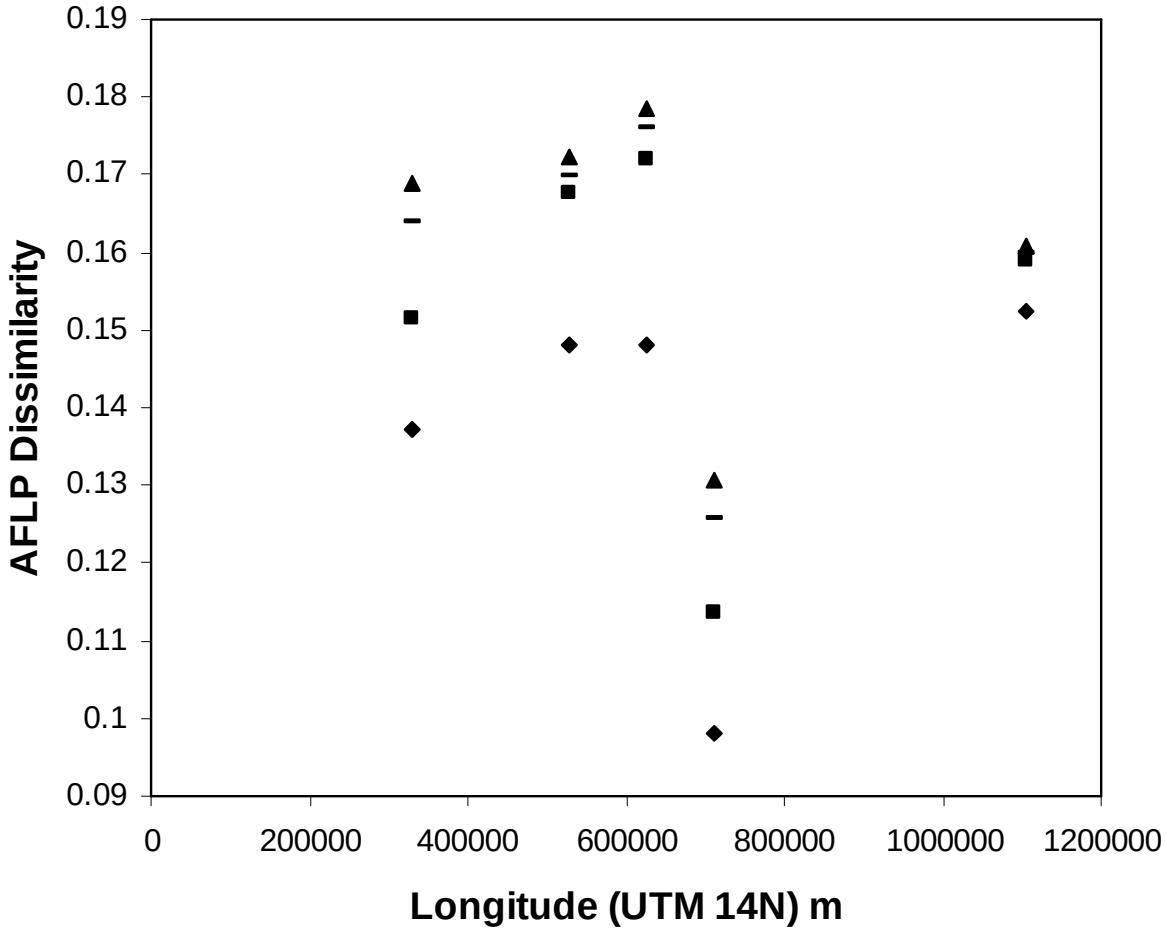

Figure 1.5 ***Rxo1*** **dissimilarity by geographic distance. The data points from left to right are plotted at 1.8 m (within subgroup), 10 m (between adjacent subgroups), 20 m, 30 m, 40 m, 2139 m (between transects), and 473,025 m (between prairies). The solid line indicates observed** ***Rxo1*** **dissimilarity and the dotted line indicates a 95% confidence interval based on a randomization test under the null hypothesis of no relationship between dissimilarity and distance.**

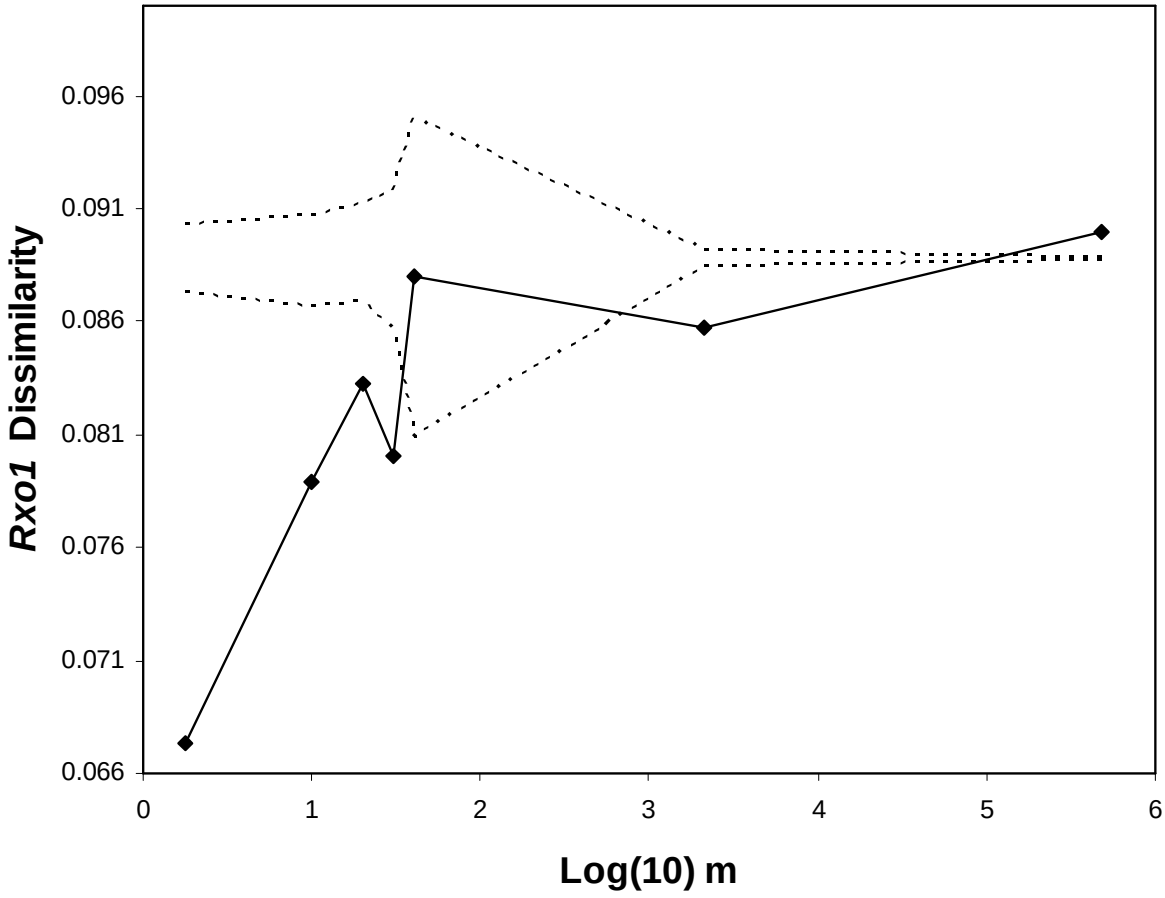

Figure 1.6 ***Rxo1*** **dissimilarity at different scales across longitude. Names of prairies west to east (left to right) are: Smoky Valley Ranch, Wilson Lake, The Land Institute, KPBS, and Tucker prairie. Diamonds represent within-subgroup dissimilarities, squares represent within-transect, among-subgroup dissimilarities, triangles represent among-transect dissimilarities, and "-"s represent within-prairie dissimilarities.**

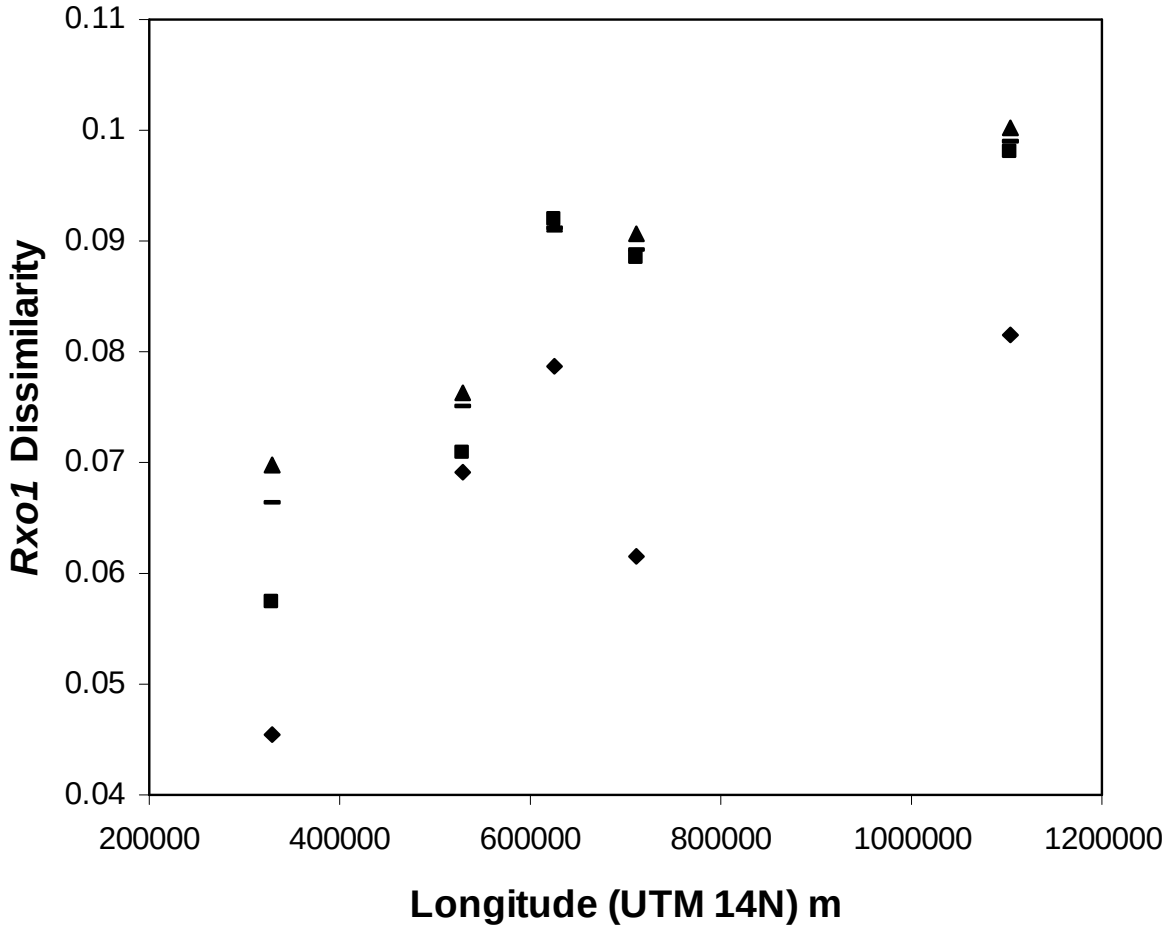

Figure 1.7 AFLP dissimilarity by distance for Tucker prairie. The solid line indicates observed AFLP dissimilarity and the dotted line indicates a 95% confidence interval based on a randomization test under the null hypothesis of no relationship between dissimilarity and distance.

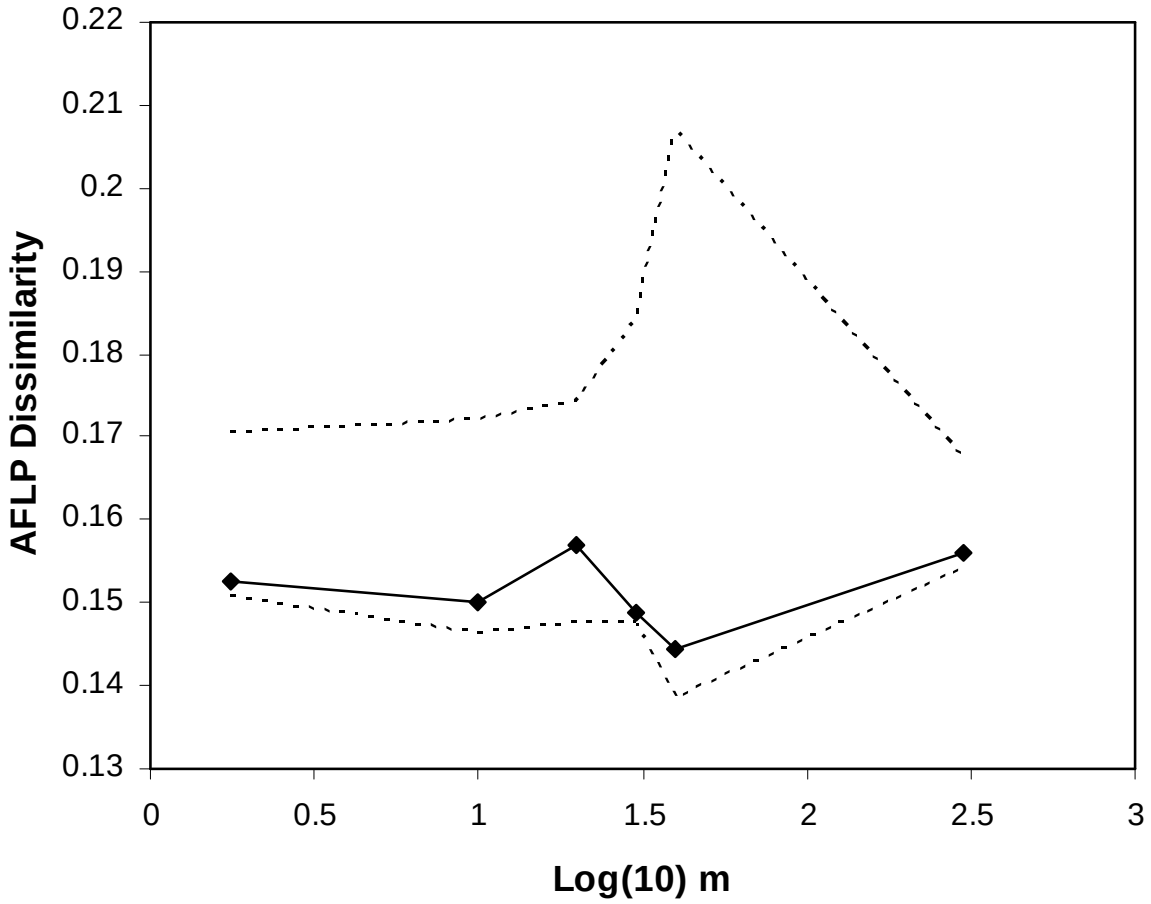

Figure 1.8 ***Rxo1*** **dissimilarity by distance for Tucker prairie. The solid line indicates observed** ***Rxo1*** **dissimilarity and the dotted line indicates a 95% confidence interval based on a randomization test under the null hypothesis of no relationship between dissimilarity and distance.**

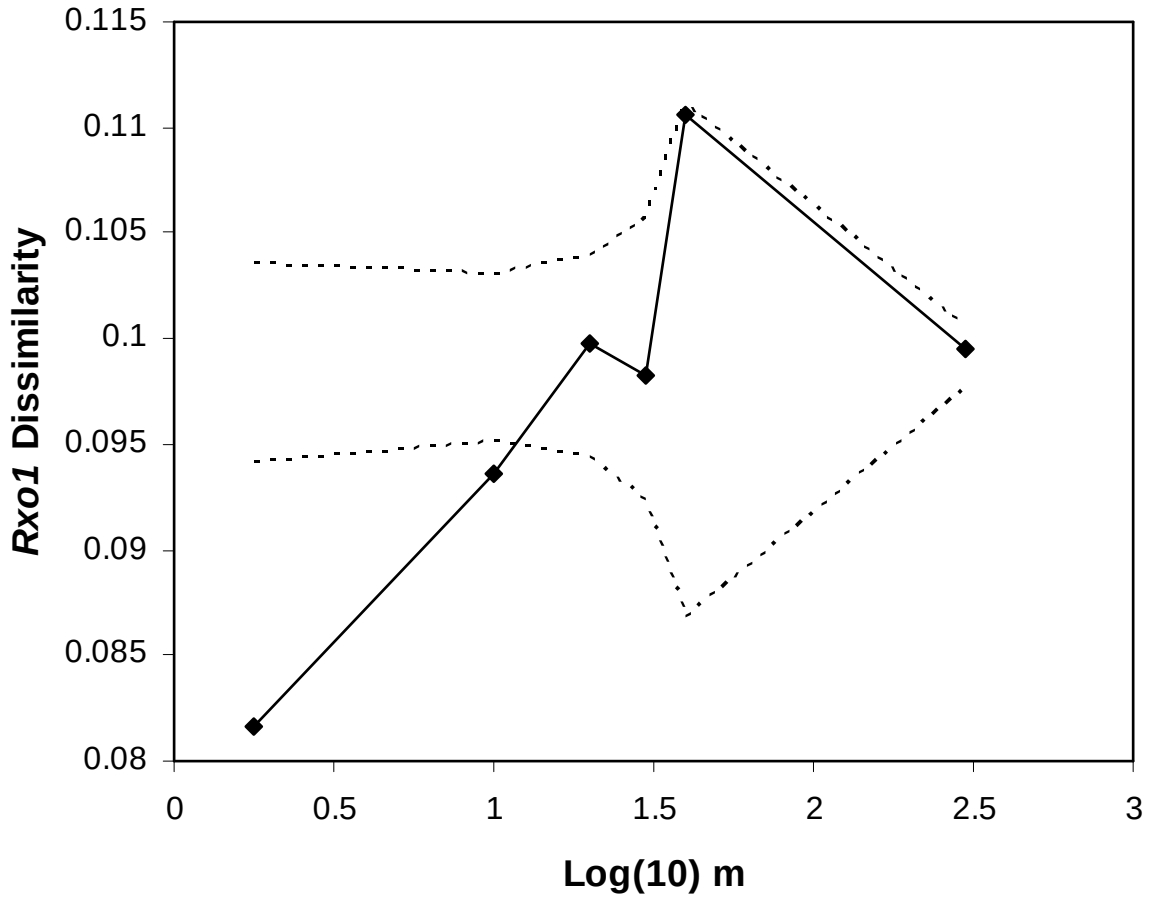

Figure 1.9 AFLP dissimilarity by distance for KPBS. The solid line indicates observed AFLP dissimilarity and the dotted line indicates a 95% confidence interval based on a randomization test under the null hypothesis of no relationship between dissimilarity and distance.

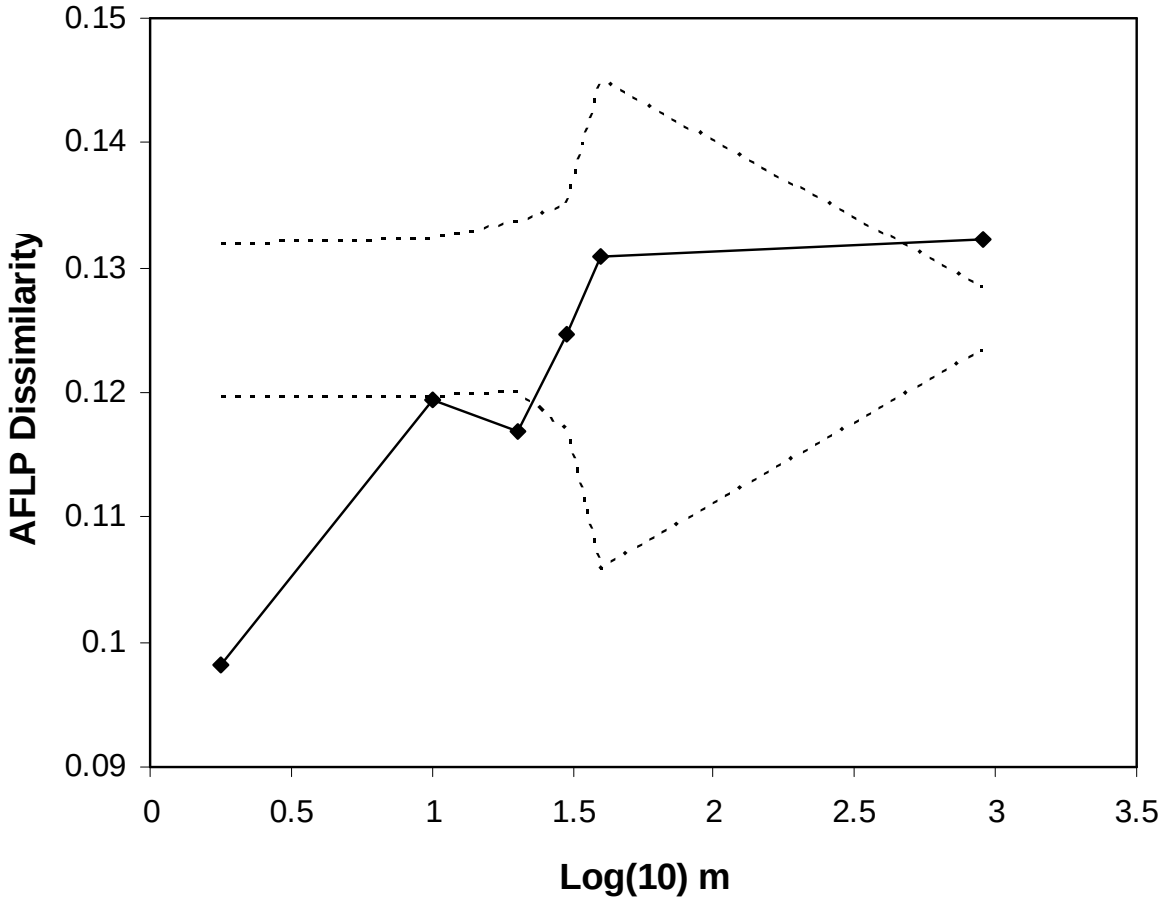

Figure 1.10 ***Rxo1*** **dissimilarity by distance for KPBS. The solid line indicates observed** ***Rxo1*** **dissimilarity and the dotted line indicates a 95% confidence interval based on a randomization test under the null hypothesis of no relationship between dissimilarity and distance.**

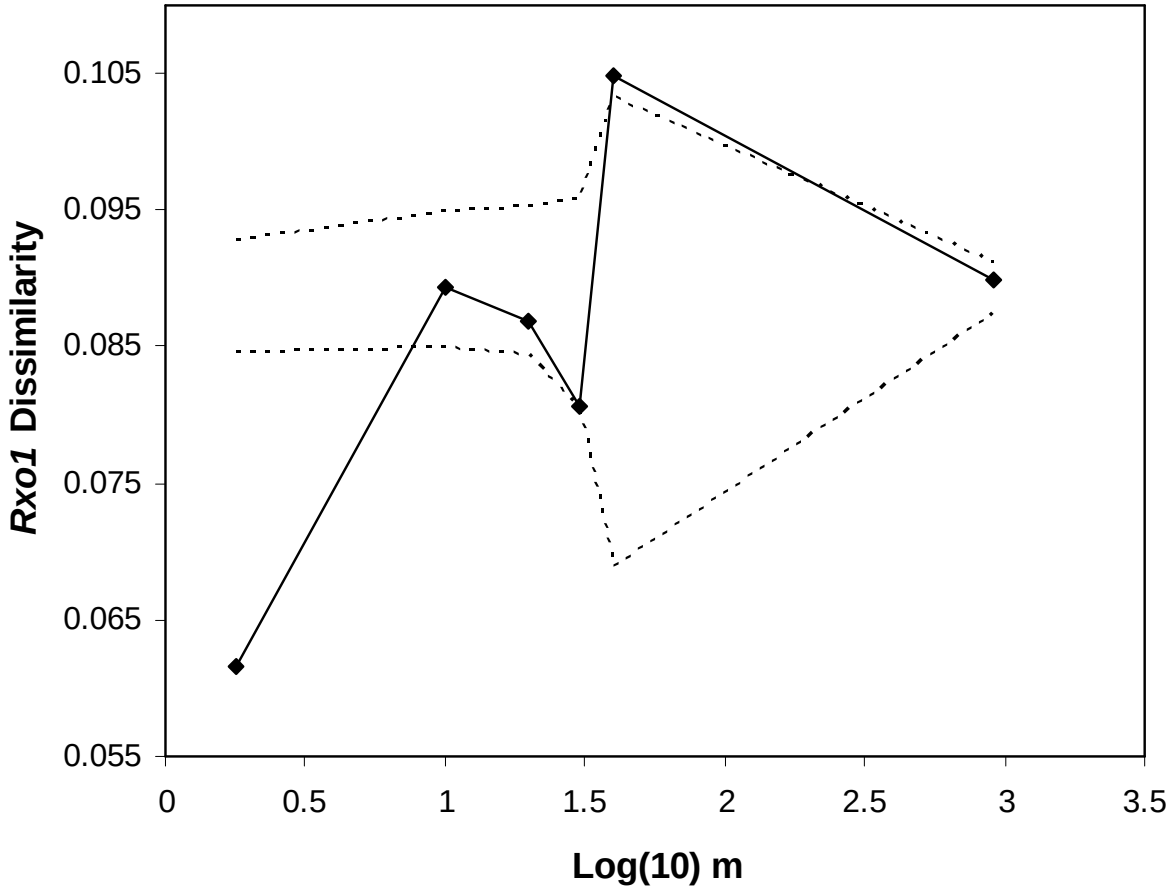

Figure 1.11 AFLP dissimilarity by distance for TLI. The solid line indicates observed AFLP dissimilarity and the dotted line indicates a 95% confidence interval based on a randomization test under the null hypothesis of no relationship between dissimilarity and distance.

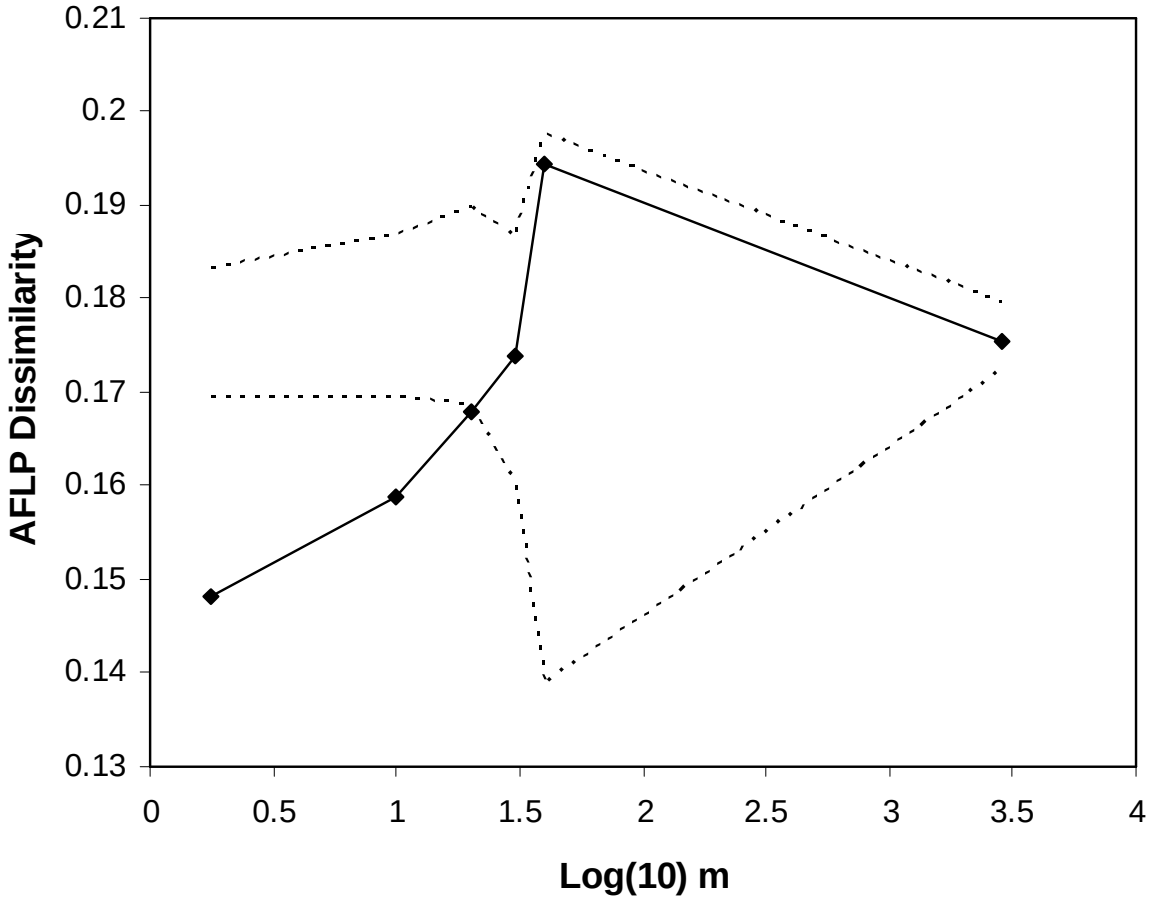

Figure 1.12 ***Rxo1*** **dissimilarity by distance for TLI. The solid line indicates observed** ***Rxo1*** **dissimilarity and the dotted line indicates a 95% confidence interval based on a randomization test under the null hypothesis of no relationship between dissimilarity and distance.**

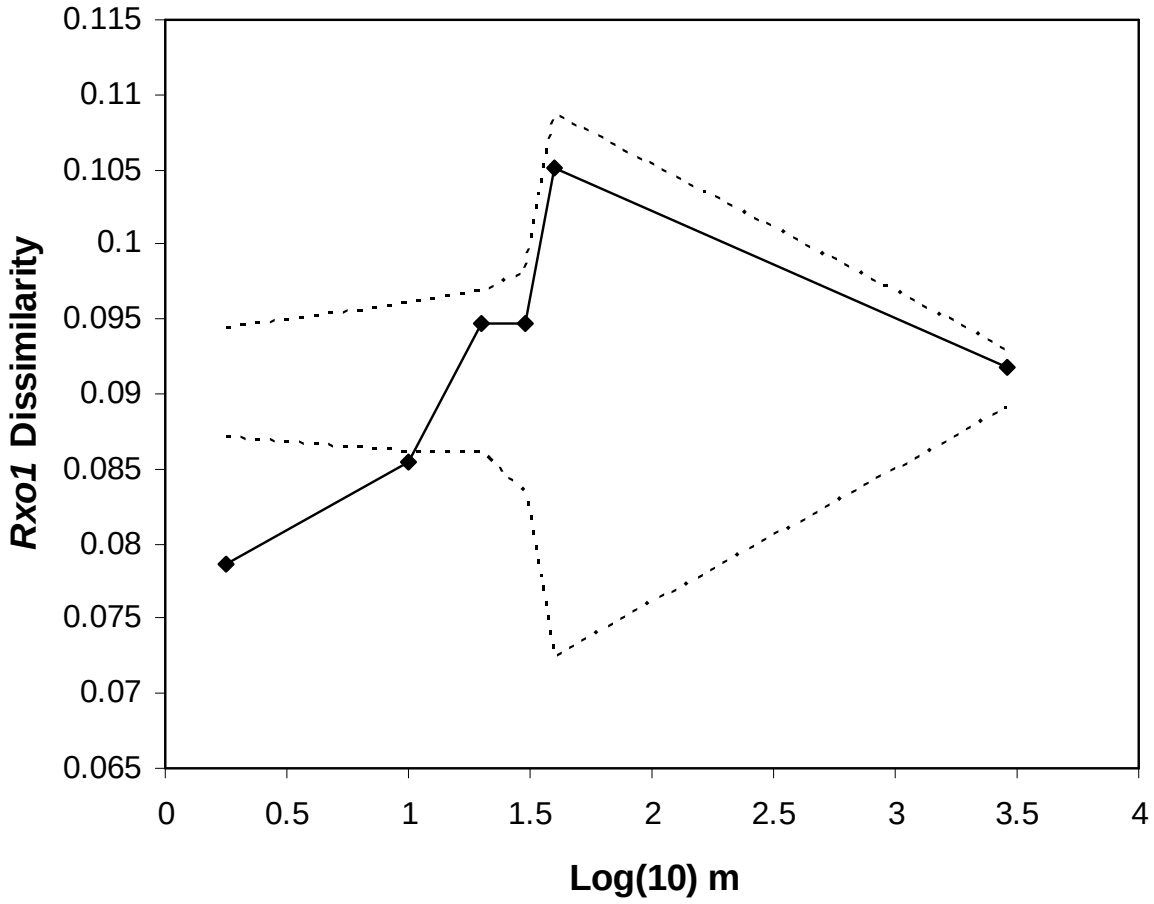

Figure 1.13 AFLP dissimilarity by distance for Wilson Lake. The solid line indicates observed AFLP dissimilarity and the dotted line indicates a 95% confidence interval based on a randomization test under the null hypothesis of no relationship between dissimilarity and distance.

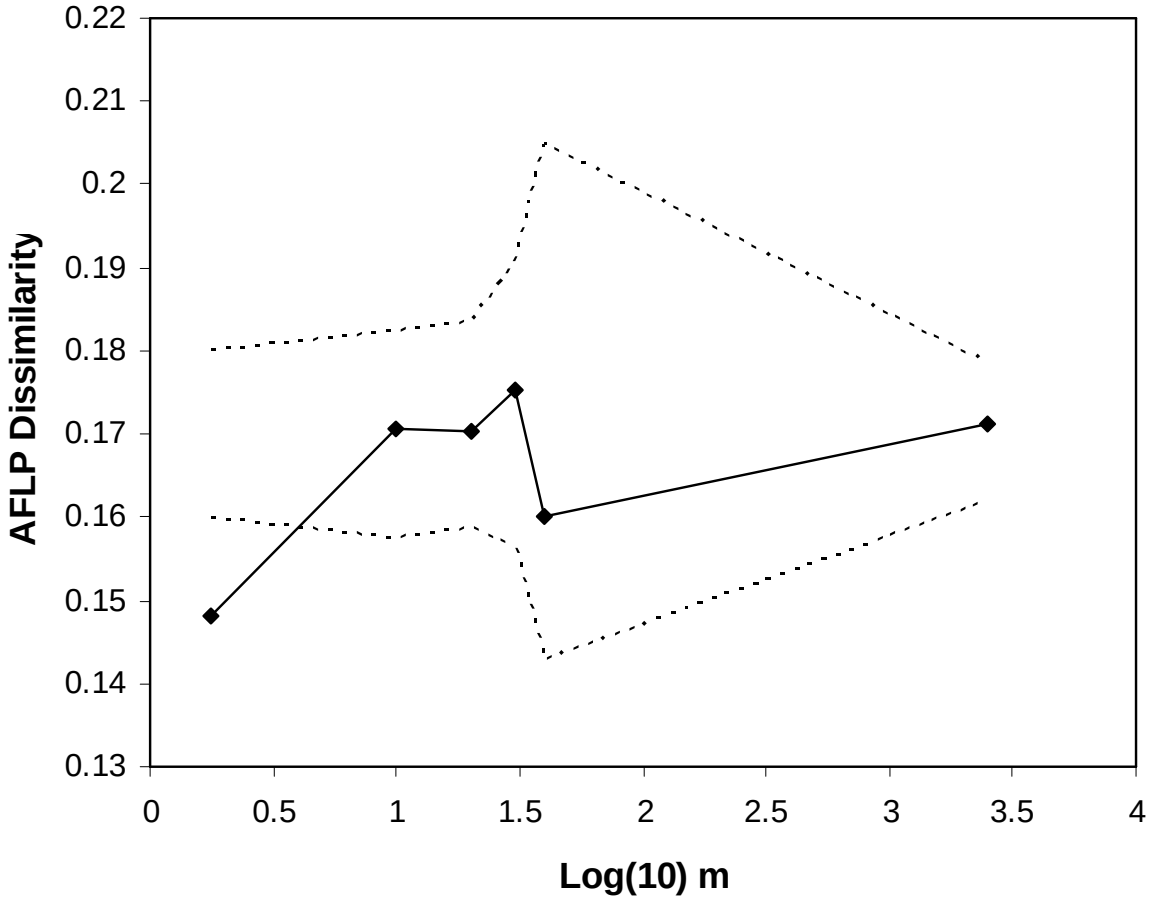

Figure 1.14 ***Rxo1*** **dissimilarity by distance for Wilson Lake. The solid line indicates observed** ***Rxo1*** **dissimilarity and the dotted line indicates a 95% confidence interval based on a randomization test under the null hypothesis of no relationship between dissimilarity and distance.**

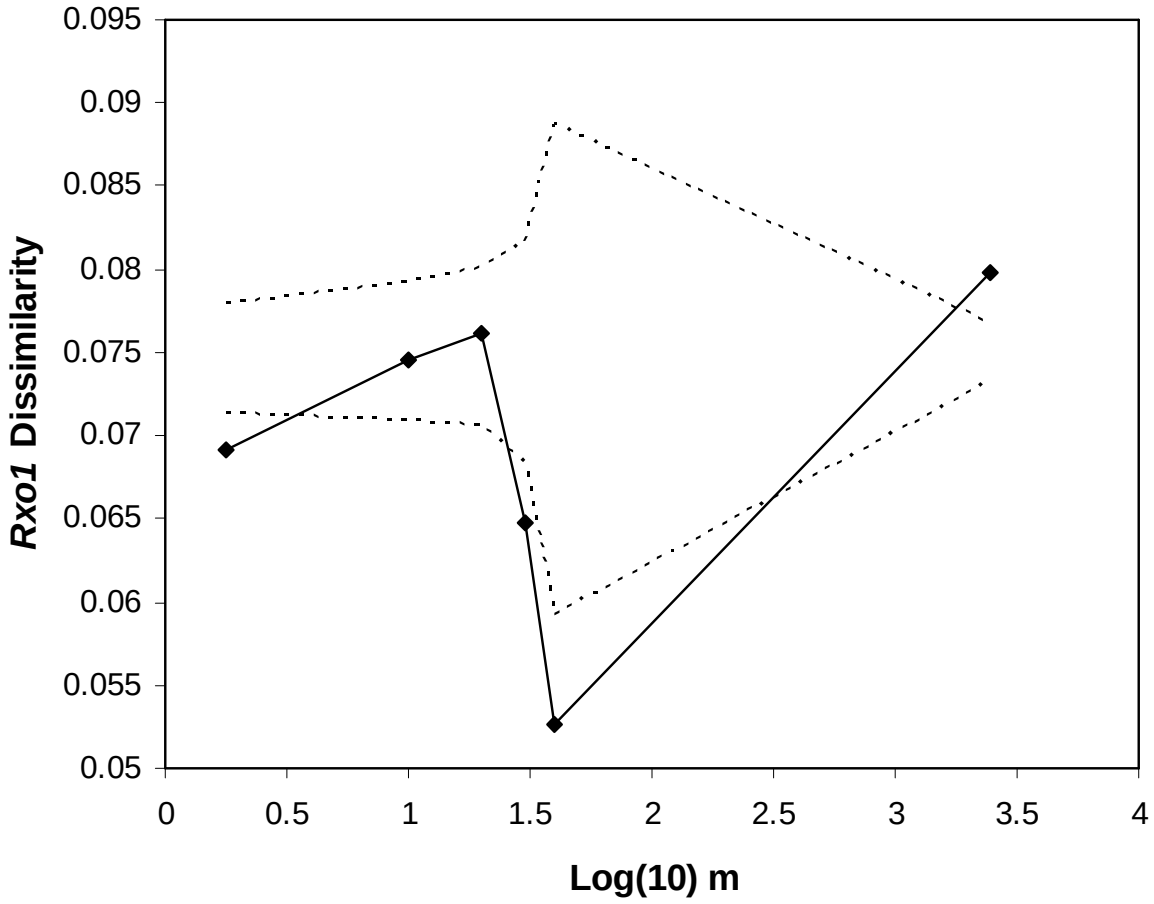

Figure 1.15 AFLP dissimilarity by distance for Smoky Valley Ranch. The solid line indicates observed AFLP dissimilarity and the dotted line indicates a 95% confidence interval based on a randomization test under the null hypothesis of no relationship between dissimilarity and distance.

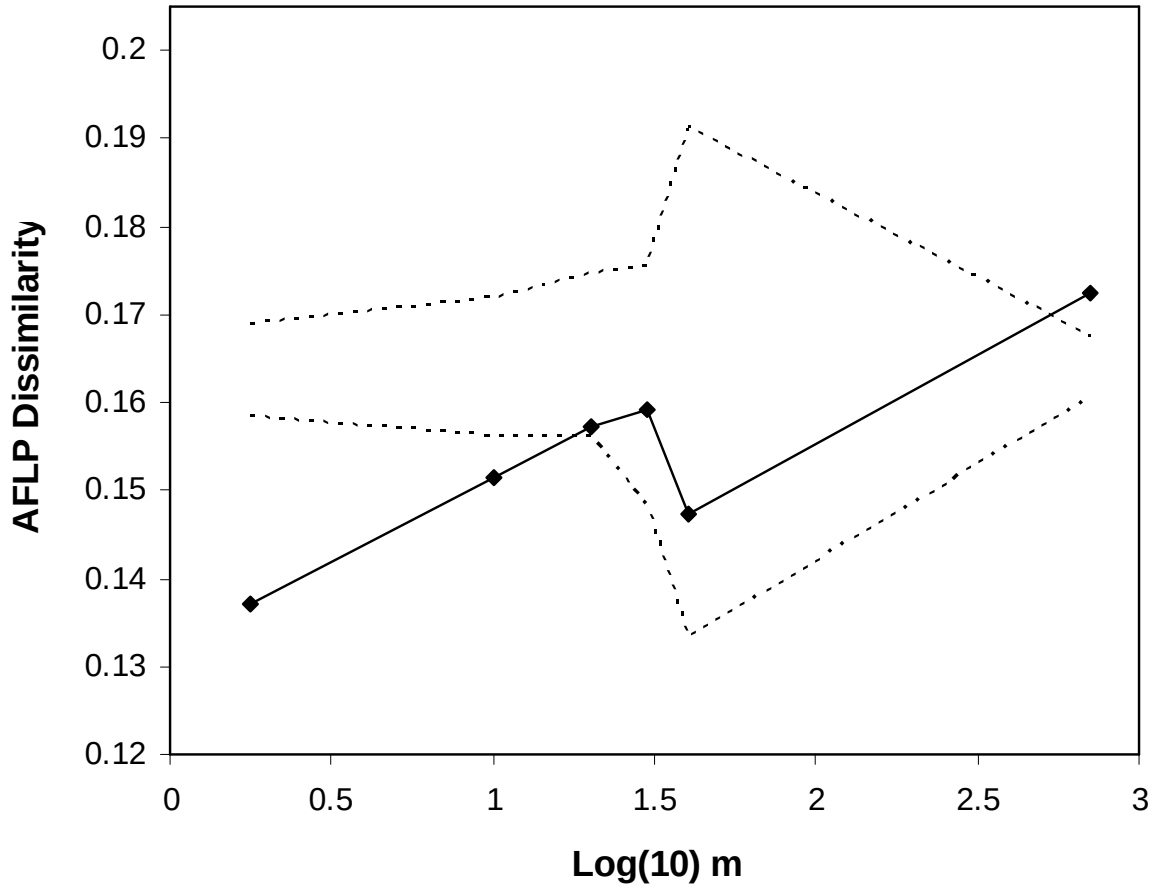

Figure 1.16 *Rxo1* dissimilarity by distance for Smoky Valley Ranch. The solid line indicates observed *Rxo1* dissimilarity and the dotted line indicates a 95% confidence interval based on a randomization test under the null hypothesis of no relationship between dissimilarity and distance.

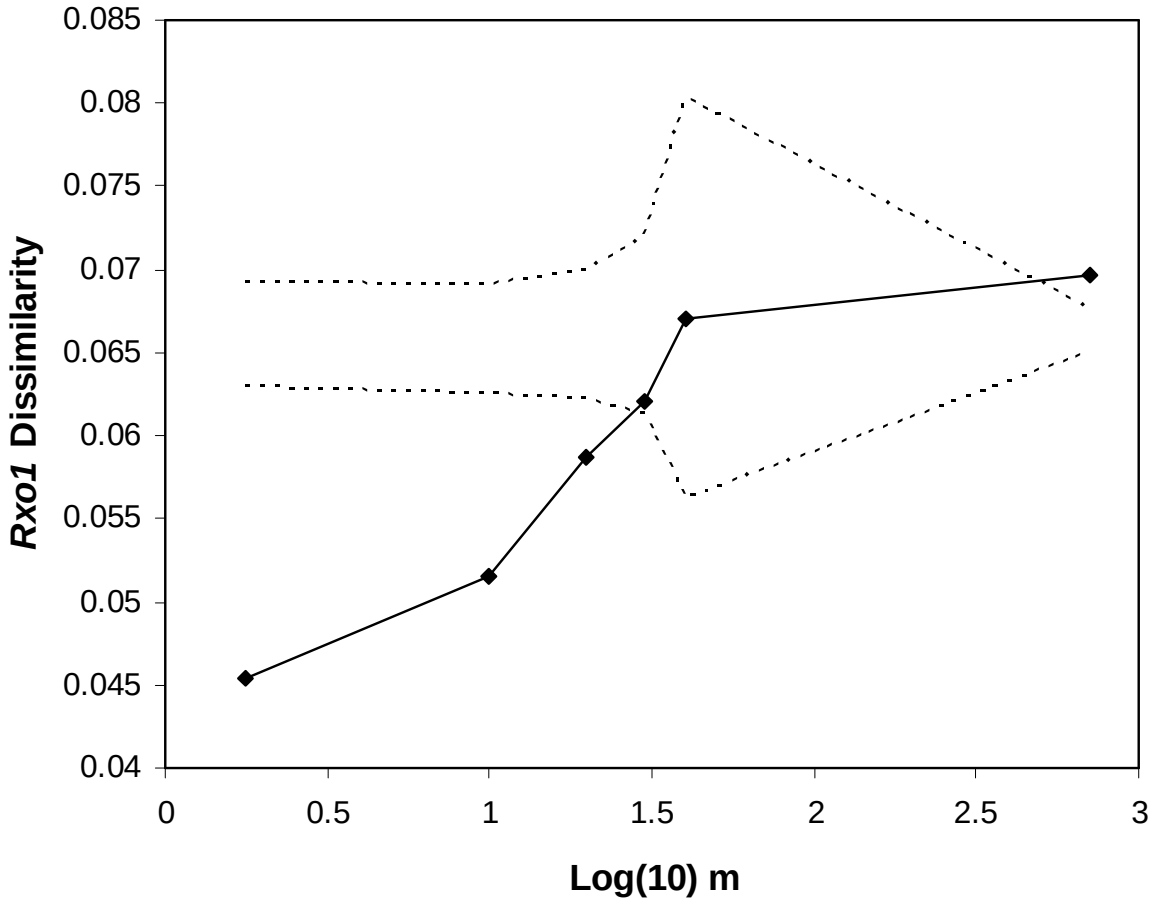

Figure 1.17 ***Rxo1*** **dissimilarity across longitude. The mean dissimilarity for two random, independent pairs of transects was calculated for each prairie. A linear regression model was applied.**

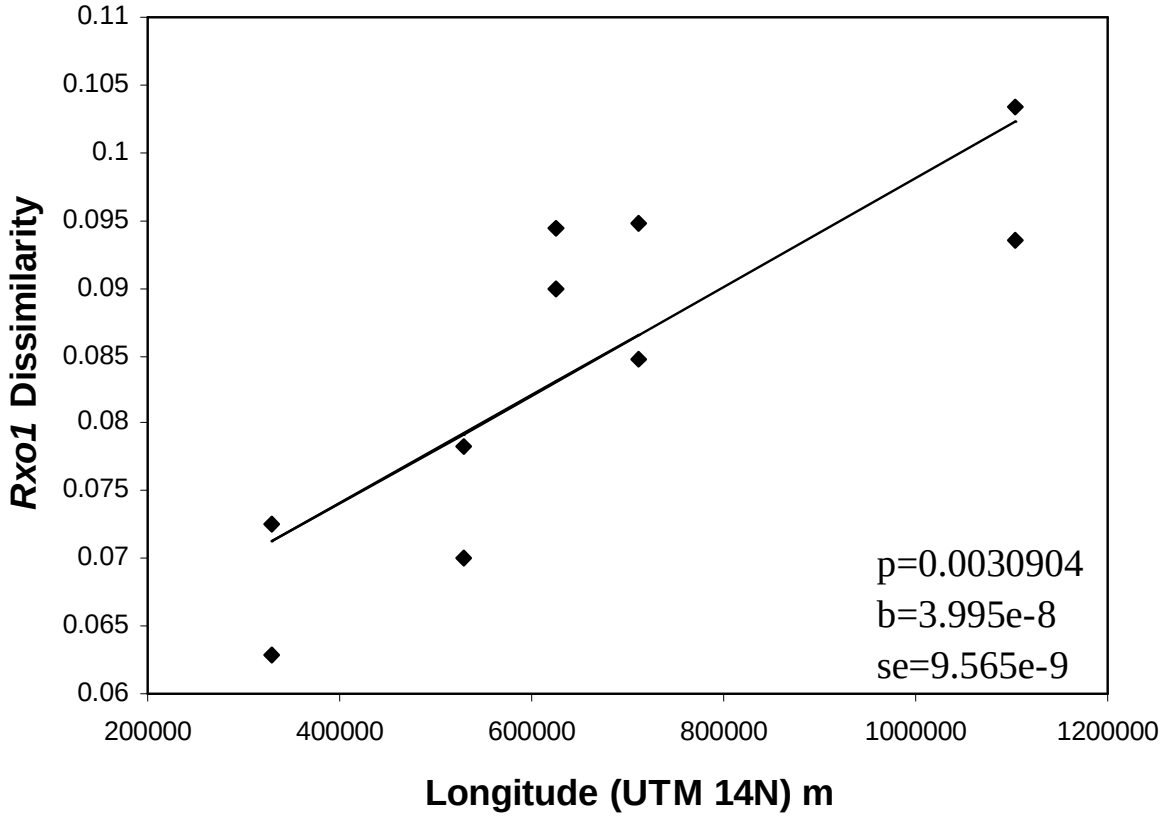

Figure 1.18 AFLP dissimilarity across longitude. The mean dissimilarity for two random, independent pairs of transects was calculated for each prairie. A linear regression model was applied.

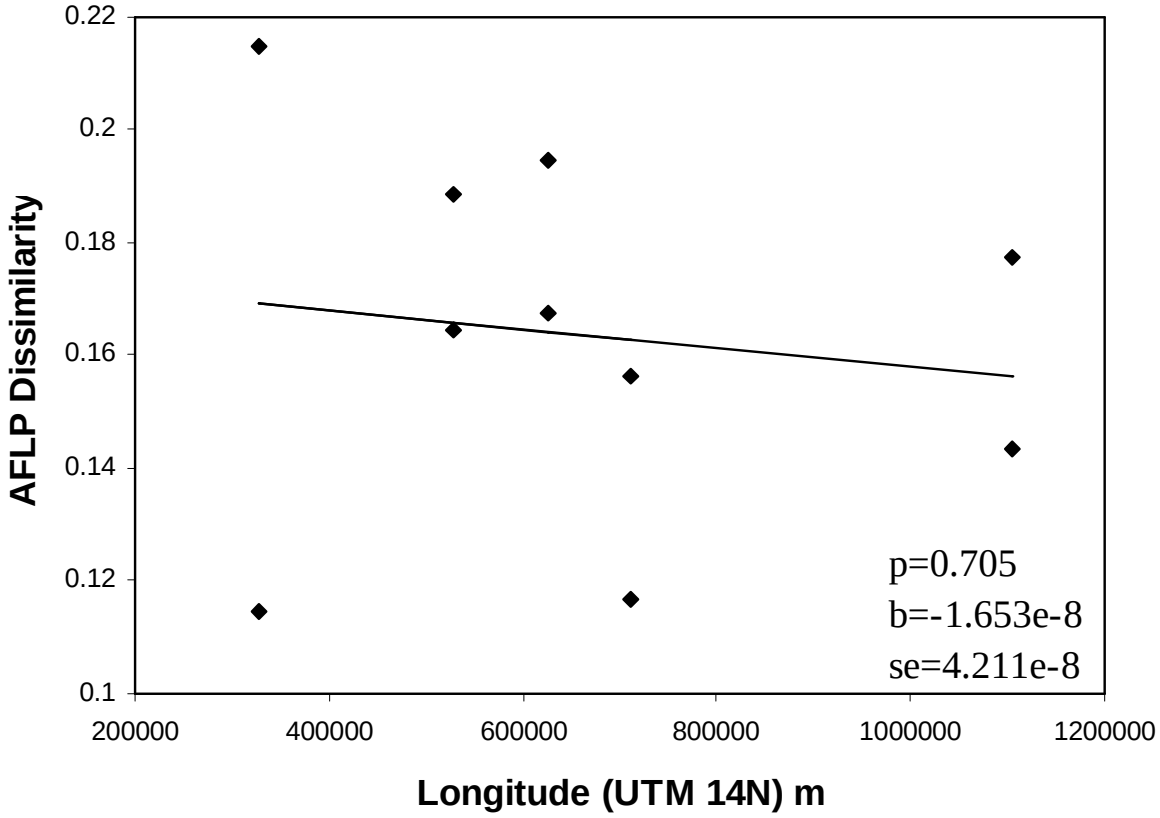

Figure 1.19 ***Rxo1*** **dissimilarity across average annual precipitation. The mean dissimilarity for two random, independent pairs of transects was calculated for each prairie. A linear regression model was applied.**

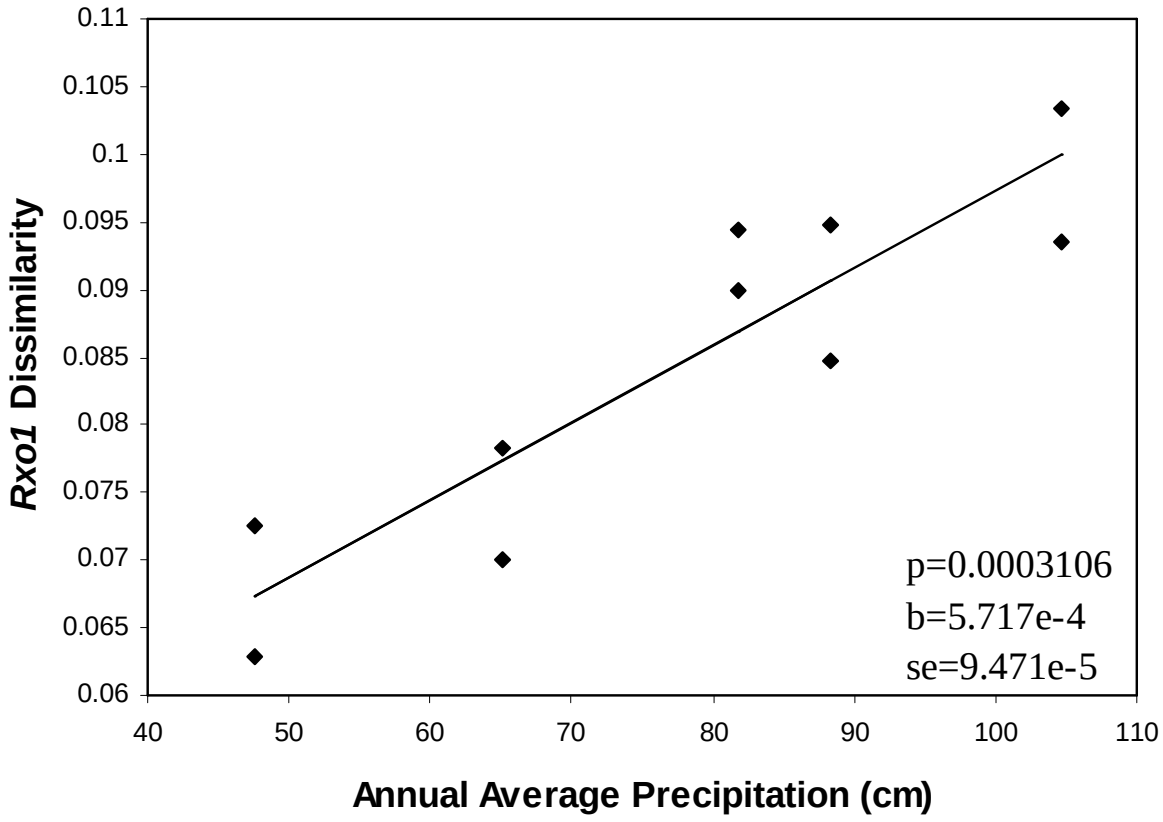

Figure 1.20 AFLP dissimilarity across average annual precipitation. The mean dissimilarity for two random, independent pairs of transects was calculated for each prairie. A linear regression model was applied.

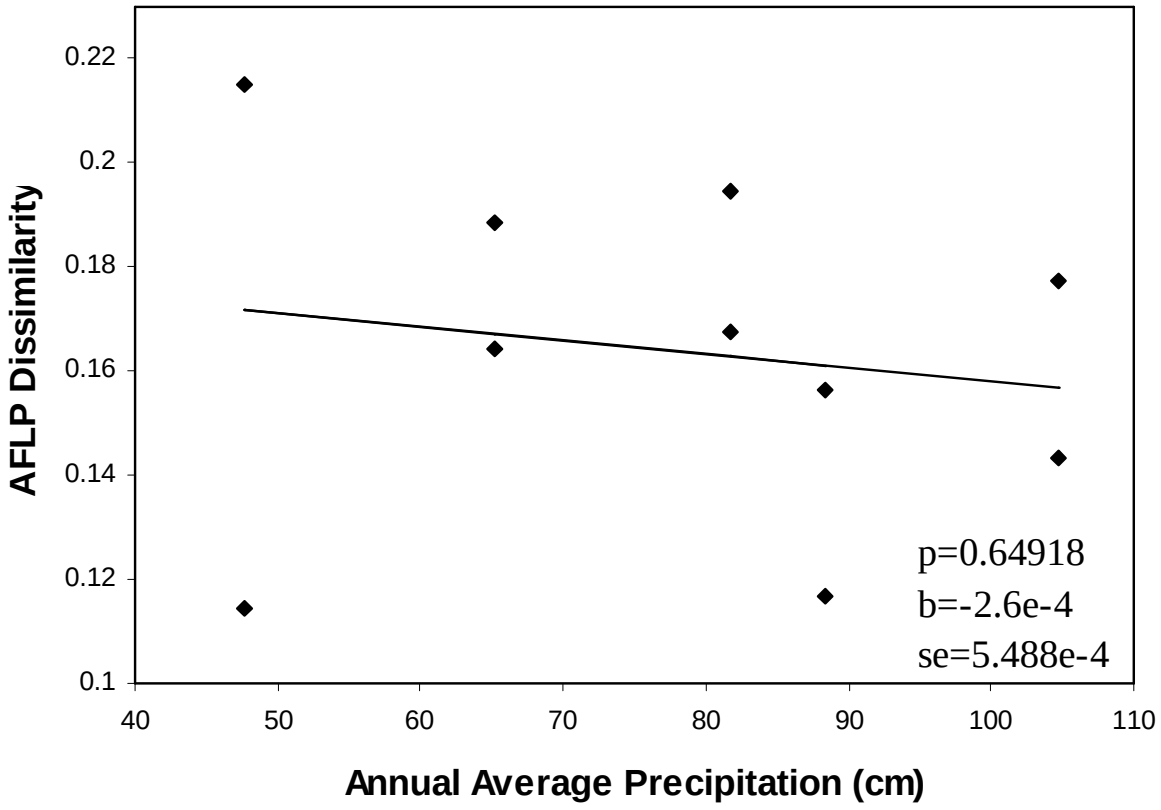

Table 1.1 Tallgrass prairies sampled, with state, UTM coordinates, AFLP sample size, *Rxo1*-RFLP sample size and average annual precipitation from the nearest weather station (1971-2000) (precipitation data from National Oceanic and Atmoshperic Administration, 2003). KPBS is Konza Prairie Biological Station and TLI is The Land Institute. Tissues from Boulder were used in ploidy analyses.

Prairie	State	UTM Coordinates	AFLP sample size	*Rxo1*-RFLP sample size	Average Annual Precipitation (cm/year)
Tucker Prairie	Missouri	15 N 587416 E 4311575 N	71	100	104.7
KPBS	Kansas	14 N 709536 E 4327791 N	84	100	88.39
TLI	Kansas	14 N 624616 E 4292829 N	85	100	81.76
Wilson Lake	Kansas	14 N 527770 E 4310686 N	70	100	65.18
Smoky Valley Ranch	Kansas	14 N 328103 E 4306146 N	97	100	47.63
Boulder	Colorado	13 N 475769 E 4427696 N	63	63	50.62

References

Alberto, F., L. Gouveia, S. Arnaud-Haond, J. L. Perez-Llorens, C. M. Duarte, and E. A. Serrao. 2005. Within-population spatial genetic structure, neighbourhood size and clonal subrange in the seagrass *Cymodocea nodosa*. Molecular Ecology 14:2669-2681.

Alexander, H. M., S. Price, R. Houser, D. Finch, and M. Tourtellot. 2007. Is there reduction in disease and pre-dispersal seed predation at the border of a host plant's range? Field and herbarium studies of *Carex blanda*. Journal of Ecology 95:446-457.

Balloux, F., L. Lehmann, and T. de Meeus. 2003. The population genetics of clonal and partially clonal diploids. Genetics 164:1635-1644.

Baums, I. B., M. W. Miller, and M. E. Hellberg. 2006. Geographic variation in clonal structure in a reed-building Caribbean coral, *Acropora palmata*. Ecological Monographs 76:503-519.

Bent, A. F., B. N. Kunkel, D. Dahlbeck, K. L. Brown, R. Schmidt, J. Giraudat, J. Leung, and B. J. Staskawicz. 1994. *RPS2* of *Arabidopsis thaliana*: A leucine-rich repeat class of plant disease resistance genes. Science 256:1856-1860.

Bergelson, J., G. Dwyer, and J. J. Emerson. 2001. Models and data on plant-enemy coevolution. Annual Review of Genetics 35:49-499.

Bergelson, J., M. Kreitman, E. A. Stahl, and D. Tian. 2001. Evolutionary dynamics of plant *R*-genes. Science 292:2281-2285.

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

Burdon, J. J. and P. H. Thrall. 1999. Spatial and temporal patterns in coevolving plant and pathogen associations. The American Naturalist 153:S15-S33.

Caicedo, A. L. and B. A. Schaal. 2004. Heterogeneous evolutionary processes affect *R* gene diversity in natural populations of *Solanum pimpinellifolium*. PNAS 101:17444-17449.

Calderon, I., N. Ortega, S. Duran, M. Becerro, M. Pascual, and X. Turon. 2007. Finding the relevant scale: clonality and genetic structure in a marine invertebrate (*Crambe crambe*, Porifera). Molecular Ecology 16:1799-1810.

Chung, M. G. and B. K. Epperson. 2000. Clonal and spatial genetic structure in *Eurya emarginata* (Theaceae). Heredity 84:170-177.

Dangl, J. L. and J. M. McDowell. 2006. Two modes of pathogen recognition by plants. PNAS 103:8575-8576.

Dinoor, A. 1970. Sources of oat crown rust resistance in hexaploid and tetraploid wild oats in Israel. Canadian Journal of Botany 48:153-161.

Dodds, P. N., G. J. Lawrence, A. –M. Catanzariti, T. The, C. –I. A. Wang, M. A. Ayliffe, B. Kobe, and J. G. Ellis. 2006. Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. PNAS 103:8888-8893.

Dorsey, E. 1937. Induced polyploidy in wheat and rye. Journal of Heredity 27:154-160.

Douhovnikoff, V. and R. S. Dodd. 2003. Intra-clonal variation and a similarity threshold for identification of clones: application to *Salix exigua* using AFLP molecular markers. Theoretical and Applied Genetics 106:1307-1315.

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

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

Etterson, J. R. and R.G. Shaw. 2001. Constraint to adaptive evolution in response to global warming. Science 294:151-154.

Falush, D., M. Stephens, and J. K. Pritchard. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular Ecology Notes 7:574-578.

Flor, H. H. 1971. Current status of the gene-for-gene concept. Annual Review of Phytopathology 9:275-296.

Garrett, K. A., S. P. Dendy, E. E. Frank, M. N. Rouse, and S. E. Travers. 2006. Climate change effects on plant disease: genomes to ecosystems. Annual Review of Phytopathology 44:498-509.

Garrett, K. A. and C. C. Mundt. 1999. Epidemiology in mixed host populations. Phytopathology 89:984-990.

Guo, Y. P., C. Vogl, M. van Loo, and F. Ehrendorfer. 2006. Hybrid origin and differentiation of two tetraploid *Achillea* species in East Asia: molecular, morphological and ecogeographical evidence. Molecular Ecology 15:133-144.

Gustafson, D. J., D. J. Gibson, and D. L. Nickrent. 2004. Conservation genetics of two co-dominant grass species in an endangered grassland ecosystem. Journal of Applied Ecology 41:389-397.

Gustafson, D. J., D. J. Gibson, and D. L. Nickrent. 1999. Random amplified polymorphic DNA variation among remnant big bluestem (*Andropogon gerardii* Vitman) populations from Arkansas' Grand Prairie. Molecular Ecology 8:1693-1701.

Hammerli, A. and T. B. H. Reusch. 2003. Genetic heighbourhood of clone structures in eelgrass meadows quantified by spatial autocorrelation of microsatellite markers. Heredity 91:448-455.

Hedren, M., M. F. Fay, and M. W. Chase. 2001. Amplified fragment length polymorphisms (AFLP) reveal details of polyploid evolution in *Dactylorhiza* (Orchidaceae). American Journal of Botany 88:1868-1880.

Huber, L. and T. J. Gillespie. 1992. Modelling leaf wetness in relation to plant disease epidemiology. Annual Review of Phytopathology 30:553-577.

Hulbert, S. H., C. A. Webb, S. M. Smith, and Q. Sun. 2001. Resistance gene complexes: evolution and utilization. Annual Review of Phytopathology 39:285-312.

Jiang, H., C. Wang, L. Ping, D. Tian, and S. Yang. 2007. Pattern of LRR nucleotide variation in plant resistance genes. Plant Science 173:253-261.

Jump, A. S. and J. Penuelas. 2007. Extensive spatial genetic structure revealed by AFLP but not SSR molecular markers in the wind-pollinated tree, *Fagus sylvatica*. Molecular Ecology 16:925-936.

Kane, N. C. and L. H. Rieseberg. 2007. Selective sweeps reveal candidate genes for adaptation to drought and salt tolerance in commom sunflower, *Helianthus annuus*. Genetics 175:1823-1834.

Keeler, K. H. 1990. Distribution of polyploid variation in big bluestem (*Andropogon gerardii* Poaceae) across the tallgrass prairie region. Genome 33:95-100.

Keeler, K. H. 1992. Local polyploid variation in the native prairie grass *Andropogon gerardii*. American Journal of Botany 79:1229-1232.

Keeler, K. H. 2004. Impact of intraspecific polyploidy in *Andropogon gerardii* (Poaceae) populations. American Midland Naturalist 152:63-74.

Keeler, K. H. and G. A. Davis. 1999. Comparison of common cytotypes of *Andropogon gerardii* (Andropogoneae, Poaceae). American Journal of Botany 86:974-979.

Keeler, K. H., C. F. Williams, and L. S. Vescio. 2002. Clone size of *Andropogon gerardii* Vitman (Big Bluestem) at Konza Prairie, Kansas. American Midland Naturalist 147:295-304.

Kosman, E., and J. Leonard. 2005. Similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid, and polyploid species. Molecular Ecology 14:415-424.

Lane, D. R., D. P. Coffin, and W. K. Lauenroth. 2000. Changes in grassland canopy structure across a precipitation gradient. Journal of Vegetation Science 11:359-368.

Leonard, K. J. 1969. Genetic equilibria in host-pathogen systems. Phytopathology 59:1858-1863.

Leonard, K. J. 1977. Selection pressures and plant pathogens. Annals of the New York Academy of Science 287:207-222.

Leonard, K. J. 1994. Stability of equilibria in a gene-for-gene coevolution model of host-parasite interactions. Phytopathology 84:70-77.

Leonard, K. J. and R. J. Czochor. 1980. Theory of genetic interactions among populations of plants and their pathogens. Annual Review of Phytopathology 18:237-258.

Lythgoe, K. A. and A. F. Read. 1998. Catching the Red Queen? The advice of Rose. Trends in Ecology and Evolution 13:473-474.

Manel, S., M. K. Schwartz, G. Luikart, and P. Taberlet. 2003. Landscape genetics: combining landscape ecology and population genetics. Trends in Ecology and Evolution 18: 189-197.

Mauricio, R., E. A. Stahl, T. Korves, D. Tian, M. Kreitman, and J. Bergelson. 2003. Natural selection for polymorphism in the disease resistance gene *Rps2* of *Arabidopsis thaliana.* Genetics 163:735-746.

Meudt, H. M., and A. C. Clarke. 2007. Almost forgotten or latest practice? AFLP applications, analyses and advances. Trends in Plant Science 12:106-117.

Michelmore, R. W. and B. C. Meyers. 1998. Clusters of resistance genes in plants evolve by divergent selection and a birth-and death process. Genome Research 8:1113-1130.

Morgan, G. W. 2003. Effects of spring burning and fungicide application on growth of warm season grasses and their pathogens in a native tallgrass prairie, greenhouse, and monoculture. M.S. thesis, Kansas State University, Manhattan, KS.

Muriithi, L. M. and L. E. Claflin. 1997. Genetic variation of grain sorghum germplasm for resistance to *Pseudomonas andropogonis*. Euphytica 98:129-132.

National Oceanic and Atmospheric Administration (NOAA). 2003. Monthly normals of temperature, precipitation, and heating and cooling degree days, 1971-2000: Kansas and Missouri. Climatography of the United States No. 81.

Nevo, E., J. G. Moseman, A. Beiles, and D. Zohary. 1984. Correlation of ecological factors and allozymic variations with resistance to *Erysiphe graminis hordei* in *Hordeum spontaneum* in Israel: Patterns and application. Plant Systematics and Evolution 145:79-96.

Norrmann, G. A. and K. H. Keeler. 2003. Cytotypes of *Andropogon gerardii* Vitman (Poaceae): fertility and reproduction of aneuploids. Botanical Journal of the Linnean Society 141:95-103.

Norrmann, G. A., C. L. Quarin, and K. H. Keeler. 1997. Evolutionary implications of meiotic chromosome behavior, reproductive biology, and hybridization in 6X and 9X cytotypes of *Andropogon gerardii* (Poaceae). American Journal of Botany 84:201-207.

Page, R. D. M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences 12:357-358.

Peck, J. R., J. M. Yearsley, and D. Waxman. 1998. Explaining the geographic distribution of sexual and asexual populations. Nature 391:889-892.

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. 2004. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. ISBN 3-900051-00-3, URL http://www.R-project.org.

Randolph, L. F. 1932. Some effects of high temperature on polyploidy and other variations in maize. PNAS 18:222-229.

Ronfort, J., E. Jenczewski, T. Bataillon, and F. Rousset. 1998. Analysis of population structure in autotetraploid species. Genetics 150:921-930.

Rousset, F. 2000. Genetic differentiation between individuals. Journal of Evolutionary Biology 13:58-62.

Rose. L. E., C. H. Langley, A. J. Bernal, and R. W. Michelmore. 2005. Natural variation in the *Pto* pathogen resistance gene within species of wild tomato (Lycopersicon). I. Functional analysis of *Pto* alleles. Genetics 171:345-357.

Rose, L. E., R. W. Michelmore, and C. H. Langley. 2007. Natural variation in the *Pto* disease resistance gene within species of wild tomato (Lycopersicon). II. Population genetics of *Pto*. Genetics 175:1307-1319.

Ruggeiro, M. V., T. B. H. Reusch, and G. Procaccini. 2005. Local genetic structure in a clonal dioecious angiosperm. Molecular Ecology 14:957-967.

Sherwin, W. B., F. Jabot, R. Rush, and M. Rossetto. 2006. Measurement of biological information with applications from genes to landscapes. Molecular Ecology 15:2857-2869.

Sork, V. L., K. A. Stowe, and C. Hochwender. 1993. Evidence for local adaptation in closely adjacent subpopulations of Northern Red Oak (*Quercus rubra* L.) expressed as resistance to leaf herbivores. The American Naturalist 142:928-936.

Stahl, E. A., G. Dwyer, R. Mauricio, M. Krietman, and J. Bergelson. 1999. Dynamics of disease resistance polymorphism at the *Rpm1* locus of *Arabidopsis*. Nature 400:667-671.

Suvanto, L. I. and T. B. Latva-Karjanmaa. 2005. Clone identification and clonal structure of the European aspen (*Populus tremula*). Molecular Ecology 14:2851-2860.

Tian, D., H. Araki, E. Stahl, J. Bergelson, and M. Kreitman. 2002. Signature of balancing selection in *Arabidopsis*. PNAS 99:11525-11530.

Tian, D., M. B. Traw, J. Q. Chen, M. Krietman, and J. Bergelson. 2003. Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. Nature 423:74-77.

Trybush, S., S. Hanley, K. Cho, S. Jahodová, M. Grimmer, I. Emelianov, C. Bayon, and A. Karp. 2006. Getting the most out of fluorescent amplified fragment length polymorphism. Canadian Journal of Botany 84:1347-1354.

Tuskan, G. A., S. DiFazio, S. Jansson, J. Bohlman, I. Grigoriev, U. Hellsten, N. Putnam, S. Ralph, S. Rombauts, A. Salamov, J. Schein, L. Sterck, A. Aerts, R. R. Bhalerao, R. P. Bhalerao, D. Blaudez, W. Boerjan, A. Brun, A. Brunner, V. Buscov, M. Campbell, J. Carlson, M. Chalot, J. Chapman, G.-L. Chen, D. Cooper, P. M. Coutinho, J. Couturier, S. Covert, Q. Cronk, R. Cunningham, J. Davis, S. Degroeve, A. Déjardin, C. dePamphilis, J. Detter, B. Dirks, I. Dubchak, S. Duplessis, J. Ehlting, B. Ellis, K. Gendler, D. Goodstein, M. Gribskov, J. Grimwood, A. Groover, L. Gunter, B. Hamberger, B. Heinze, Y. Helariutta, B. Henrissat, D. Holligan, R. Holt, W. Huang, N. Islam-Faridi, S. Jones, M. Jones-Rhaodes, R. Jorgensen, C. Joshi, J. Kangasjärvi, J. Karlsson, C. Kelleher, R. Kirkpatrick, M. Kirst, A. Kohler, U. Kalluri, F. Larimer, J. Leebens-Mack, J.-C. Leplé, P. Locascio, Y. Lou, S. Lucas, F. Martin, B. Montanini, C. Napoli, D. R. Nelson, C. Nelson, K. Nieminen, O. Nilsson, V. Pereda, G. Peter, R. Philippe, G. Pilate, A. Poliakov, J. Razumovskaya, P. Richardson, C. Rinaldi, K. Ritland, P. Rouzé, D. Ryaboy, J. Schmutz, J. Schrader, B. Segerman, H. Shin, A. Siddiqui, F. Sterky, A. Terry, C.-J. Tsai, E. Uberbacher, P. Unneberg, J. Vahala, K. Wall, S. Wessler, G. Yang, T. Yin, C. Douglas, M. Marra, G. Sandberg, Y. Van de Peer, and D. Rokhsar. 2006. The genome of Black Cottonwood *Populus trichocarpa* (Torr. & Gray). Science 313:1596-1604.

Vekemans, X., T. Beauwens, M. Lemaire, and I. Roldan-Ruiz. 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragement size. Molecular Ecology 11:139-151.

Vekemans, X. and O. J. Hardy. 2004. New insights from fine-scale spatial genetic structure analyses in plant populations. Molecular Ecology 13:921-935.

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

Wright, S. 1943. Isolation by distance. Genetics 28:114-138.

Zhao, B. Y., E. Ardales, E. Brasset, L. E. Claflin, J. E. Leach, and S. H. Hulbert. 2004. The *Rxo1/Rba1* locus of maize controls resistance reactions to pathogenic and non-host bacteria. Theoretical and Applied Genetics 109:71-79.

Zhao, B., X. Lin, J. Poland, H. Trick, J. Leach, and S. Hulbert. 2005. A maize resistance gene functions against bacterial streak disease in rice. PNAS 102:15383-15388.

Appendix A - Review of Methods for Determining Spatial Genetic Structure

Several methods are available for measuring spatial structure (see Bonin et al., 2007 for review). Sherwin et al. (2006) suggest the use of the Shannon index as a standardized measure of diversity for different hierarchical levels of information such as genetic diversity and species diversity. Baums et al. (2006) used Simpson's index, Fager's evenness measure and statistics involving genotype number and frequency to quantify diversity of a coral, *Acropora palmata*. AMOVA (Analysis of Molecular Variance; Excoffier et al., 1992) can be used to analyze molecular data at different population subdivisions. Conventionally, measures of spatial autocorrelation are used to estimate spatial structure when the specific geographic locations of organisms sampled is known. Statistics include measures such as Moran's I (Suvanto and Latva-Karjanmaa, 2005), '*Sp*' (Vekemans and Hardy, 2004; Jump and Penuelas, 2007; Ruggiero et al., 2005; Alberto et al., 2005), and F-statistics with random permutations to test for significance (Mantel tests)(Jump and Penuelas, 2007; Vekemans and Hardy, 2004; Alberto et al., 2005).

The study system and questions determine which measure is the most appropriate. For example, Moran's I is used for determining not only spatial genetic structure, but also whether clone aggregation or seed dispersal were the mechanisms behind spatial genetic structure in *Eurya emarginata* by comparing spatial autocorrelations of data with and without clones (Chung and Epperson, 2000). Also, the shape of F-statistic by distance curve may be used to infer relative importance of pollen vs. seed dispersal (Vekemans and Hardy, 2004). All of the above techniques except the simple diversity indices (Shannon and Simpson's) rely upon the use of gene flow estimates such as F-statistics that have been derived for organisms with simply ploidy and inheritance. The simple diversity indices can only be used when analyzing data categorically.

Though some F-statistics have been developed for autotetraploids (Ronfort et al., 1998), we know of no F-statistics based on assumptions that allow for intraspecific variation in ploidy and mode of inheritance, as is the case for *A. gerardii* (Norrmann et al., 1997).

The recently developed program STRUCTURE can assign individuals into populations based on sequence or molecular marker data (Pritchard et al., 2000). An update to the program allows dominant marker data (such as AFLP) to be analyzed for diploids, and codominant marker data for polyploids (Falush et al., 2007). For complex polyploid organisms such as *A. gerardii*, analysis of dominant markers in STRUCTURE is not yet possible.

AFLP data has been used to assay spatial genetic structure. In a wind pollinated tree (*Fagus sylvatica*), Jump and Penuelas (2007) found that spatial genetic structure exists up to 110 meters using AFLP markers and that 100-150 AFLP markers are sufficient to identify spatial genetic structure. AFLP markers have also been used in studies of individuals of varying cytotypes. Hedren et al. (2001) found that AFLP markers reveal polyploid evolution in *Dactylorhiza* (allopolyploids have a single, not multiple origin). Similarly, Guo et al. (2006) resolved the evolution of tetraploid *Achillea* species using AFLP markers. Experimental and modeling data suggest AFLP analyses identify a much higher proportion of smaller size fragments (Vekemans et al. 2002). Size homoplasy can be a problem with only 88% of amplified fragments expected to be detected after assuming 65 fragments are detected. When AFLP markers are applied to studies of clonal populations, two types of variation in AFLP banding patterns arise within clones: (1) scoring error and (2) somatic mutation within clones. Douhovnikoff and Dodd (2003) found that variation among stems within a clone accounted for more variation within a clone than did lab error. A threshold similarity value was established by Douhovnikoff and Dodd (2003), above which samples were thought to be of the same clone.