CONSERVATION GENETICS OF THE FLORIDA GRASSHOPPER SPARROW
(*Ammodramus savannarum floridanus*)

By

NATALIE LYNNETTE BULGIN, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

In Partial Fulfillment of the Requirements for the Degree

Master of Science

McMaster University

MASTER OF SCIENCE (2000) McMaster University

(Biology) Hamilton, Ontario

TITLE: Conservation Genetics of the Florida Grasshopper Sparrow (*Ammodramus savannarum floridanus*)

AUTHOR: Natalie Bulgin

SUPERVISOR: Dr. H.L. Gibbs

NUMBER OF PAGES: x, 138

ACKNOWLEDGEMENTS

I would first like to thank my supervisor, Dr. H. Lisle Gibbs for his time, patience, and commitment to my transition from a naïve undergraduate geneticist to a less, but still slightly naïve, molecular ecologist. I also thank my committee members, Drs. Bradley N. White and Susan A. Dudley, for stimulating discussion and suggestions regarding my thesis and science in general. I am grateful to our collaborator, Dr. Peter D. Vickery, for organizing the project, for insight into the biology of the species, and for making me realize I am not as much of a naturalist as I once tended to believe. Also, I thank Dr. Allan J. Baker, Centre for Biodiversity and Conservation Biology at the University of Toronto for the use of his automated sequencer, and technical and analytical support. Also, I thank Dr. John C. Avise for donating purified mtDNA from various *Ammodramus* species. I am grateful to Dr. Brian Golding and Liisa Koski for assistance with the phylogenetic analyses.

My utmost appreciation goes to Tylan Dean and Dusty Perkins for collection of Grasshopper Sparrow blood samples throughout the United States - God and Bonnie know I couldn't have done it. Special thanks is given to Oliver Haddrath, Centre for Biodiversity and Conservation Biology at the University of Toronto, for help designing the mitochondrial DNA primers, and for technical and analytical support. My sincerest thanks to Liliana De Sousa also for technical and analytical support. Je remercie Marylène Boulet pour avoir commentè les versions préliminaires du manuscrit. Finally, I would like to acknowledge my appreciation of all members of Dr. H. Lisle Gibbs' lab (Liliana De Sousa, Marylène Boulet, Bonnie Woolfenden, Lisa Tabak, Becky Pearce, and

Kevin Shonk) and Dr. Allan J. Baker's lab (Tara Paton, Maryann Burbidge, Nicola Wade, Oliver Haddrath, Andrew Given, Annette Greenslade, and Carol Croke).

In upholding my unsentimental-type reputation, I will not address my utmost appreciation for the emotional support from my friends Chantalle, Tim, and Carla, my best-friends Jennifer and Jeff, and my parents, Joyce Rhoda Ethel and Edwin Claude (thank goodness my family isn't one for inheriting names). However, I do hope that they realize this appreciation, and that I have, can, and will return the same.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 2

Chapter 3

LIST OF FIGURES

Chapter Two

Chapter Three

Chapter 1 – General Introduction

DNA-based evolutionary studies can provide valuable insight into the historical and contemporary genetic relationships of populations. Such studies have been especially valuable in studies of threatened species, where the information gained can be used to delimit boundaries of genetic continuity (Avise *et al.* 1987; Zink 1997). This information can then be used to identify conservation units (Moritz 1994). Specifically, management units are populations which exhibit significant differences in allele frequencies, and are thus thought to have only recently begun to diverge. In contrast, evolutionarily significant units show reciprocal monophyly and are believed to have been genetically isolated from each other for some time. Using the same datasets, conservation geneticists can also estimate dispersal rates among these populations to assess whether a metapopulation best fits a source-sink (Pulliam 1988) or a balanced dispersal (Diffendorfer *et al.* 1995) model.

Since the taxa of interest in conservation studies are generally populations, rapidly evolving, polymorphic DNA markers are needed in order to detect significant genetic differences, if present, among the closely related populations. To date, most studies have used mitochondrial DNA (mtDNA) control region sequence, which evolves at a rate of approximately 20.8% per million years (Quinn 1992). The mitochondrial genome is haploid, maternally inherited, and contained within the mitochondrium organelle separate from the nuclear genome (Brown *et al.* 1979). Although the majority of mtDNA codes

for proteins involved in respiration, a small (~1100 bp) fragment called the control region is non-coding and is believed to contain the mitochondrial genome's origin of replication (Hoelzel 1993). Because of its rapid rate of evolution and technical advantages (haploidy and uni-parental inheritance), conservation geneticists have been able to use control region sequence to identify genetic differentiation among populations which have only been separated on the order of 10 000 years (Avise and Nelson 1989; Quinn 1992; Milot *et al.* 2000).

Although mtDNA is well-suited for elucidating phylogenetic splits which have occurred very recently with respect to evolutionary time, the technical benefits of mtDNA are also some of its weaknesses. Avise *et al.* (1987) suggested that results based on mtDNA data alone should also be evaluated using a second independent dataset. Recently, conservation geneticists have begun using hypervariable nuclear microsatellite loci, which evolve at a rate comparable to mtDNA, and are bi-parentally inherited (Weber and Wong 1993). Because they are diploid and bi-parentally inherited, microsatellites are said to have an effective population size four times greater than that of mtDNA (Moore 1995), and thus generally show lower levels of genetic differentiation for the same populations. However, microsatellites have proven useful in detecting low levels of significant differentiation among closely related populations (Oyler-McCance *et al.* 1999; McDonald *et al.* 1999), and therefore offer a second independent genetic dataset to complement results based on mtDNA data.

The Grasshopper Sparrow (*Ammodramus savannarum*) is a New World grassland species (Ridgely and Tudor 1989). Unlike other *Ammodramus* species, the Grasshopper Sparrow is widely distributed throughout the Americas. Currently, four North American

subspecies are recognized primarily on the basis of plumage (Vickery 1996). Although most North American populations have suffered severe population declines over the last century, the Florida Grasshopper Sparrow (*A. s. floridanus*) is the only subspecies which has been designated as endangered under United States Fish and Wildlife Service guidelines (US Federal Register). A recent census of potentially reproductive individuals estimated the effective population size of Florida Grasshopper Sparrows at 1200 breeding adults (P.D. Vickery pers. comm.). The subspecies presently exists as four heterogeneously distributed sub-populations in south-central Florida prairie-like grasslands (Delany *et al.* 1999) and is discernable from the other three North American subspecies in that is non-migratory, prefers higher density nesting habitat, and is darker in plumage (Delany *et al.* 1995; Delany and Linda 1998; Vickery 1996).

Given that the endangered Florida subspecies has experienced severe population declines over a relatively short period of time and its sedentary behavior, it is quite possible that genetic differentiation may exist on a local spatial scale among the four Florida sub-populations, on a continental scale between Florida and populations of the other North American subspecies, or both. In any case, such information would be vital for identifying Grasshopper Sparrow conservation units in order to design appropriate management plans for the Florida subspecies.

In this study, I used mtDNA control region sequence and nuclear microsatellite markers to examine genetic relationships among locally- and continentally-distributed Grasshopper Sparrow populations including the Florida Grasshopper Sparrow. The ultimate goals of this project are: 1) to identify genetically differentiated groups of Grasshopper Sparrows on a local and continental scale; 2) to evaluate the extent to which

these groups currently exchange individuals; and 3) to assess the metapopulation dynamics of the Florida sub-populations. Based on the outcome of the study, I will recommend how conservation efforts for the Florida Grasshopper Sparrow should be directed in order to preserve the subspecies and to maximize overall genetic diversity within the species.

Literature Cited

Avise, J.C., J. Arnold, R.M. Ball, E. Bermingham, T. Lamb, J.E. Neigel, C.A. Reeb, and N.C. Saunders. 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Ann. Rev. Ecol. Syst.* 18: 489-522.

Avise, J.C. and W.S. Nelson. 1989. Molecular genetic relationships of the extinct dusky seaside sparrow. *Science* 243: 646-648.

Brown, W.M., M. George Jr., and A.C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A.* 76: 1967-1971.

Delany, M.F. and S.B. Linda. 1998. Characteristics of Florida grasshopper sparrow nests. *Wilson Bulletin* 110: 136-139.

Delany, M.F., C.T. Moore, D.R. Progulske Jr. 1995. Territory size and movements of Florida Grasshopper Sparrows. *J. Field Ornithol.* 66: 305-309.

Delany, M.F., H.M. Stevenson, and R. McCracken. 1985. Distribution, abundance, and habitat of the Florida Grasshopper Sparrow. *J. Wildl. Manage.* 49: 626-631.

Delany, M.F., P.B. Walsh, B. Pranty, and D.W. Perkins. 1999. A previously unknown

population of the Florida grasshopper sparrows on Avon ark Air Force Range. *Florida Field Naturalist* 27: 52-56.

Diffendorfer, J.E., M.S. Gaines, and R.D. Holt. 1995. Habitat fragmentation and movements of three small mammals (*Sigmodon, Microtus,* and *Peromyscus*) *Ecology* 76: 827-839.

Hoelzel, A.R. 1993. Evolution by DNA turnover in the control region of vertebrate mitochondrial DNA. *Current Opinion in Genetics and Development* 3: 891-895.

McDonald, D.B., W.K. Potts, J.W. Fitzpatrick, and G.E. Woolfenden. 1999. Contrasting genetic structures in sister species of North American scrub-jays. *Proc. R. Soc. Lond.* B. 266: 1117-1125.

Milot, E., H.L. Gibbs, and K.A. Hobson. 2000. Phylogeography and genetic structure of northern populations of the yellow warbler (*Dendroica petechia*). *Molecular Ecology* 9: 667-681.

Moore, W.S. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution* 49: 718-726.

Moritz, C. 1994. Applications of mitochondrial DNA analysis in conservation: a critical review. *Molecular Ecology* 3: 401-411.

Oyler-McCance, S., N.W. Kahn, K.P. Burnham, C.E. Braun, and T.W. Quinn. 1999. A population genetic comparison of large- and small-bodied sage grouse in Colorado using microsatellite and mitochondrial DNA markers. *Molecular Ecology* 8: 1457-1465.

Pulliam, H.R. 1988. Sources, sinks, and population regulation. *Am. Nat.* 132: 652-661.

Quinn, T.W. 1992. The genetic legacy of Mother Goose - phylogeographic patterns of lesser snow goose *Chen caerulescens caerulescens* maternal lineages. *Molecular Ecology* 1: 105-117.

Ridgely, R.S. and G. Tudor. 1989. The Birds of South America: The Oscine Passeriformes. University of Texas Press. p. 465.

Vickery, P.D. 1996. Grasshopper Sparrow (*Ammodramus savannarum*). *In*: The Birds of North America, no 239 (A. Poole and F. Gills, Eds.). Academy of Natural Sciences, Philadelphia, and American Ornithologists' Union, Washington, D.C.

Weber, J.L. and C. Wong. 1993. Mutation of human short random repeats. *Hum. Mol. Gen.* 2: 1123-1128.

Zink, R.M. 1997. Phylogeographic studies of North American birds. *In*: " Avian

Molecular Evolution and Systematics." (D.P. Mindell, Ed.) pp 301-324.

Academic Press.

Chapter 2 - Out of Florida: Phylogeography and Genetic Structure of North American Grasshopper Sparrows (*Ammodramus savannarum*)

ABSTRACT

DNA-based genetic analyses of bird populations can provide insights into the evolutionary history of a species as well as information which is useful for developing conservation plans for threatened populations. Here I present the results of an analysis which used mitochondrial DNA (mtDNA) control region sequence and six microsatellite loci to examine genetic differentiation and phylogeographic structure among five continentally-distributed populations of Grasshopper Sparrows (*Ammodramus savannarum*). These populations represented three subspecies: the southwestern subspecies, *A. s. ammolegus*, represented by a population from Arizona; three populations (Georgia, Ohio, and Massachusetts) of the eastern race, *A. s. pratensis*; and the endangered *A. s. floridanus* subspecies from Florida. I found significant differentiation between Florida and each of Georgia, Ohio, and Arizona based on mtDNA (F_{ST} = 0.114 to 0.119; $p < 0.008$), and between Florida and each of the other four populations in one measure of microsatellite variation (F_{ST}: θ = 0.014 to 0.026; $p <$ 0.008), but not in another (R_{ST}: ρ = 0.004 to 0.079; $p > 0.008$). In contrast, I did not observe significant genetic structure in any pair-wise comparison among Georgia, Ohio, Massachusetts, and Arizona in mtDNA (F_{ST} = -0.032 to 0.042) or in microsatellite DNA variation (θ = 0.005 to 0.012; ρ = -0.014 to 0.031). Thus, the major pattern of differentiation in both types of markers was between Florida and the other four

populations. Migration rates between Florida and the other populations were high based on different measures involving either type of marker. For example, based on the mtDNA data and a coalescence-based approach, I observed 48.0 female migrants per generation from Florida to the other four populations and 4.6 female migrants per generation from the four populations (pooled) into Florida. Estimates of migration based on microsatellite data was nearly symmetrical with 8.6 individuals per generation migrating from Florida to the pooled populations and 7.6 individuals per generation migrated from the pooled populations to Florida. F_{ST}-based measures of migration ranged from 4.6 female migrants per generation based on mtDNA data, to 15.4 migrants per generation based on microsatellite data. In contrast to the significant genetic structure observed, I detected a Type III (star-like) phylogeographic pattern based on both distance-based and parsimony analyses, although there is evidence that two Florida haplotypes were ancestral to all other North American Grasshopper Sparrow haplotypes. Finally, I did not find evidence supporting a recent (i.e. 0.2 to $4N_e$ generations ago) bottleneck event in any of the populations, or evidence suggesting that these populations may have expanded from a historical bottleneck (genetic imbalance index = 0.419 to 0.653). Together, these results imply that *floridanus* is genetically differentiated from and possibly ancestral to populations of the *pratensis* and *ammolegus* subspecies. Furthermore, the high ($Nm > 1$) migration rates suggest that the expansion occurred rapidly and recently. Given the significant genetic differentiation between Florida and each of Georgia, Ohio, Massachusetts, and Arizona based on two types of markers, I suggest that conservation plans for the *floridanus* subspecies focus on managing the

Florida population as a single management unit separate from the other four continental populations.

INTRODUCTION

The field of population biology has been revolutionized over the past decade following the recent application of genetic data to studies of intraspecific evolutionary history and population structure (Avise and Walker 1998). In particular, phylogenetic analysis of rapidly evolving, neutral DNA markers can provide insight into the historical relationships of closely related taxa (Avise *et al.* 1987). Also, assessing patterns of genetic differentiation can be used to identify populations which have only recently diverged from one another. Although the markers most often used in these studies are neutral (i.e. non-functional), they are used as indicators of the unique variation in fitness-associated traits which may only be present in certain groups of individuals. Ultimately, this information is useful for inferring factors promoting speciation and/or extinction and for developing conservation plans for threatened species.

Mitochondrial DNA Variation

To date, most studies of avian population genetics have used mitochondrial DNA (mtDNA) to examine intraspecific evolutionary relationships. The genome of this organelle is contained within the mitochondrium and hence evolves and is inherited separately from the nuclear genome (Brown *et al.* 1979). All animal mitochondrial genomes are haploid, maternally-inherited, lack introns, and do not recombine. In birds, the mitochondrial genome is roughly 16 kb in length and consists primarily of protein-coding sequence (Quinn 1997). A one kilobase fragment of non-coding sequence contains the mitochondrial origin of replication and displacement loop (D-loop), and is referred to as the control region (Hoelzel 1993). As a result of being non-coding, the

control region is believed to be selectively neutral and accumulates substitutions at a rate nearly 20 times greater than the rest of the mitochondrial genome (Quinn 1992). This region is further subdivided into three domains, of which domain III is most variable, followed by domain I, and finally II (Marshall and Baker 1997). The function of any of the domains remains unknown, but it is speculated that domain II contains the origin of replication because it is highly conserved across taxa, as would be needed for recognition by the replication machinery (Hoelzel 1993). Because of the genome's simple mode of inheritance, lack of recombination, and rapid rate of evolution, mtDNA is considered ideal for revealing low levels of significant differentiation among populations which may not be detected with more slowly evolving protein-coding nuclear DNA markers (Avise *et al.* 1987).

MtDNA has proven useful in elucidating the evolutionary history of a number of bird species. For instance, Milot *et al.* (2000) used mtDNA control region sequence to examine evolutionary relationships of the Yellow Warbler (*Dendroica petechia*) across seven Canadian provinces and Alaska. Phylogenetic analysis of the sequences revealed an east-west split of the haplotypes, and suggested that certain eastern haplotypes were ancestral to the western lineage. Furthermore, the results provided support for Mengel's (1964) theory that North American Yellow Warbler populations radiated westward post-Pleistocene from an eastern refugium approximately 10 000 years before present (ybp).

Similarly, Edwards (1993) used mtDNA control region sequence to examine phylogeographic structure among three populations of the Grey-Crowned Babbler (*Pomatostomus temporalis*) in Australia's Northern Territory. In contrast to the Yellow

Warbler study, Edwards observed no phylogeographic structure among haplotypes from different locations and therefore concluded that the paraphyletic pattern observed was a result of ongoing gene flow.

As a third example of the utility of mtDNA in evolutionary studies, Quinn (1992) used control region sequence to examine phylogenetic relationships of the Lesser Snow Goose (*Chen caerulescens caerulescens* L.). Interestingly, Quinn found two distinct mtDNA lineages among sympatrically distributed geese. Following Avise *et al.* (1987), Quinn concluded that the unusual pattern was a result of secondary contact between two maternal lineages which inhabited allopatric refugia during the Pleistocene ice age.

Nuclear DNA Markers

Based on these examples, it is clear that mtDNA variation is useful for inferring the evolutionary history of a species. However, Avise *et al.* (1987) pointed out that conclusions based solely on mtDNA should be viewed with caution because it is a single locus and only represents the evolutionary history of maternal lineages. Accordingly, they suggested that other, independent datasets be used to support findings based on mtDNA variation alone.

One possible source of such markers are nuclear microsatellite loci which have received much attention for their usefulness in parentage analyses and evolutionary studies (Neilsen and Palsbøll 1999). These selectively neutral DNA markers are short (1-5 bp; Messier *et al.* 1996), hypervariable repeat loci which evolve at a rate of 10^{-6} to 10^{-2} mutations/locus/generation (Weber and Wong 1993) comparable to the mtDNA control region (20.8% sequence divergence per million years; Quinn 1992). In contrast to

mtDNA loci, microsatellite DNA loci are diploid and are bi-parentally inherited. As a consequence of these characteristics, microsatellite loci are said to have an effective population size four times that of haploid, maternally inherited mtDNA markers (Moore 1995). This means that microsatellite loci require a longer period of time to show divergence among genetically isolated populations than mtDNA loci and therefore may be less sensitive to detecting low, but significant, levels of genetic differentiation. However, because of their rapid rate of evolution and high levels of polymorphism, microsatellite loci remain useful DNA markers which are capable of detecting low levels of differentiation similar to mtDNA loci (Jarne and Lagoda 1996).

Relatively few studies have analyzed population structure in birds using microsatellite data. Gibbs *et al.* (submitted) used six microsatellite loci to assay genetic differentiation among populations of the Yellow Warbler in Canada and Alaska. The results of this study implied low but significant genetic differentiation between some pairs of eastern and western populations. Because Milot *et al.* (2000) found strong phylogeographic structure between eastern and western North America Yellow Warbler mtDNA haplotypes, the study by Gibbs *et al.* (submitted) is an example of where microsatellite loci can detect genetic differentiation, but may not be as sensitive to revealing low levels of differentiation as mtDNA markers.

McDonald *et al.* (1999) used ten microsatellite loci to compare patterns of variation among populations of patchily distributed Florida Scrub-Jays (*Aphelocoma coerulescens*) and continuously distributed Western Scrub-Jays (*A. californica*). The results of the study revealed significant differentiation among all pairs of the patchily

distributed Florida Scrub-Jay populations and no significant differentiation among the continuously distributed Western species. From this, the authors suggested that any future management plans for the Florida Scrub-Jay should focus on individual populations versus the species as a whole.

Genetic Studies of Ammodramus *Sparrows*

Ammodramus species represent a genus of New World grassland sparrows found throughout North America and northern South America (Ridgely and Tudor 1987). Species of this genus are typically small and dull in coloration. However, many North American *Ammodramus* species exhibit substantial intraspecific plumage variation, which has been shown in some cases to be correlated with patterns of genetic variation.

For example, Avise and Nelson (1989) used restriction fragment-length polymorphism (RFLP) analysis to examine phylogenetic relationships among subspecies of the Seaside Sparrow (*A. maritimus*) in order to determine the closest relative of the extinct Dusky Seaside Sparrow (*A. m. nigrescens*). The analysis revealed two distinct mitochondrial lineages within the species, an Atlantic Coast lineage and a Gulf Coast lineage, and suggested that the Dusky Seaside Sparrow's belonged to the Gulf Coast lineage despite its geographic location on the Atlantic Coast.

Similarly, Rising and Avise (1993) used mtDNA RFLP analysis to examine evolutionary relationships among five Sharp-Tailed Sparrow (*A. caudacutus*) subspecies. Phylogenetic reconstruction of the restriction fragments revealed two clades separated by 1.2% sequence divergence. The two clades corresponded to a split between a northern lineage consisting of two subspecies and a southern lineage stemming from the three

other subspecies. As a result of this study, Rising and Avise suggested that the two mtDNA lineages be granted species versus subspecies status. Furthermore, Rising and Avise were also able to use this data to test alternative hypotheses regarding the species biogeographic history. In applying a molecular clock to the RFLP data, they concluded that the two lineages arose from two Pleistocene refugia approximately 600 000 ybp instead of one refugium 10 000 ybp.

In contrast to all other North American *Ammodramus* species, the Grasshopper Sparrow (*A. savannarum*) is a widely distributed grassland species (Vickery 1996). Currently four Grasshopper Sparrow subspecies are recognized in North America, all of which are described on the basis of plumage variation: the eastern race, *A. s. pratensis*; the southwestern United States subspecies, *A. s. ammolegus*; the northwestern United States and southwestern Canada subspecies, *A. s. perpallidus*; and a subspecies endemic to Florida, *A. s. floridanus* (Vickery 1996). The Florida Grasshopper Sparrow has experienced severe population declines over the last century primarily due to human-induced habitat fragmentation and destruction and received official United States Fish and Wildlife Service (USFWS) endangered status in 1991 (USFWS status report).

Conservation efforts for the recovery of the subspecies are currently focused on preservation of suitable habitat (Vickery 1996), but very little is known as to the genetic structure of the subspecies. Such information is vital for determining the historical relationships of these populations which may have had profound affects on the contemporary structure of the species (Moritz 1994). Once obtained, this information can be used to design management plans which focus on maintaining genetically

differentiated conservation units, which in essence maximizes intraspecific genetic diversity.

In the present study, I used two types of neutral DNA markers, mtDNA control region sequence and microsatellite DNA loci, to examine phylogeographic structure and genetic differentiation among five North American Grasshopper Sparrow continental populations representing *A. s. floridanus*, *A. s. ammolegus*, and *A. s. pratensis*. These data will be used to: 1) examine the evolutionary history of the species in North America; and 2) identify boundaries of management and evolutionarily significant units among the five populations to obtain information that can be used for designing an effective recovery plan for the endangered Florida Grasshopper Sparrow.

METHODS

Sample Collection

A total of 214 adult Grasshopper Sparrows were captured from five North American populations representing three subspecies: three *A. s. pratensis* populations (Georgia, $n = 19$; Ohio, $n = 34$; and Massachusetts, $n = 39$); one *A. s. floridanus* population (Florida, $n = 105$); and one *A. s. ammolegus* population (Arizona, $n = 17$) (Figure 1). All individuals were captured during the 1995 to 1998 breeding seasons (inclusive) using a combination of play backs and mist nets. Approximately 100 μL blood was collected from the brachial vein of each individual, suspended in 1 mL lysis buffer, and stored at –20 °C. DNA was then extracted using either DNAzol (Gibco BRL), QIAMP Tissue Kit (QIAGEN), or a standard phenol / chloroform procedure (e.g. Sambrook *et al.* 1989).

Data Collection

Mitochondrial DNA sequence

To avoid amplification of a nuclear mtDNA homologue (Numt; Sorenson and Quinn 1998), I initially obtained mtDNA control region sequence from a Cesium-Chloride purified Grasshopper Sparrow mtDNA sample as described below and then used this information to design species-specific mtDNA control region primers. Specifically, I first amplified a 9 kb fragment of the mtDNA genome using the A5Rev and 16S-CR primers (O. Haddrath pers. comm.) and the Expand Long Template PCR kit (Roche Diagnostics). I then used primers FCR13 (Baker and Marshall 1997) and BDW (O. Haddrath pers. comm.) to obtain sequence for designing Grasshopper Sparrow-

specific control region primers GRSP1 (5' CAC TCT TTG CCC CAT CAG ACA G 3') and 2 (5' ATA ATC TAC AGG GAC GTG GGG AT 3'), respectively, which together amplify a 1100 kb fragment of the control region (Figure 2). Subsequently, I used sequence obtained from this fragment to design a third control region primer, GRSP3 (5' GCC GAC CAT GAA TGG GGT CAA AT 3'; Figure 2). GRSP3 was necessary for obtaining high quality sequence of the first and second control region domains (approximately 700 bp total) since sequences obtained from GRSP1 and 2 were poor on either side of an inverted repeat in the third domain (Marshall and Baker 1997).

GRSP1 and GRSP2 were then used to amplify an 1100 bp fragment for 171 Grasshopper Sparrows, a Jamaican Grasshopper Sparrow (*A. s. savannarum*), and two outgroups (the Grassland Sparrow, *A. humeralis*; and the Yellow-Browed Sparrow, *A. aurifrons*). Interestingly, these primers did not amplify the control region of Le Conte's Sparrow (*A. lecontei*), Henslow's Sparrow (*A. henslowii*), or Sharp-Tailed Sparrow (*A. caudacutus*) which I initially intended to use as outgroups. Products were amplified via the polymerase chain reaction (PCR) by mixing 50 ng DNA with 3 μL PCR buffer (10X MBI-Fermentas), 75 pmol $MgCl_2$ (MBI-Fermentas), 150 pmol forward primer (GRSP1), 150 pmol reverse primer (GRSP2), 100 pmol dNTPs, and 0.75 U Taq Polymerase (MBI-Fermentas) in a total volume of 30 μL. The samples were subjected to an initial incubation at 94 °C for 3 minutes, followed by 32 cycles of 94 °C for 30 seconds, 62 °C for 30 seconds, and 72 °C for 1 minute in a PTC-100 thermocycler (MJ Research Inc.).

Amplified samples were then electrophoresed for one hour at 115 V on a 0.8% agarose/1X TAE gel containing ethidium bromide alongside a 1-kb DNA ladder (MBI-

Fermentas). The PCR products were viewed under low-intensity ultraviolet light, excised from the gel, and purified from the agarose by spinning for eight minutes at 7000 rpm through a polyester-stuffed 1 mL pipette tip into a 1.5 mL Eppendorf tube (Dean and Greenwald 1995). The first and second mtDNA control region domains were sequenced from 10 μL of the eluted PCR product using infrared fluorescently-labeled GRSP1 and GRSP3 primers, and the ThermoDYEnamic sequencing kit (Pharmacia-Amersham) according to the manufacturer's directions with an annealing temperature of 64 °C in a Techne 9600 thermocycler (Hybaid). The products were then electrophoresed overnight on a 6% acrylamide gel at 2000 V using a Li-COR automated sequencer. Sequences were automatically read from 5' to 3' in each direction per sample using the Li-COR BaseImagIR software package. Ambiguous sites were corrected by eye or from the chromatogram.

Finally, approximately 200 bp of the third control region domain was sequenced manually. A 10 μL volume of the eluted mtDNA amplification product was sequenced with GRSP2 using the Thermosequenase kit (Pharmacia-Amersham) according to the manufacturer's instructions with an annealing temperature of 62 °C. The sequencing products were electrophoresed 1.25 hours on a 6% acrylamide gel at 70 W. The gel was dried for 2 hours and exposed 48 hours to autoradiography film. Sequences were read by eye, added to the 3' end of the first and second control region domain sequences, and aligned using XESEE 3.2 (Cabot 1998).

Microsatellite DNA

I genotyped each individual at six microsatellite loci using primers for loci cloned from Swainson's Thrush (*Catharus ustulatus*) (*Ca*μ02; Gibbs *et al.* 1999), Brown-Headed Cowbird (*Molothrus ater*) (*Ma*μ23; Alderson *et al.* 1999); Yellow Warbler (*Dp*μ16; Dawson *et al.* 1997); and three which I isolated from a Grasshopper Sparrow genomic library using standard techniques (Dawson *et al.* 1997) (*As*μ09, 15, and 18; see Table 1). For radioactive amplifications, 100 pmol of the forward primer was end-labeled by incubating the primer with 0.6 μL T4 Polynucleotide Kinase (PNK) buffer (MBI-Fermentas), 0.4 μL water, 30 U PNK enzyme (MBI-Fermentas), and 25 μCi γ^{33}ATP (Amersham-Pharmacia Biotech) in a 6 μL reaction volume. This mixture was then incubated for 30 minutes at 37 °C followed by 10 minutes at 68 °C. Each sample was subsequently amplified at each locus by using a mix consisting of 50 ng DNA, 2.4 pmol end-labeled forward primer cocktail, 4.6 pmol cold forward primer, 7.0 pmol cold reverse primer, 20 pmol dNTPs, 25 pmol $MgCl_2$ (MBI-Fermentas), 1.0 μL Taq Buffer (MBI-Fermentas 10X), and 0.25 U Taq polymerase (MBI-Fermentas) in a total volume of 10 μL. Each sample was denatured for 2 minutes at 94 °C, followed by 45 seconds denaturation at 94 °C, 45 seconds the optimal annealing temperature (Table 1), and 45 seconds extension at 72 °C for 30 cycles.

Amplified samples were electrophoresed 2 to 2.5 hours on a 6% acrylamide gel at 70 W. Gels were dried and exposed for a minimum of 18 hours to autoradiography film

at room temperature. Individuals were scored at each locus against a clone of known size and two previously scored samples which were run on all gels.

Population Analysis

Patterns of Variation

For the mtDNA data, I used the program Arlequin (Schneider *et al.* 1996) to identify the number of haplotypes, nucleotide composition, number and type of nucleotide substitutions throughout the mtDNA dataset, average number of pair-wise differences among the haplotypes, and nucleotide diversity per site (π) both per and across all populations. For the microsatellite data, I tested for significant heterozygote deficiency per locus per population using the Hardy-Weinberg Exact Test for Heterozygote Deficiency sub-routine in GENEPOP (Raymond and Rousset 1995). I also tested for significant linkage disequilibrium per locus pair across populations (Linkage Disequilibrium sub-routine in GENEPOP), and the number and frequency of alleles per locus for each population (Basic Information sub-routine of GENEPOP).

Population structure

Genetic differentiation between pairs of populations was assayed for both mtDNA and microsatellite DNA variation. For mtDNA, I calculated a distance-based F_{ST} between each pair of populations using the program Arlequin. Distances were estimated using a Kimura two-parameter mutation model (Kimura 1980) and corrected for substitution rate heterogeneity among sites by applying a gamma correction value (Yang 1996) of 0.001 calculated in Modeltest (Posada and Crandall 1998). However, because Arlequin cannot use gamma values less than 0.01, I used a value of 0.01 in place of the

empirically derived gamma value. F_{ST} values were then tested for significant difference from zero using 1000 permutations.

I also tested genetic structure between population pairs using six microsatellite loci assuming i) an infinite alleles model (IAM) to generate theta (θ), a modified measure of F_{ST} (Weir and Cockerham 1984), and ii) a step-wise mutation model (SMM) to generate *rho* (ρ), a measure of R_{ST} (Slatkin 1995). Since the processes by which microsatellites mutate are unclear (DiRienzo *et al.* 1994), I assayed population structure using both models. Theta was calculated for each population pair and tested for significant difference from zero using 1000 permutations of alleles between samples using the program FSTAT (Goudet 1995). *Rho* values were calculated using the program RSTCALC (Goodman 1997). The variance in allele size was first standardized to account for differences in both sample sizes among populations and variances among loci. Values were then permuted 1000 times to test for statistical significance from zero. All *p*-values associated with F_{ST}, ρ, or θ values were adjusted for multiple comparisons within a given analysis using a sequential Bonferroni correction (Rice 1989).

Migration

To estimate the number of migrants per generation (Nm) between populations, I used a coalescence-based approach implemented in the program *Migrate* (Beerli and Felsenstein 1999). This approach is superior to conventional F_{ST}-based estimates of Nm in that it does not assume symmetric migration between populations. To estimate migration from the mtDNA dataset, I used the default DNA sequence model parameters and a transition / transversion ratio of 3.6 as calculated from Modeltest. Beerli and

Felsenstein recommend that analyses are repeated several times using theta (representing effective population size in this application) and Nm estimates from the previous runs until consistent values are obtained. Accordingly, I ran the analysis four times in total, the first using F_{ST}-based parameters as starting values for theta and Nm, and the second to fourth using the theta and Nm values estimated from the previous analysis (Beerli and Felsenstein 1999). The Nm values converged within four sets of simulations. For the microsatellite data I carried out simulations using the default parameters for a Brownian motion microsatellite model. Again, the microsatellite analyses were repeated and converged within four sets of simulations. For comparison, I also estimated symmetric F_{ST} migration rates among the groups based on F_{ST} values from i) the mtDNA data in Arlequin, and ii) the microsatellite data by substituting the F_{ST} value derived from a comparison of two populations into the equation $F_{ST} = 1/(4Nm + 1)$ and solving for Nm (Whitlock and McCauley 1998). In contrast to *Migrate*, F_{ST}-based migration estimates assume an island model in which the populations are assumed to be at mutation-drift equilibrium, have constant population sizes over time, exchange individuals symmetrically, and the genetic variants are not under selection and do not mutate. Although *Migrate* has fewer and different assumptions, the two migration estimates share the assumption that the genetic variation used to generate the estimate is neutral.

A major assumption of all migration estimates is that all shared haplotypes between or among populations are the result of gene flow. In other words, the methods do not compensate for haplotypes which may be shared as a result of recent common ancestry, or retained ancestral polymorphisms. To deal with this potential problem,

Edwards (1993) used an approach proposed by Slatkin and Maddison (1989) to determine whether the presence of a single haplotype in more than one population was a reflection of current migration or common ancestry. In their approach, Slatkin and Maddison found an inverse relationship between the number of between-population migration events (*s*), and the quotient of time since divergence (in generations) divided by long-term effective population size, N_e. In counting *s*, one can determine the expected quotient of t/N_e. Knowing the approximate time since divergence, the hypothetical N_e can be determined. Assuming that populations lose variation by drift at a constant rate (Avise *et al.* 1987), genetically isolated populations will no longer share haplotypes after a period of time that is dependent on the N_e of the ancestral population. Thus, ongoing gene flow is inferred between the two populations if the observed N_e is much greater than that expected for the species in question, whereas if the calculated N_e is much less than that expected, the migration estimates are assumed to be biased upwards due to the presence of retained ancestral polymorphisms. I used the neighbor-joining tree reconstructed from the mtDNA sequence data to determine *s* between genetically differentiated groups, which I then used to estimate t/N_e and calculate a long-term N_e for the Grasshopper Sparrow.

Bottleneck Events

I investigated if a recent or historical bottleneck event had left a genetic imprint on the population-level data. To detect a recent bottleneck (i.e. within 0.2 to $4N_e$ generations), I used the program *Bottleneck* (Cornuet and Luikart 1996; Luikart and Cornuet 1998) which relies on the theoretical assumption that populations which have experienced a recent bottleneck lose rare alleles prior to a decrease in heterozygosity.

The greater observed heterozygosity than that expected given the number of alleles in this particular situation is referred to as heterozygosity excess (Nei *et al.* 1975). Applying this assumption, the program tests the null hypothesis that the observed number of loci with heterozygosity excess is not significantly different from the expected number of loci with heterozygosity excess under an IAM, a SMM, and a two-parameter model (TPM) which combines the IAM and SMM. I tested the null hypothesis using the TPM, specifically a SMM to IAM ratio of 9:1, since this ratio is currently recognized as the most accurate microsatellite mutation model (Luikart and Cornuet 1998) and the qualitative statistical Sign test.

To detect a historical population bottleneck, I calculated the average genetic imbalance index across loci per population (Kimmel *et al* 1998; discussed in Gibbs *et al.* 1999). Briefly, this index measures the ratio of variance in allele size to expected homozygosity per locus:

$$\beta = \frac{2V}{(1/H_e^2 - 1)/2} \quad (1)$$

where β is the imbalance index, V represents the variance in allele size and H_e represents expected homozygosity. Indices are averaged across loci for a population index. This index is greater than one if a historical bottleneck has occurred and less than one if not. I calculated the index per and across loci for each population using the variance in allele size obtained from R_{ST} CALC and the expected homozygosity obtained from GENEPOP.

Phylogenetic Analysis

To detect phylogenetic structure among the five populations, I generated a minimum-spanning tree using the program MINISPNET (Excoffier *et al.* 1993). Distances among the 58 haplotypes, the Jamaican haplotype, and the two outgroup haplotypes were calculated as the number of base pair substitutions (including indels) between pairs of haplotypes using MEGA (Kumar, Tamura, and Nei 1993).

I also generated a neighbor-joining tree using PHYLIP (Felsenstein 1993). A total of 1300 datasets were generated in SEQBOOT, from which distances among haplotypes for each dataset were calculated in DNADIST. Distances were calculated using default parameters for a Jin and Nei substitution model (Jin and Nei 1990) with a gamma correction value of 0.001 and an empirically determined transition / transversion ratio of 3.6. These distances were then used to construct 1300 phylogenetic trees based on neighbor-joining (NEIGHBOR) algorithms. The final topology was constructed as a consensus neighbor-joining tree (CONSENSE) of the 1300 previously generated trees with bootstrap values to indicate the robustness of each node.

Finally, I constructed a neighbor-joining tree from microsatellite allele frequency data to examine evolutionary relationships among the five populations. Allele frequencies per locus per population were calculated in GENEPOP. In a recent review of the accuracy of microsatellite distance measures in phylogenetic reconstruction, Takezaki and Nei (1996) found that Nei *et al.*'s (1983) genetic distance and Cavalli-Sforza and Edwards' (1967) chord distance were equally reliable for generating the correct tree topology. Accordingly, I used both Nei's genetic distance (D_A) and the chord distance

(D_C) to calculate interpopulation genetic distances in PHYLIP's GENDIST sub-program. Two neighbor-joining trees were then constructed from each of the distance calculations in NEIGHBOR. To test the effect of any one locus on the topology of the trees, I repeated the reconstructions six times for each distance measure by removing one locus per simulation as a pseudo-jackknife re-sampling technique. Ultimately, if the topology of the tree is driven by the allele frequencies at any particular locus, the removal of this locus will result in a dramatically altered topology (G.B. Golding pers. comm.). Thus, this pseudo-jackknife re-sampling technique will ensure that the overall neighbor-joining tree based on microsatellite variation is not biased by the allele frequencies at any one of the six loci.

RESULTS

Levels of Variability: MtDNA

A total of 58 haplotypes were identified among the 171 Grasshopper Sparrow samples sequenced. Haplotypes were identified using an arbitrarily assigned haplotype, GRSP1, as the reference sequence for all other Grasshopper Sparrow haplotypes (Appendix 1). The 58 haplotypes were distinguishable by 45 polymorphic sites, including 32 transitions, 12 transversions, and two deletions (Figure 3). All but one polymorphic site had only two character states, the exception being site 131 with three states: A, G, and T. Light-strand nucleotide composition was relatively equal among cytosine (28.4%), adenine (28.9%), and thymine (28.3%), with a substantially lower percentage of guanine (14.4%) which is consistent with the general pattern seen in avian control region light-strand sequence (Baker and Marshall 1997).

Pairs of haplotypes differed by an average of 8.85 ± 4.10 substitutions (range: 1 to 14 nucleotide differences). The probability of two homologous sites being different among sequences (nucleotide diversity) was 0.010 ± 0.005. Forty-three of the 58 (74.1%) haplotypes were found in single populations (Appendix 2). Haplotypes GRSP4 and 6 were sampled in all five populations and had an observed frequency of 0.111 (19 of 171 individuals) each. Interestingly, four Florida individuals exhibited double-banding along the length of the sequence suggesting heteroplasmy (Mundy *et al.* 1996) and thus could not be included in analyses of mtDNA variation.

Individual populations showed substantial variation in the number of haplotypes per population, the mean number of substitutions among haplotypes, and nucleotide

diversity (Table 2). Specifically, the number of haplotypes per individual ranged from 0.378 (37 haplotypes for 98 individuals) in Florida to 0.857 (18 for 21) in Massachusetts. The average number of substitutions among haplotypes ranged from 5.531 ± 2.806 in Arizona to 11.859 ± 5.616 in Georgia. Nucleotide diversity ranged from 0.006 ± 0.004 in Arizona to 0.013 ± 0.007 in Georgia.

Levels of Variability: Microsatellite DNA

Patterns of variation in the six microsatellite loci are summarized in Table 3. The total number of alleles per locus across populations averaged 16 and varied from 8 (*Ma*μ23) to 22 (*As*μ18). Generally speaking, loci obtained from Grasshopper Sparrow libraries (*As*μ) were more variable (mean number of alleles = 20) than loci from the other three species (*Ma*μ, *Dp*μ, and *Ca*μ; mean number of alleles = 11). Each locus had at least three private alleles ranging in frequency from 0.2 to 1.0% of the total number of alleles surveyed per locus. Observed mean heterozygosity was 0.730 and ranged from 0.654 in Arizona to 0.780 in Massachusetts. A test of significant linkage disequlibrium for any pairs of the microsatellite loci revealed that the markers were unlinked since I did not observe significant linkage disequilibrium for any pair of the six loci ($p > 0.05$). Finally, I examined the possible existence of null alleles and non-random mating at each microsatellite locus per population by testing for significant deviations from Hardy-Weinberg equilibrium. Significant deviations ($p < 0.05$) from Hardy-Weinberg equilibrium were observed at loci *Dp*μ16, *Ca*μ02, and *As*μ15 in Florida; *As*μ18 in Ohio; *Dp*μ16 in Massachusetts, and *Dp*μ16 in Arizona (Table 4). However, the lack of

significant heterozygote deficiency across all loci within a population suggests that inbreeding and null alleles are rare or absent within these populations.

Population Structure

The overall F_{ST} value based on mtDNA was low ($F_{ST} = 0.074$; $p < 0.05$) but significant and suggested the presence of moderate genetic differentiation among the five populations. Closer examination of pair-wise comparisons of the five populations showed that the major pattern of differentiation was between Florida and each of the other four populations. Specifically, F_{ST} values were significantly different from zero in all pair-wise comparisons with Florida (range: 0.114 to 0.119; $p < 0.005$) except Massachusetts, but not in any pair-wise comparison of the other four populations (range: -0.032 to 0.042) (Table 5).

Based on microsatellite data, the overall theta value was also significantly different from zero ($\theta = 0.016$; $p < 0.001$). In particular, I observed significant theta values between Florida and each of Georgia, Ohio, Massachusetts, and Arizona (range: 0.014 to 0.026; $p < 0.005$, and a lack of significant genetic differentiation among Georgia, Ohio, Massachusetts, and Arizona (range: 0.005 to 0.009) (Table 6). In contrast, *rho* values were not significantly different for any pair-wise comparison (range: -0.014 to 0.031) (Table 6) or across all five populations ($\rho = 0.004$; $p > 0.005$).

Since two independent analyses revealed significant genetic differentiation between Florida and each of the other four populations, I repeated the population analyses by comparing two groups: i) Florida; and ii) a pooled group consisting of Georgia, Ohio, Massachusetts, and Arizona. F_{ST} values between the two groups were

significant ($p < 0.05$) based on mtDNA ($F_{ST} = 0.098$) and microsatellite ($\theta = 0.016$) variation based on an IAM. However, microsatellite variation based on a SMM was not significant ($\rho = 0.003$).

Migration

Based on the significant genetic differentiation observed between Florida and each of the other four populations, I used *Migrate* to estimate migration rates between Florida and the other four populations pooled. Based on the mtDNA data, migration estimates were generally high ($Nm > 1$) and asymmetric. Specifically, I observed twelve times the number of female migrants from Florida ($Nm = 48.0$; 95% CI: 21.5 to 60.3) than into Florida ($Nm = 4.6$; 95% CI: 1.8 to 6.7). High symmetric migration rates were estimated from the microsatellite data. Specifically, 7.6 (95% CI: 7.1 to 8.0) migrants per generation immigrated to Florida compared to 8.6 (95% CI: 7.8 to 8.8) migrants per generation moving from Florida to the other populations.

I also estimated migration between the two groups using conventional F_{ST}-based calculations of Nm. As stated previously, F_{ST}-based migration estimates assume an island model and provide only a symmetrical measure of migration between two populations. In contrast to the *Migrate* estimates, the mtDNA-based F_{ST} migration estimate was 4.6 female migrants per generation while the two-population theta-based estimate was 15.4 migrants per generation. The results of the F_{ST}-based estimates suggest that some information may be lost in assuming symmetrical migration, since *Migrate* calculated obviously asymmetrical migration between the two populations using either DNA marker.

To determine whether the migration estimates reflect historical or contemporary gene flow, I counted a total of 28 migration events (*s*) in the neighbor-joining tree. From the inverse relationship of *s* and t/N_e, 28 migration events extrapolates to a t/N_e quotient of $<<0.1$. Assuming that Grasshopper Sparrows radiated throughout North American post-Pleistocene, I estimated a time since divergence (t) of 10 000 generations. If I use a minimum quotient t/N_e of 0.1 and solve for N_e using 10 000 generations for t, then the minimum long-term effective population size of Grasshopper Sparrows is 100 000. This N_e value suggests that if the phylogeny solely reflects retained ancestral polymorphism, then the N_e of the ancestral populations would have to have been $\geq$ 100 000. Since this N_e is much larger than that assumed for typical passerine species (Barrowclough and Shields 1984), this suggests that the high degree of similarity between the two groups is likely due to contemporary gene flow.

Bottleneck Events

Two independent tests suggested that population bottlenecks did not occur in each of the five populations greater than $4N_e$ generations ago. The results of the recent bottleneck analysis detected a significantly fewer ($p < 0.05$) number of loci with heterozygosity excess than that expected in all five populations except Arizona (Table 7). Specifically, I observed heterozygosity deficiency at a minimum of four loci in each population suggesting that a bottleneck could have occurred in each population prior to $4N_e$ generations ago.

In terms of a historical population bottleneck event, the average genetic imbalance index per locus was 0.501, and ranged from 0.027 to 2.973. The average index across

loci per population was 0.501 and ranged from 0.419 in Florida to 0.653 in Arizona (Table 8). Since the indices are all < 1, I can conclude that a historical bottleneck did not occur in any of the five populations. In contrast, these values suggest that each population expanded from a non-bottlenecked ancestral population. Furthermore, the observation that all values overlap 1 when the 95% confidence levels are taken into consideration, suggests that these populations are in equilibrium (Kimmel *et al.* 1998).

Phylogenetic Patterns: MtDNA

A minimum-spanning tree based on raw pair-wise differences among the haplotypes showed extensive radiations from GRSP4 and GRSP6 (Figure 4). These "mother" haplotypes were connected by an intermediate haplotype, GRSP35. Of the 58 haplotypes sampled, 12 were shared between Florida and at least one of Georgia, Arizona, Massachusetts, and Ohio. The number of substitutions between any two haplotypes ranged from 1 to 14, while the two outgroup species, the Yellow-Browed Sparrow and the Grassland Sparrow, were separated from a Florida-specific haplotype, GRSP39, by a minimum of 117 substitutions. A Jamaican Grasshopper Sparrow sample was also sequenced for possible use as an outgroup, but appears to be an ingroup since it only differed by two nucleotide substitutions from GRSP24, which was unique to Massachusetts.

When rooted with control region sequence from the Yellow-Browed Sparrow, a neighbor-joining analysis revealed three distinct, but not significant, Grasshopper Sparrow clades (Figure 5). Overall, haplotypes were best described as sharing a Type III (or star-like) phylogeographic pattern (Avise *et al.* 1987). Specifically, 15 haplotypes

were shared by at least two populations, and in all cases haplotypes from the same population were paraphyletic. Bootstrapping confirmed the lack of statistical support for most of the clades in the tree as might be expected with an unsorted, shallow phylogeny. However, the basal clade to all other North American Grasshopper Sparrow haplotypes consisted of two haplotypes unique to Florida and had 72% bootstrap support.

Phylogeographic Patterns: Microsatellite DNA

Neighbor-joining trees based on the microsatellite DNA data using Nei's genetic distance (D_A) and the Cavalli-Sforza and Edwards chord distance (D_C) showed similar topologies (Figure 6). In both cases, two distinct groups were revealed: the first a combination of Arizona, Massachusetts, and Ohio; and the second a combination of Florida and Georgia. Branch lengths were typically longer using Nei's genetic distance and ranged from 0.006 to 0.054 compared to 0.001 to 0.027 using a chord distance measure. However, in four of the six jackknife simulations, the topology of the trees changed when loci *As*μ09 and *As*μ18 were removed (Appendix 4). For these simulations, Arizona, Georgia, and Ohio formed one cluster, and Florida and Massachusetts formed the other. Although the clustering observed when all six loci are included may be slightly affected by the inclusion of *As*μ09 and *As*μ18, two clades are still clearly visible. Additional loci will need to be sampled if further resolution of the true topology is necessary. However, I can conclude from the present analyses that Florida is most closely related to a *pratensis* population (either Massachusetts or Georgia), and that Ohio and Arizona are more closely related to each other than either is to Florida.

DISCUSSION

The Florida population showed significant differentiation in both mtDNA and microsatellite DNA variation from other sampled populations. I did not detect any significant genetic differences among the four other continental populations of these birds. In addition, I found weak evidence that some Florida haplotypes are ancestral to all other North American Grasshopper Sparrow haplotypes. Unrooted phylogenetic reconstruction of the populations using microsatellite data also suggested rapid and recent expansion of the species given the generally shallow topology observed. Below, I discuss the genetic differentiation and phylogeographic patterns observed in terms of biogeography and contemporary migration. I conclude with a brief discussion of the conservation implications of my results.

Methodological Considerations

For microsatellites, the two measures of differentiation that I used (F_{ST}, based on an IAM; and R_{ST}, based on a SMM) gave different results. Low but significant levels of differentiation between Florida and each of the other four populations were detected using F_{ST}, whereas no significant differentiation was observed with R_{ST}. The observed discrepancy between the F_{ST} and R_{ST} results has also been observed among populations of the Northern Water Snake, *Nerodia sipedon sipedon* (Prosser *et al.* 1999), the Yellow-Pine chipmunk (*Tamias amoenus*; Schulte-Hostedde *et al.* in prep.), the Komodo dragon (Ciofi and Bruford 1999), and the Yellow Warbler (Gibbs *et al.* 2000). This result also conflicts with Slatkin's (1995) conclusion that R-statistics are more sensitive to detecting population structure than are F-statistics. I suggest that the non-significant *rho* values are

a consequence of small sample sizes per population in combination with the effects of rare alleles on variance in allele size per locus.

Rho tests for significant differences in the within versus between population variance in allele size (Goodman 1997). In essence, *rho* is the fraction of the overall variance in allele size that is contributed by the difference between the total and average within population variance in allele size. Thus, populations with a substantially larger sample size will contribute more to the overall variance in allele size than those with smaller sample sizes, hence biasing the difference between the overall and the within population variance in allele size towards zero (Ruzzante 1998). Similarly, IAM-based measures of differentiation also calculate the proportion of overall variation that is contributed by the difference between the total and average within population variation, however, the frequency of alleles is used in place of variance in allele size in this situation (Whitlock and McCauley 1998).

Ruzzante (1998) tested the accuracy of various population structure parameters under specific conditions for six Atlantic cod (*Gadus morrhua*) microsatellite loci. He concluded that both F_{ST} and R_{ST} are accurate estimators of population structure regardless of the number of alleles per locus, the number of loci used, and differences in sample sizes among populations when $n > 50$ per population. However, exceptionally large variances in R_{ST} are observed when sample sizes are less than 50, leading to an inability to detect significant structure between populations which are, in fact, differentiated. Although it seems intuitive that differences in sample size would have the same effect on F_{ST}, Ruzzante found that differences in sample size did not affect the accuracy of F_{ST}.

The reasons for this observed difference were unclear. In addition, Ruzzante (1998) also found that extreme-lying rare alleles can substantially affect the variance in allele size (generally increasing variance) at any particular locus, especially where sample sizes are low (<50).

Using computer simulations of the effect of sample size and microsatellite mutation model, Gaggiotti *et al.* (1999) also determined that F_{ST} estimates are more reliable than R_{ST} when there are low sample sizes ($n < 500$), and a SMM to IAM ratio ranging from 0.75 to 0.90. Given both Ruzzante's (1998) and Gagiotti's (1999) observations, and the similar discrepancies observed in additional studies with limited sample sizes, I feel that there is substantial evidence in the literature to support my suggestion that the lack of population structure detected based on a SMM was due to low sample sizes in combination with rare alleles.

To test my predictions, I arbitrarily removed all alleles with a frequency less than 5% per population and then re-calculated *rho* values for all pair-wise comparisons of the five populations. All *rho* values increased following removal of rare alleles. Specifically, previously non-significant *rho* values were now significant for pair-wise comparisons of Florida and each of Georgia, Massachusetts, and Arizona (range before: -0.007 to 0.021; range after: 0.030 to 0.065), and for comparisons of Arizona and each of Ohio and Massachusetts (values before: 0.017 and 0.031, respectively; values after: 0.064 and 0.048, respectively).

To test the affect of differences in sample size, I repeated the analysis using all alleles and a randomly selected 25 Florida samples. After adjusting for differences in

sample size among the populations, *rho* values were significant for pair-wise comparisons of Florida and each of Georgia and Arizona (values before: 0.007 and 0.019, respectively; values after: 0.029 and 0.056, respectively), and for the comparison between Massachusetts and Arizona (0.031 before; 0.037 after). A comparison of the variances per locus per population for the original *rho* values, and the tests concerning the effect of low frequency alleles and differences in sample size can be found in Appendix 3.

Together, these findings suggest that rare alleles may reduce the overall power of the analysis for detecting significant genetic differences among populations where there are large differences in sample sizes. Given the results of my tests, and the findings of Ruzzante (1998) and Gaggiotti *et al.* (1999), it appears as if both differences in sample size when $n < 50$ and rare alleles can bias *rho* values towards zero. Therefore, I believe that the F_{ST} values may give a more accurate picture of the differentiation in microsatellite variation between sparrow populations. Furthermore, I suggest that researchers studying genetic differentiation of natural populations where sample sizes are inherently low should take great caution in interpreting R_{ST} results in the absence of F_{ST} values.

A second methodological issue I need to address is the general observation that the mtDNA-based F_{ST} values were almost six-fold greater in magnitude than microsatellite-based theta values even though the same patterns of differentiation were seen in both types of variation. This discrepancy may be attributed to: 1) male-biased dispersal; or 2) differences in the effective population size (N_e) of the markers. Since it is generally known that dispersal in passerine species is female-biased (Greenwood and

Harvey 1982), it seems unlikely that differences in the magnitude of F_{ST} versus theta are due to male-biased dispersal. Although I cannot reject this possibility entirely, the difference seems most likely due to the inherent differences in the effective population size of the markers. Being haploid and maternally inherited, mtDNA has one-quarter the effective population size of diploid, bi-parentally inherited microsatellite loci (Moore 1995). Consequently, mtDNA is expected to reach an equilibrium state four times faster than nuclear markers and thus is more likely to show stronger differentiation than the latter marker (Moritz 1994). Second, mtDNA evolves faster than nuclear DNA and therefore is believed to reach significant levels of differentiation earlier than nuclear markers (Gibbs *et al.* submitted). Furthermore, Birky *et al.* (1989) have shown that in order for nuclear DNA markers to show more subdivision than organellar loci, the ratio of the male N_e (N_m) to the female N_e (N_f) must be $< 1/7$. This is an unlikely scenario for the Grasshopper Sparrow since the species is monogamous (Vickery 1996), and therefore N_f is equal to N_m. Thus, it is possible that the higher levels of differentiation observed using mtDNA loci in comparison to microsatellite DNA loci may be a result of differences in the N_e of the markers.

Biogeographic Considerations and Ancestry

Two lines of evidence suggest that the Grasshopper Sparrow population ancestral to all others in North America was located in Florida. First, both distance- and parsimony-based phylogenetic analyses showed two Florida haplotypes as basal to all other North American haplotypes. Specifically, GRSP39 and 53, both of which were only found in Florida, were basal to all other Grasshopper Sparrow haplotypes with 72%

bootstrap support as seen in the neighbor-joining tree. Second, of the eight most basal haplotypes (GRSP2, 9, 16, 39, 41, 45, 50, and 53), five (GRSP39, 41, 45, 50, and 53) were unique to Florida and three (GRSP2, 9, and 16) were found in Florida and some combination of Massachusetts, Ohio, and Arizona at low frequencies. None of these haplotypes were only found in any of Massachusetts, Ohio, or Arizona. At first glance, all of these observations may seem like a sampling effect whereby Florida clusters most closely to the outgroup simply because more Florida birds were sampled in comparison to any one of the other populations. In essence, the larger the sample size, the better the representation of the variation in a given population. Thus, GRSP39 and 53 may have fallen out as basal because Florida's variation was more thoroughly sampled than any one of Georgia, Ohio, Massachusetts, or Arizona since more than four times the number of individuals were sampled in Florida than any one of the four other populations. However, if one considers only two groups of samples (i.e. Florida versus the others), sample sizes are comparable (98 and 73, respectively), and it can be assumed that I have sequenced a sample size representative of the actual amount of genetic variation in each of the groups.

North America experienced several major glacial advances and retreats during the late Pleistocene era (0.01 to 2 million ybp; Watts and Hansen 1994). Accordingly, I can hypothesize that the glaciation events could have played a major environmental role in the evolution of contemporary North American Grasshopper Sparrow populations since the northern migratory populations would have been forced to occupy a southern refugium to escape the advances of the ice sheet. If I assume that the Florida population

was ancestral, then a possible scenario for the current Grasshopper Sparrow distribution in North America was that the other continental populations were recently founded by radiation from a refugium in Florida. Consistent with many North American avian species evolutionary histories, radiation most likely followed retreat of the final Pleistocene ice sheet some 10 – 15 000 years ago in response to the newly available habitat (Zink 1997). Subsequent divergence between Florida and the non-Florida populations then occurred due to genetic drift associated with a cessation of gene flow (Avise and Walker 1998).

To support this hypothesis, I would need to estimate when the two groups began to diverge. However, this date is not possible to determine from my data. To do so, one requires the percent sequence divergence separating reciprocally monophyletic clades of mtDNA haplotypes (Avise and Walker 1998). Otherwise, precise dating of the ancestral coalescence of the lineages is obscured by overlap of within and between population genetic variation, which tends to overestimate dating when applying molecular clock methods. However, to gain a rough estimate of the age of North American Grasshopper Sparrows, I can apply a rate of 20.8% sequence divergence per million years for control region domain I (Quinn 1992) to my dataset to estimate the earliest coalescence event of the haplotypes versus the populations. The haplotypes differed by a maximum of nine nucleotide substitutions over the 282 bp (3.2%) of control region domain I sequence that I collected. By dividing 3.2% by 20.8%, I estimate the ancestral coalescence event of the haplotypes at 150 000 ybp. Although this event predates the retreat of the final

Pleistocene ice sheet, the date is also not corrected for within versus between population genetic variation and so could be less.

While it is possible that North American Grasshopper Sparrow populations radiated from a refugium in Florida, two alternative hypotheses exist. First, it is possible that the ancestors of North American Grasshopper Sparrows occupied two or more refugia during the Pleistocene, one of which was Florida while the other was located elsewhere. In this scenario, the birds inhabiting Florida remained in close proximity to this refugium following retreat of the last ice sheet, while at the same time birds inhabiting the alternative refugium(ia) rapidly expanded throughout continental North America. Minimal secondary contact subsequently occurred between the groups leading to the presence of some mtDNA haplotypes throughout the species North American distribution. Alternatively, the Florida population may not have been a refugium at all, but it is possible that the North American populations, including Florida, were instead formed from emigrants of an alternative refugium. However, this last hypothesis seems unlikely because if North American Grasshopper Sparrows radiated from a Western refugium, the Florida haplotypes would not be observed as basal (Milot *et al.* 2000). Further phylogenetic analysis including samples from the other subspecies is required to determine which, if any, alternative subspecies are ancestral to the North American populations.

Aspects of the paleoecological history of Florida during the Pleistocene can be used to assess these hypotheses. Grasshopper Sparrows presently occupy two major biogeographic realms which were not covered by ice in the Pleistocene: (i) the Southern

United States, including Florida; and (ii) southern Mexico and northern South America (Vickery 1996). The species typically exploits arid prairie-like habitats with a moderate density of tall-grass for nest cover (Delany and Stevenson 1985). Paleoecological surveys of central peninsular Florida have shown that the region currently inhabited by Grasshopper Sparrows was similar to that covering the area throughout the Pleistocene (Grimm *et al.* 1993). Although there is no direct evidence for or against the statement that Florida was used as a refugium during the Pleistocene, evidence exists that suitable habitat for the Grasshopper Sparrow was available at this time.

My second hypothesis was that the currently extant continental populations radiated from two or more allopatric refugia. It has been shown that populations which most likely occupied allopatric refugia during the Pleistocene exhibit Type II phylogeographic structure (Avise *et al.* 1987; Quinn 1992). This theory speculates that populations which inhabited allopatric refugia evolved independent of one another resulting in two or more discrete mtDNA lineages. Subsequent population expansion and secondary contact is then observed as distinct mtDNA clades observed in symptrically distributed populations (Avise *et al.* 1987) as seen in some bird species: the mallard (*Anas platyrhynchos*) and black ducks (*Anas rubripes*) (Avise *et al.* 1990); and blue tits (*Parus caerulescens*) (Taberlet *et al.* 1992). In contrast, my data exhibit Type III, or star-like, phylogeographic structure, which is characterized by phylogenetic continuity and geographical separation (Avise *et al.* 1987). This pattern is believed to be the result of rapid, recent population expansion associated with limited gene flow in the absence of strong physical barriers and is seen in many North American bird species (Ball and Avise

1990; Zink 1996; reviewed in Zink 1997). Given the two distinctly different phylogeographic patterns representing different intraspecific evolutionary relationships, it is unlikely that *floridanus* occupied a refugium separate from that of the *pratensis* and *ammolegus* populations.

Inconsistent with my hypothesis that Florida is ancestral is that such populations are generally believed to have more variation on average than derived populations (Bouzat *et al.* 1997). In contrast, the Florida population showed less variability than at least two other populations in the number of mtDNA haplotypes per individual, number of pair-wise differences, nucleotide diversity, and the number of microsatellite alleles per locus per individual than any of the other four populations. A possible explanation for this discrepancy is that the *floridanus* subspecies has suffered severe population declines over the last century due to extensive habitat destruction and fragmentation throughout its range (Delany and Linda 1998). Typically populations lose substantial genetic variation via genetic drift following a bottleneck event in combination with minimal immigration subsequent to the event (Bouzat *et al.* 1997). Thus, the population declines observed in Florida within the last 100 generations may have had a "bottleneck-like" affect which resulted in a loss of genetic variation beyond the pre-human impact levels. The current N_e of Florida Grasshopper Sparrows is approximately 1200 individuals (P.D. Vickery pers. comm.). Since the ability of *Bottleneck* to detect a recent bottleneck is $0.2N_e$, this means that the method can only detect a bottleneck in this particular population if it occurred at least 240 years ago. Hence, it is quite possible that the low genetic variation observed in Florida compared to the other four continental populations could reflect a

human-induced bottleneck event that occurred some time in the last century and thus is not detectable using the analyses currently available.

Migration

Migration analyses based on mtDNA showed strong asymmetric mtDNA-based migration rates between Florida and the other populations, with 10.4 times as many birds leaving Florida as arriving from the pooled population. The most plausible explanation for the strongly asymmetric migration is history. *Migrate* estimates migration rates averaged over the total number of generations since the ancestral coalescent event (Beerli and Felsenstein 1999). As a consequence, the migration rate may be biased upwards if high levels of migration were occurring over a short period of time during the taxon's past evolutionary history relative to the minimal contemporary migration rates. Thus, if the species radiated out of Florida post-Pleistocene and subsequently ceased to exchange migrants between the two groups, *Migrate* would average the high level of historical movement out of Florida against the low level of contemporary movement in either direction, resulting in an overestimate of current migration between the two groups.

The apparent genetic differentiation between two groups of North American Grasshopper Sparrows would lead one to expect little to no migration between the two units. To the contrary, I estimated high (i.e. $Nm > 1$) migration rates between the two groups regardless of the type of analysis or datum. While it is highly possible that migration can occur between the two groups, I believe that the magnitude of these estimates are further biased upwards due to not meeting all of *Migrate's* assumptions. For instance, *Migrate* makes the assumption that there is some degree of phylogenetic

structure among the populations (P. Beerli pers. comm.). In other words, it assumes that all alleles shared by two or more populations are a consequence of migration and does not compensate for alleles which are shared due to common ancestry. Thus, in a situation such as the Grasshopper Sparrow where there is incomplete lineage sorting, *Migrate* has the potential to overestimate the rate of migration between the two groups.

However, I must also consider the possibility that contemporary gene flow could be occurring despite the significant measures of genetic differentiation observed between Florida and the other four populations. Following Edwards' (1993) approach to distinguishing between ongoing gene flow and retained ancestral polymorphism, I calculated that in order for the number of putative migration events to be a result of common ancestry versus migration, the minimum long-term N_e of North American Grasshopper Sparrows was 100 000. This N_e is much larger than that expected for typical passerine species (Barrowclough and Shields 1984), and thus I must acknowledge that the phylogeny may in actuality reflect some degree of ongoing gene flow, versus the possibility of shared ancestry proposed above. Although this may at first seem to be in contradiction to the results of genetic differentiation between Florida and the other four populations, Edwards (1993) found a similar pattern among three Grey-Crowned Babbler species but concluded that both can simultaneously exist; ongoing gene flow inferred from an unstructured phylogeny may be occurring, but may be at low enough levels to be insufficient to counteract low, yet significant, levels of genetic differentiation which occur due to genetic drift.

While this explanation can also be applied to the Grasshopper Sparrow, an alternative possibility is that North American Grasshopper Sparrows did not radiate immediately following retreat of the Pleistocene glacier, which would directly affect the magnitude of N_e. Considering that Florida birds are highly sedentary, and that observations of either group outside its recognized breeding distribution are rare (P.D.Vickery unpublished data), I conclude that the latter explanation is most likely, but cannot completely discount the possibility of ongoing gene flow between the two groups.

Genetic Structure

Generally speaking, explanations for structure among North American avian populations have all been based on physical barriers to dispersal, including mountains, islands, and distance (Zink 1997). However, several studies including the present have observed structure between migrating and sedentary populations, and among populations of wholly sedentary species: the Canada Goose (*Branta canadensis*); Small-Bodied Sage Grouse (Oyler-McCance *et al.* 1999); the Spotted Owl (*Strix occidentalis*) (Barrowclough *et al.* 1999); the Ring-Necked Pheasant (*Phasianus colchicus*; Giesel *et al.* 1997); Clapper Rails (*Rallus longirostris*; Fleischer *et al.* 1995); and Seaside Sparrows (*Ammodramus maritimus*; Avise and Nelson 1989).

Similar patterns of genetic differentiation seen among Grasshoppper Sparrow populations were observed by Buerkle (1999) between two subspecies of Prairie Warblers (*D. discolor discolor* and *D. d. paludicola*), one of which (*paludicola*) is found in Florida. *D. d. discolor* is a migratory, habitat-generalist subspecies which inhabits the Atlantic states. In contrast, *D. d. paludicola* is sedentary and endemic to coastal

mangrove habitats in Florida; the two subspecies do not occur in sympatry. As in my study, Buerkle observed significant genetic differentiation in mtDNA variation between the two subspecies and a lack of reciprocal monophyly. However, in contrast to what I have proposed for the Grasshopper Sparrow, Buerkle (1999) postulated that the *paludicola* subspecies did not inhabit Florida during the Pleistocene since coastal mangrove habitats were believed to be absent in Florida at this time. Thus, he concluded that the current distribution of Prairie Warblers is a result of radiation from two allopatric refugia.

Another distinction between my study and Beurkle's is that although the current distributions of Prairie Warblers and Grasshopper Sparrows are both believed to have been established following retreat of the Pleistocene ice sheet, F_{ST} values were much higher in general among the Prairie Warbler populations. The existence of two or more Prairie Warbler refugia versus the one North American Grasshopper Sparrow refugium suggested might explain the higher F_{ST} values and stronger phylogeographic structure observed in the former species, as a result of prolonged lack of gene flow between the refugia during the Pleistocene. This historical separation may also explain why *paludicola* haplotypes are not basal to *discolor* haplotypes if the species did not radiate from a single refugium in Florida.

McDonald *et al.* (1999) also found similar patterns of variation among eastern and western populations of Scrub-Jays. In particular, they observed significant differences in microsatellite variation between populations of the Florida Scrub-Jay and the Western Scrub-Jay. As a result of this study, McDonald *et al.* suggested that populations of the

Florida Scrub-Jay be managed separately from the Western Scrub-Jay due to the inferred lack of migration among populations.

In conclusion, the results of my study and previous studies of non-migratory species in Florida support the notion that sedentary behavior may also be an important microevolutionary force in avian species.

Conclusions

Based on the significant genetic structure found between Florida and four additional North American Grasshopper Sparrow populations, and the lack of phylogenetic structure among these groups, I suggest that the Florida population be recognized as a management unit (Moritz 1994), separate from the other North American Grasshopper Sparrow populations. I have also proposed that sedentary behavior be considered an important microevolutionary force among avian populations in addition to the conventionally recognized physical barriers to gene flow. Finally, my study has suggested a novel North American radiation pattern following the Pleistocene glaciations. These results should thus also be added to the growing database concerning comparative phylgeography of North American avian species.

Future work of interest with respect to the Grasshopper Sparrow would include sampling individuals from the other nine subspecies, which may increase the resolution of the species phylogeny. In doing so, it may be possible to determine if North American populations are actually derived from a tropical subspecies as proposed in our second hypothesis regarding biogeography and ancestry. Also, the low variation observed in Florida despite the strong evidence of its ancestral role suggests that the

recent human-induced habitat fragmentation and degradation have had detrimental effects on the viability of the subspecies. Accordingly, an assessment of the genetic variation within the subspecies prior to the 1900's may prove useful in assessing the impacts of such anthropogenic effects.

LITERATURE CITED

Alderson, G.A., H.L. Gibbs, and S.G. Sealy. 1999. Parentage and kinship analyses in an obligate brood parasitic bird, the brown-headed cowbird (*Molothrus ater*) using microsatellite DNA markers. *Journal of Heredity* 90: 182-190.

Avise, J.C., J. Arnold, R.M. Ball, E. Bermingham, T. Lamb, J.E. Neigel, C.A. Reeb, and N.C. Saunders. 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Ann. Rev. Ecol. Syst.* 18: 489-522.

Avise, J.C. and W.S. Nelson. 1989. Molecular genetic relationships of the extinct dusky seaside sparrow. *Science* 243: 646-648.

Avise, J.C. and D. Walker. 1998. Pleistocene phylogeographic effects on avian populations and the speciation process. *Proc. R. Soc. Lond. B.* 457-463.

Baker, A.J. and H.D. Marshall. 1997. Mitochondrial control region sequences. *In*: "Avian Molecular Evolution and Systematics" (D.P. Mindell, Ed.), pp. 51-82. Academic Press.

Ball Jr., R.M., S. Freeman, F.C. James, E. Bermingham, and J.C. Avise. 1988.

Phylogeographic population structure of red-winged blackbirds assessed by mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A.* 85: 1558-1562.

Ball Jr., R.M. and J.C. Avise. 1992. Mitochondrial DNA phylogeographic differentiation among avian populations and the evolutionary significance of subspecies. *Auk* 109: 626-636.

Barrowclough, G.F. and G.F. Shields. 1984. Karyotypic evolution and long-term effective population sizes of birds. *Auk* 101: 99-102.

Barrowclough, G.F., R.J. Guttierez, and J.G. Groth. 1999. Phylogeography of spotted owl (*strix occidentalis*) populations based on mitochondrial DNA sequences: Gene flow, genetic structure, and a novel biogeographic pattern. *Evolution* 53: 919-931.

Beerli, P. and J. Felsenstein. 1999. Maximum likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics* 152:763-773.

Birky, C.W., Jr., P. Fuerst, and T. Maruyama. 1989. Organelle gene diversity under migration, mutation, and drift: equilibrium expectation, approach to equilibrium,

effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics* 121: 613-627.

Bouzat, J.L., H.H. Cheng, H.A, Lewin, R.L. Westemeier, J.D. Brawn, and K.N. Paige. 1998. Genetic evaluation of a demographic bottleneck in the greater prairie chicken. *Conservation Biology* 12: 836-843.

Brown, W.M., M. George Jr., and A.C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A.* 76: 1967-1971.

Buerkle, C.A. 1999. The historical pattern of gene flow among migratory and nonmigratory populations of prairie warblers (Aves: Parulinae). *Evolution* 53: 1915-1924.

Cavalli-Sforza., L.L. and A.W.F. Edwards. 1967. Phylogenetic Analysis: Models and estimation procedures. *American Journal of Human Genetics* 19: 233-257.

Ciofi, C. and M.W. Bruford. 1999. Genetic structure and gene flow among komodo dragon populations inferred by microsatellite loci analysis. *Molecular Ecology* 8: S17-S30.

Cornuet, J.-M. and G. Luikart. 1996. Description and power analysis of two tests for

detecting recent population bottlenecks from allele frequency data. *Genetics* 144: 2001-2014.

Dawson, R.J.G., H.L. Gibbs, K.A. Hobson, and S. Yzerniac. 1997. Isolation of microsatellite DNA markers from a passerine bird, the yellow warbler. *Heredity* 79: 506-514.

Dean, A.D. and J.E. Greenwald. 1995. Use of filtered pipet tips to elute DNA from agarose gels. *Biotechniques* 18: 980.

Delany, M.F. and S.B. Linda. 1998. Characteristics of Florida grasshopper sparrow nests. *Wilson Bulletin* 110: 136-139.

DiRienzo, A., A.C. Peterson, J.C. Garza, A.M. Valdes, M. Slatkin, and N.B. Friemer. 1994. Mutational processes of simple-sequence repeat loci in human populations. *Proc. Natl. Acad. Sci. U.S.A.* 91: 3166-3170.

Edwards, S.V. 1993. Mitochondrial gene genealogy and gene flow among island and mainland populations of a sedentary songbird, the grey-crowned babbler (*Pomatostomus temporalis*). *Evolution* 47: 1118-1137.

Excoffier, L. 1993. MINISPNET. Genetics and Biometry Lab, Department of

Anthropology, University of Geneva, Geneva, Switzerland.

Felsenstein, J. 1993. Phylogeny Inference Package, v. 3.5c, University of Washington.

Fleischer, R.C., G. Fuller, and D.B. Ledig. 1995. Genetic structure of endangered clapper rail (*Rallus longirostris*) populations in southern California. *Conservation Biology* 9: 1234-1243.

Gaggiotti, O.E., O. Lange, K. Rassman, and C. Gliddon. 1999. A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology* 8: 1513-1520.

Giesel, J.T., D. Brazeau, R. Koppelman, and D. Shriver. 1997. Ring-necked pheasant population structure. *Journal of Wildlife Management* 61: 1332-1338.

Gibbs, H.L., L.M. Tabak, and K. Hobson. 1999. Characterization of microsatellite loci for a neotropical migrant songbird, the Swainson's thrush (*Catharus ustulatus*). *Molecular Ecology* 8: 1151.

Goodman, S.J. 1997. R_{ST}Calc: A collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance. *Molecular Ecology* 6: 881-885.

Goudet, J. 1995. FSTAT version 1.2: A computer program to calculate F-statistics. *Journal of Heredity* 86: 485-486.

Greenberg, R., R.J. Cordero, S. Droege, and R.C. Fleischer. 1998. Morphological adaptation with no mitochondrial DNA differentiation in the coastal plain swamp sparrow. *Auk* 115: 706-712.

Greenwood, P.J. and P.H. Harvey. 1982. Natal and breeding dispersal in birds. *Ann. Rev. Ecol. Syst.* 13: 1-21.

Grimm, E.C., G.L. Jacobson, Jr., W.A. Watts, B.C.S. Hansen, and K.A. Maasch. 1993. A 50 000-year record of climatic oscillations from Florida and its temporal correlation with the Heinrich events. *Science* 261: 198-200.

Hoelzel, A.R. 1993. Evolution by DNA turnover in the control region of vertebrate mitochondrial DNA. *Current Opinion in Genetics and Development* 3: 891-895.

Jarne, P. and P.J.L. Lagoda. 1996. Microsatellites, from molecules to populations and back. *TREE* 11: 424-429.

Jin, L. and M. Nei. 1990. Limitations of the evolutionary parsimony method of

phylogenetic analysis. *Mol. Biol. Evol.* 7: 82-102.

Kimmel, M., R. Chakraborty, J.P. King, M. Bamshad, W.S. Watkins, and L.B. Jorde. 1998. Signatures of population expansion in microsatellite repeat data. *Genetics* 148: 1921-1930.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16: 111-120.

Kumar, S., K. Tamura, and M. Nei. 1993. MEGA Molecular Evolutionary Genetics Analysis (Version 1.01). Institute of Molecular Evolutionary Genetics, University Park, Pennsylvania.

Luikart, G. and J.-M. Cornuet. 1998. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conservation Biology*. 12: 228-237.

Luikart, G. and P.R. England. 1999. Statistical analysis of microsatellite DNA Data. *TREE* 14: 253-256.

Marshall, H.D. and A.J. Baker. 1997. Structural conservation and variation in the

mitochondrial control region of fringilline finches (*Fringilla* spp.) and the greenfinch (*Carduelis chloris*). *Mol. Biol. Evol.* 14: 173-184.

McDonald, D.B., W.K. Potts, J.W. Fitzpatrick, and G.E. Woolfenden. 1999. Contrasting genetic structures in sister species of North American scrub-jays. *Proc. R. Soc. Lond.* B. 266: 1117-1125.

Mengel, R.M. 1964. The probable history of species formation in some northern wood warblers (Parulidae). *Living Bird* 3: 9-43.

Messier, W., S.-H. Li, and C.-B. Stewart. 1996. The birth of microsatellites. *Nature* 381: 483.

Milot, E., H.L. Gibbs, and K.A. Hobson. 2000. Phylogeography and genetic structure of northern populations of the yellow warbler (*Dendroica petechia*). *Molecular Ecology* 9: 667-681.

Moore, W.S. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution* 49: 718-726.

Moritz, C. 1994. Applications of mitochondrial DNA analysis in conservation: a critical review. *Molecular Ecology* 3: 401-411.

Mundy, N.I., C.S. Winchell, and D.S. Woodruff. 1996. Tandem repeats and heteroplasmy in the mitochondrial DNA control region of the loggerhead shrike (*Lanius ludovicianus*). *Journal of Heredity* 87: 21-26.

Nei, M., T. Maruyama, and R. Chakraborty. 1975. The bottleneck effect and genetic variability in populations. *Evolution* 29: 1-10.

Nei, M., F. Tajima, and Y. Tateno. 1983. Accuracy of phylogenetic trees from molecular data. *Journal of Molecular Evolution* 19: 153-170.

Neilsen, R. and P.J. Palsbøll. 1999. Single-locus tests of microsatellite evolution: multi-step mutations and constraints on allele size. *Mol. Phyl. Evol.* 11: 477-484.

Oyler-McCance, S., N.W. Kahn, K.P. Burnham, C.E. Braun, and T.W. Quinn. 1999. A population genetic comparison of large- and small-bodied sage grouse in Colorado using microsatellite and mitochondrial DNA markers. *Molecular Ecology* 8: 1457-1465.

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

Prosser, M.R., H.L. Gibbs, and P.J. Weatherhead. 1999. Microgeographic population genetic structure in the northern water snake, *Nerodia sipedon sipedon* detected using microsatellite DNA loci. *Molecular Ecology* 8: 329-333.

Quinn, T.W. 1992. The genetic legacy of Mother Goose - phylogeographic patterns of lesser snow goose *Chen caerulescens caerulescens* maternal lineages. *Molecular Ecology* 1: 105-117.

Quinn, T.W. 1997. Molecular Evolution of Mitochodrial Genome. *In*: "Avian Molecular Evolution and Systematics" (D.P. Mindell, Ed.), pp. 4-28. Academic Press.

Raymond, M. and F. Rousset. 1995. GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86: 248-249.

Rice, W.R. 1989. Analyzing tables of statistical tests. *Evolution* 43: 223-225.

Ridgely, R.S. and G. Tudor. 1989. The Birds of South America: The Oscine Passeriformes. University of Texas Press. p. 465.

Rising, J.D. and J.C. Avise. 1993. Application of genealogical-concordance principles to the taxomony and evolutionary history of the sharp-tailed sparrow (*Ammodramus caudacutus*). *Auk* 110: 844-856.

Ruzzante, D.E. 1998. A comparison of several measures of genetic distance and population structure with microsatellite data: bias and sampling variance. *Can. J. Fish. Aquat. Sci.* 55: 1-14.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular Cloning. Cold Springs Harbor Laboratory, Cold Springs Harbor, New York.

Schneider, S. J.-M. Kueffer, D. Roessli, and L. Excoffier. 1996. Arlequin: A software package for population Genetics Genetics and Biometry Lab, Department of Anthropology, University of Geneva, Geneva, Switzerland.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139: 457-462.

Slatkin, M. and W.P. Maddison. 1989. A cladistic measure of gene flow inferred from the phylogenices of alleles. *Genetics* 123: 603-613.

Sorenson, M.D. and T.W. Quinn. 1998. Numts: A challenge for avian systematics and population biology. *Auk* 115: 214-221.

Takezaki, N. and M. Nei. 1996. Genetic distances and reconstruction of phylogenetic

trees from microsatellite DNA. *Genetics* 144: 389-399.

Taberlet, P., A. Meyer, and J. Bouvet. 1992. Unusual mitochondrial DNA polymorphism in two local populations of blue tit *Parus caeruleus*. *Molecular Ecology* 1: 27-36.

Vickery, P.D. 1996. Grasshopper Sparrow (*Ammodramus savannarum*). *In*: The Birds of North America, no 239 (A. Poole and F. Gills, Eds.). Academy of Natural Sciences, Philadelphia, and American Ornithologists' Union, Washington, D.C.

Watts, W.A. and B.C.S. Hansen. 1994. Pre-Holocene and Holocene pollen records of vegetation history from the Florida peninsula and their climatic implications. *Paleogoegraphy, Paleoclimatology, and Paleoecology* 109: 163-176.

Weber, J.L. and C. Wong. 1993. Mutation of human short random repeats. *Hum. Mol. Gen.* 2: 1123-1128.

Weir, B.S. and C.C. Cockerham. 1984. Estimating *F*-Statistics for the analysis of population structure. *Evolution* 38: 1358-1370.

Whitlock, M.C. and D.E. McCauley. 1999. Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$. *Heredity* 82: 117-125.

Yang, Z. 1996. Among-site rate variation and its impact on phylogenetic analyses. *TREE* 11: 367-371.

Zink, R.M. 1994. The geography of mitochondrial DNA variation, population structure, hybridization, and species limits in the fox sparrow. *Evolution* 48: 96-111.

Zink, R.M. 1996. Comparative phylogeography in North American birds. *Evolution* 50: 308-317.

Zink, R.M. 1997. Phylogeographic studies of North American birds. *In*: " Avian Molecular Evolution and Systematics." (D.P. Mindell, Ed.) pp 301-324. Academic Press.

Table 1. Grasshopper Sparrow microsatellite primer sequences, repeat motif, size of cloned fragment, and annealing temperatures. F denotes the forward primer and R denotes the reverse primer.

Primer	Primer sequences (5' – 3')	Repeat motif	Size (bp)	T_m (°C)
Asμ09	F CTT TGA TTA CAG AAA TAT GTC TTC T	$(CA)_{24}$	153	55
	R GAA AGA GGC ATG CTC GTA T			
Asμ15	F AAT AGA TTC AGG TGC TTT TTC	$(TG)_9$	135	53
	R TAG CAC ATG TTG GTT TTT G			
Asμ18	F ACA CAG AGA GAC ACA AAT TCA T	$(AC)_7TC(AC)_9$	132	53
	R AAA TGC TAC TGA GGT AAA GTC C			

Table 2. Summary of the mtDNA variation per population including the number of individuals genotyped (N), number of haplotypes observed (No. haplotypes), number of polymorphic sites, average number of pair-wise differences among haplotypes, and nucleotide diversity per site (π) per population.

	Florida	Georgia	Ohio	Massachusetts	Arizona
N	98	19	17	21	16
No. Haplotypes	37	12	12	18	9
Polymorphic sites	34	23	21	23	15
Pair-wise differences	8.332	11.859	7.376	8.861	5.531
π	0.009	0.013	0.008	0.010	0.006

Table 3. Summary of the microsatellite DNA variation per locus and per population including the number of individuals genotyped (N), number of alleles detected (No. alleles), and the frequency of the most common allele.

	Locus					
	*Dp*μ16	*Ma*μ23	*Ca*μ02	*As*μ09	*As*μ15	*As*μ18
Florida						
N	105	105	105	101	105	105
No. alleles	11	6	11	12	18	21
Frequency	0.59	0.68	0.21	0.44	0.21	0.24
Georgia						
N	19	19	17	19	19	19
No. alleles	6	5	9	11	12	14
Frequency	0.55	0.82	0.26	0.29	0.24	0.18
Ohio						
N	34	34	31	34	34	34
No. alleles	6	5	11	15	16	18
Frequency	0.47	0.79	0.44	0.25	0.24	0.22
Massachusetts						
N	36	39	39	38	37	39
No. alleles	10	6	10	13	15	14
Frequency	0.49	0.68	0.35	0.33	0.24	0.28

Table 3. Cont'd.

	Locus					
	*Dp*μ16	*Ma*μ23	*Ca*μ02	*As*μ09	*As*μ15	*As*μ18
Arizona						
N	17	15	14	17	17	17
No. alleles	5	4	7	7	13	13
Frequency	0.56	0.80	0.43	0.35	0.18	0.18

Table 4. Observed (H_o) and expected (H_e) heterozygosity for each of the six microsatellite loci per population as determined using GENEPOP. Loci with a significant heterozygote deficiency are indicated with an asterisk.

		Locus					
Population		*Dp*μ16	*Ma*μ23	*Ca*μ02	*As*μ09	*As*μ15	*As*μ18
Florida	H_o	0.60*	0.47	0.75*	0.73	0.83*	0.89
	H_e	0.61	0.50	0.79	0.74	0.90	0.90
Georgia	H_o	0.80	0.37	0.77	0.95	0.90	0.84
	H_e	0.63	0.32	0.82	0.84	0.90	0.84
Ohio	H_o	0.65	0.35	0.61	0.94	0.91	0.94*
	H_e	0.71	0.35	0.68	0.88	0.88	0.88
Massachusetts	H_o	0.64*	0.62	0.82	0.76	0.95	0.90
	H_e	0.69	0.49	0.80	0.84	0.89	0.85
Arizona	H_o	0.47*	0.40	0.64	0.59	0.88	0.94
	H_e	0.59	0.33	0.71	0.77	0.88	0.94

Table 5. F_{ST} values between population pairs based on distance-weighted mtDNA sequence data. * $p \leq 0.008$ † significant prior to adjustment for multiple comparisons.

	Florida	Georgia	Ohio	Massachusetts
Georgia	0.119*			
Ohio	0.114*	-0.032		
Massachusetts	0.042†	-0.021	0.008	
Arizona	0.119*	-0.014	-0.027	0.021

Table 6. Theta (above diagonal) and *rho* (below diagonal) values for pair-wise comparisons among the five populations. * $p \leq 0.008$ following adjustment for multiple comparisons; † values significant prior to adjustment.

	Florida	Georgia	Ohio	Massachusetts	Arizona
Florida	-	0.014*	0.020*	0.017*	0.026*
Georgia	0.007	-	0.005	0.011†	0.009
Ohio	0.005	-0.014	-	0.007†	0.006
Massachusetts	0.004	0.021	0.009	-	0.012†
Arizona	0.019	-0.007	0.017	0.031†	-

Table 7. Results of the recent bottleneck test. The two right-hand columns indicate the number of loci exhibiting heterozygosity excess or deficiency assuming a 9:1 ratio two-parameter model. * $p \leq 0.05$ for populations having significantly *fewer* number of loci exhibiting heterozygosity excess than expected given the number of alleles per locus. If a population has experienced a recent population bottleneck, there would be a significantly *greater* number of loci exhibiting heterozygosity excess than expected.

	Deficiency	Excess
Florida	6	0*
Georgia	5	1*
Ohio	6	0*
Massachusetts	6	0*
Arizona	4	2

Table 8. Genetic imbalance index values for individual Grasshopper Sparrow populations. Variance in allele size (θ_v), expected homozygosity (θ_{po}), and genetic imbalance index (β). All values are averaged across six microsatellite loci per population.

	θ_v	θ_{po}	$\beta \pm$ SE	95% CI
Florida	1.000	0.260	0.419 ± 0.197	0.032 – 0.0806
Georgia	0.991	0.261	0.470 ± 0.300	-0.117 – 1.057
Ohio	0.923	0.255	0.498 ± 0.344	-0.177 – 1.173
Massachusetts	1.062	0.235	0.465 ± 0.308	-0.140 – 1.069
Arizona	0.940	0.284	0.653 ± 0.468	-0.264 – 1.569
Mean across populations	0.983	0.260	0.501	

Figure 1. Locations of sampled populations (dots) and subspecies boundaries in the United States. *A. s. ammolegus* is represented by the Arizona population (n = 17); *A. s. pratensis* is represented by three populations, Massachusetts[1] (n = 39), Ohio[2] (n = 34), and Georgia[3] (n = 19); and *A. s. floridanus* is represented by the Florida population (n = 105).

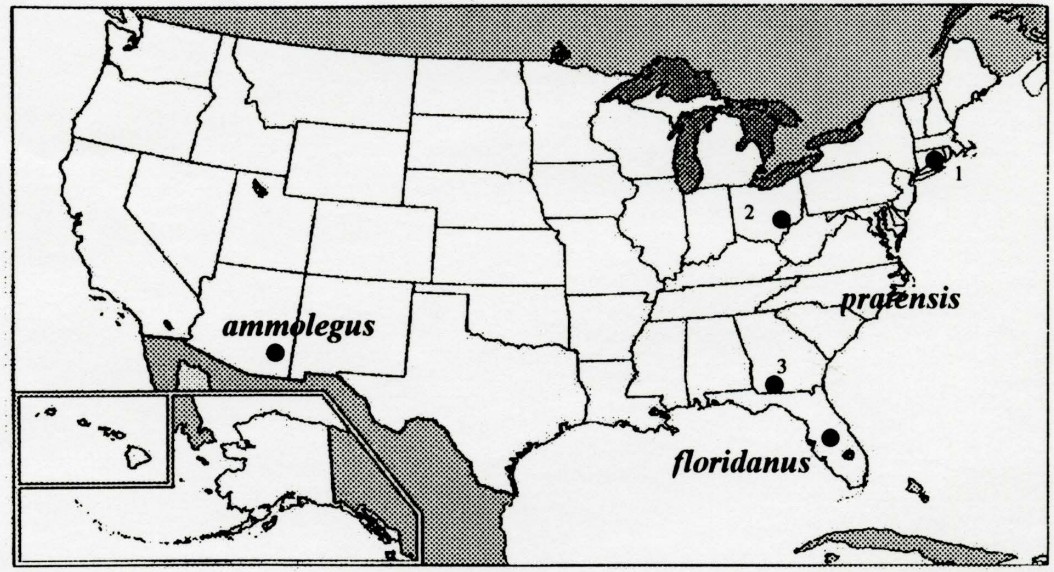

1
2
pratensis
ammolegus
3
floridanus

Figure 2. Relative location of primers used to amplify and sequence the Grasshopper Sparrow mtDNA control region and their amplification product sizes. Arrows beside primer names indicate the 5' to 3' orientation of the primer.

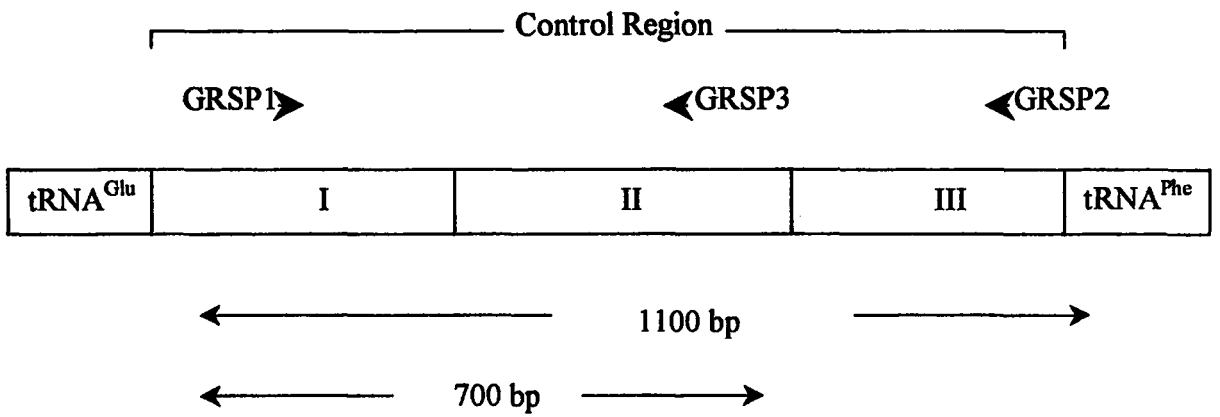

Control Region
GRSP1
GRSP3
GRSP2
tRNA^Glu
I
II
III
tRNA^Phe
1100 bp
700 bp

Figure 3. Variable sites among the 58 Grasshopper Sparrow haplotypes. All variation is with respect to GRSP1. Dots indicate homologous sites.

```
                 1111111111122222222333344445666677777777788
               467902345899999123337791799045692289013678890 2
               269613103905689806898932278634762834892611518 3
GRSP1          tgacttaccccactaatccggggaaaaaatctcaccgcacaaaacc
GRSP2          .a.t.c.....................................
GRSP3          .....cg........g..a........gt..............
GRSP4          .....c.........g..a........................
GRSP5          .a...c.........g..a........................
GRSP6          .a...c.....................................
GRSP7          .a...c....................c.............g.g...
GRSP8          .a...c....................g.....................
GRSP9          .a.t.c...t.....................................
GRSP10         .....c....tt...g..a............a..............
GRSP11         .....c.........g..a..........................t.
GRSP12         .....c....t..c.g..a........................
GRSP13         ....cc.........g..a...................g.......
GRSP14         .a...cg.....t..g...a....t.................-..
GRSP15         .....c.........g..a...................g.......
GRSP16         .a.t.c..............................t.........
GRSP17         .a.................................t.......g...
GRSP18         .....c....t..c.g..a...................g.......
GRSP19         .at..c.....................................
GRSP20         ..........t............................t.......
GRSP21         ....................................t.........
GRSP22         .a...c.........c.........t.....................
GRSP23         .a...c..t..............................a.........
GRSP24         ....c.....................................t.....
GRSP25         .....c....t................................
GRSP26         .....c.........g..a...a....................
GRSP27         c....c.........cg..a.......................
GRSP28         c....c.........g..a........................
GRSP29         ....c.....................................t-.....
GRSP30         .a....t...t....g........t...................
GRSP32         .....c.........g..a..............t.........
GRSP34         .....c.........g..a............a...........
GRSP35         .....c.....................................
GRSP37         .a...c.........c............................
GRSP38         .a...c.....................................g...
GRSP39         ...t.c.............................ta.........
GRSP40         .....c....tt...g..a............a............a
GRSP41         .a.t.c...t............................t.........
GRSP42         .....c.........g..a.a......................
GRSP43         ....................................t.....
GRSP44         .a...c.....................................a
GRSP45         .a.t.c..........c..................t.........
GRSP46         .a...c.........c.........t..........t.........
GRSP47         .a...c.........c.........t...t...............
GRSP48         ...t.c.....................c.............g.g...
GRSP49         ..........................................t.
GRSP50         .a.t.c...t.......t.................t.........
GRSP51         .a...c..t..................................
GRSP52         .a...c.....................c...........g.g.g...
GRSP53         ...t.c.t..............................ta.........
GRSP54         .....c....tt...g..a..t.........a............
GRSP55         .a...c................................a.........
GRSP56         .....c.........g..a.....................g.g...
GRSP57         .....c.........g..a.........................a
```

```
                111111111112222222233334444566667777777788
          467902345899999123337791799045692289013678890 2
          269613103905689806898932278634762834892611518 3
GRSP1     tgacttaccccactaatccggggaaaaaatctcaccgcacaaaacc
GRSP59    .....c....t..c.g..a....g.....ct..............
GRSP61    ....c..........c.............................
Jam005    ....c....t.....................a.-.....t......
```

Figure 4. Minimum-spanning tree of the 58 Grasshopper Sparrow haplotypes, the Jamaican Grasshopper Sparrow (*A.s.s.*), and two outgroup species [the Yellow-Browed Sparrow (*A.a.*) and the Grassland Sparrow (*A.h.*)]. Each circle represents one haplotype indicated by the number in bold. The top-left legend indicates the population(s) in which each haplotype is found. Numbers in boxes along connecting lines indicate the number of substitutions separating the two haplotypes; lines lacking a number are separated by one substitution.

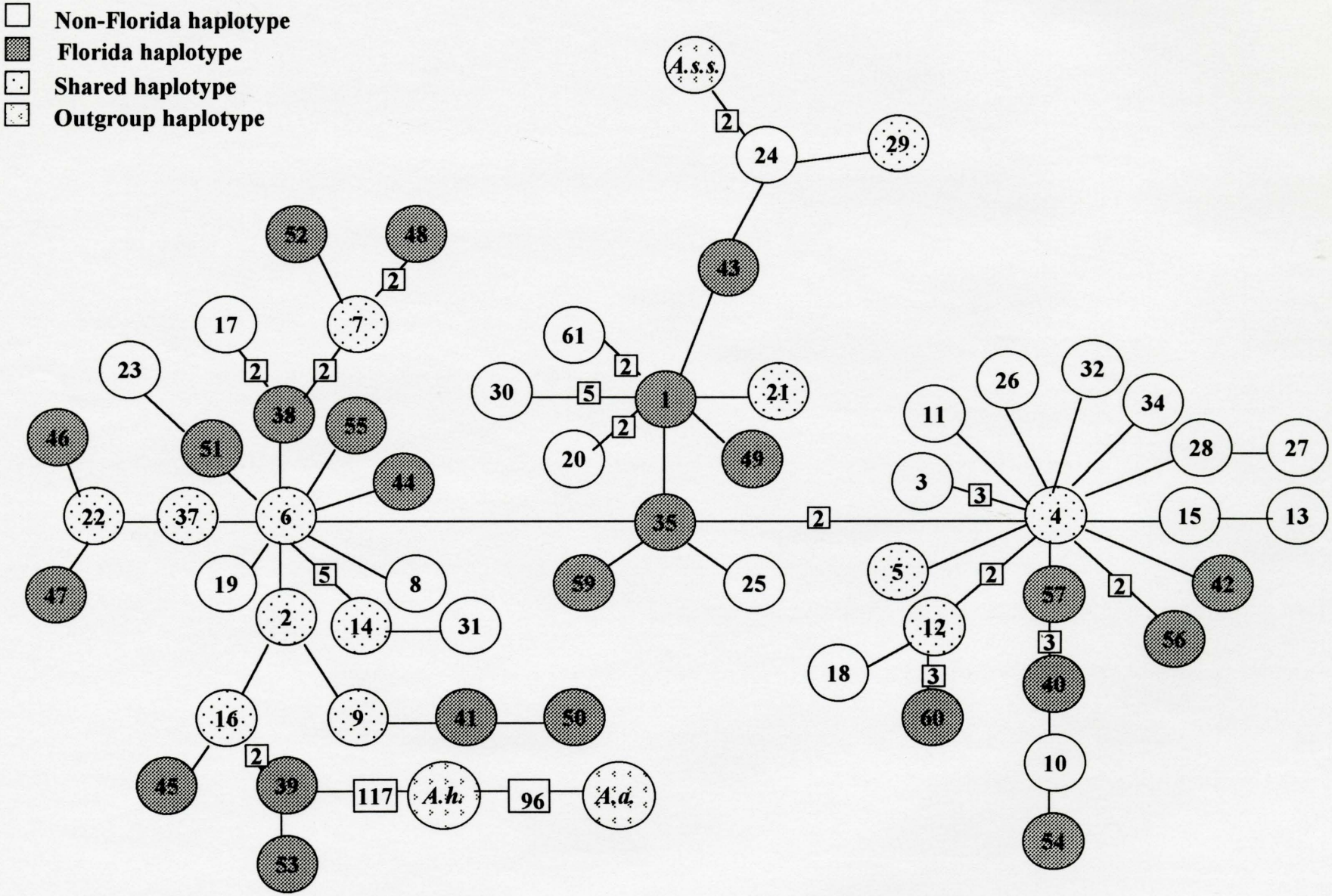

Non-Florida haplotype
Florida haplotype
Shared haplotype
Outgroup haplotype
A.s.s.
A.h.
A.a.
1
2
4
5
6
7
8
9
10
11
12
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
34
35
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
59
60
61
2
3
5
117
96

Figure 5. Consensus neighbor-joining tree of the 58 Grasshopper Sparrow haplotypes, a Jamaican sample (*A. s. savannarum*), and two outgroup species (*A. humeralis* and *A. aurifrons*). The tree was rooted with the *A. aurifrons* mtDNA sequence. Bootstrap values greater than 50% are indicated at the node. Numbers at branch tips indicate the haplotype identity. The population(s) in which it was found are shown in parentheses. The scale bar at the bottom right is proportional to branch length, measured as the number of DNA substitutions per site. Due to size constraints, the broken lines connecting the two outgroups to the ingroup clade are not proportional to branch length. The branch length separating *A. aurifrons* from the nearest ingroup node is 0.11.

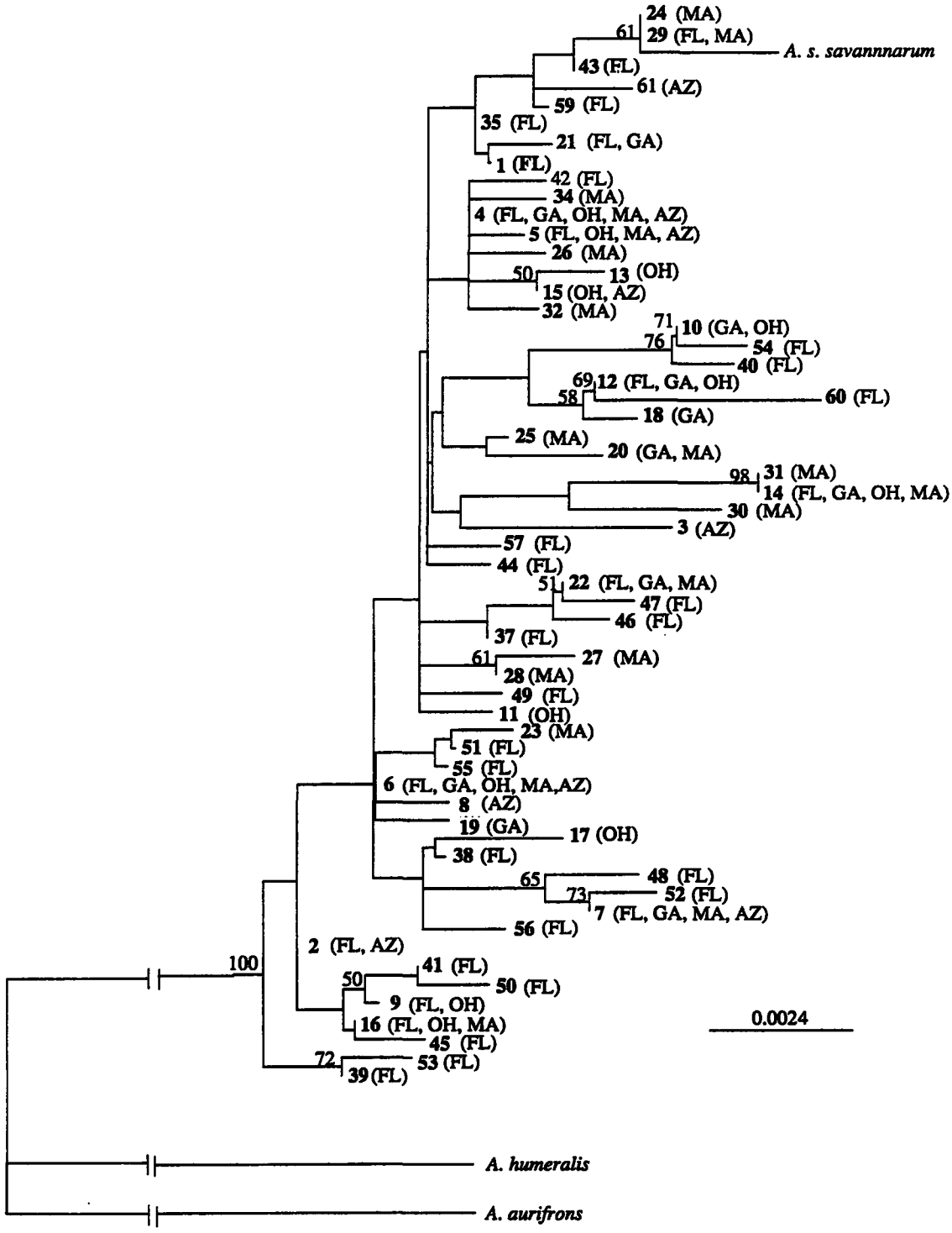
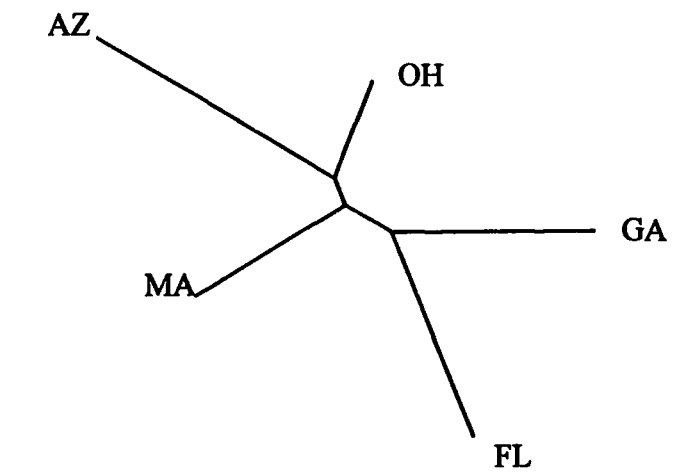
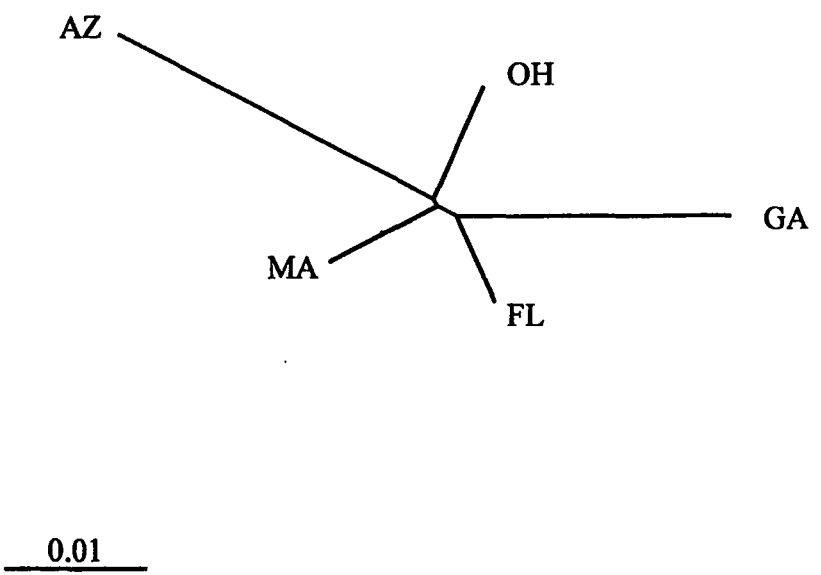

24 (MA)
61
29 (FL, MA)
A. s. savannnarum
43 (FL)
61 (AZ)
59 (FL)
35 (FL)
21 (FL, GA)
1 (FL)
42 (FL)
34 (MA)
4 (FL, GA, OH, MA, AZ)
5 (FL, OH, MA, AZ)
26 (MA)
50
13 (OH)
15 (OH, AZ)
32 (MA)
71
10 (GA, OH)
76
54 (FL)
40 (FL)
69
12 (FL, GA, OH)
58
60 (FL)
18 (GA)
25 (MA)
20 (GA, MA)
98
31 (MA)
14 (FL, GA, OH, MA)
30 (MA)
3 (AZ)
57 (FL)
44 (FL)
51
22 (FL, GA, MA)
47 (FL)
46 (FL)
37 (FL)
61
27 (MA)
28 (MA)
49 (FL)
11 (OH)
23 (MA)
51 (FL)
55 (FL)
6 (FL, GA, OH, MA,AZ)
8 (AZ)
19 (GA)
17 (OH)
38 (FL)
65
48 (FL)
73
52 (FL)
7 (FL, GA, MA, AZ)
56 (FL)
2 (FL, AZ)
100
41 (FL)
50
50 (FL)
9 (FL, OH)
16 (FL, OH, MA)
45 (FL)
72
53 (FL)
39 (FL)
0.0024
A. humeralis
A. aurifrons

Figure 6. Unrooted neighbor-joining trees depicting evolutionary relationships of the five populations using (a) Nei's genetic distance and (b) Cavalli-Sforza and Edwards chord distance. The scale bar represents branch length. Populations are indicated at the branch tips by their two-character state abbreviation: GA (Georgia); MA (Massachusetts); OH (Ohio); FL (Florida); and AZ (Arizona).



(a)



(b)

Appendix 1. Sequence for Grasshopper Sparrow haplotype (GRSP) 1. Refer to Figure 3 for substitutions in haplotypes GRSP2 to 61. Roman numerals indicate the 5' end of the respective control region domain. The numbers on the right-hand side of the figure indicate the base-pair site. The line separating the end of domain II and beginning of domain III indicates an approximately 200 bp region which could not be sequenced because of an inverted repeat in the third domain. This region has not be included in the length of the sequence.

```
5' Itgataatccatagtatatgtaatgctcttccattagaaacctaaacattat  51

ctccaaaacagatggtatttggacacaatatccaccaggcacatccttgtttcag 106

ggaccatagagcccaatctctccacctacgaccagatgcaagcgtcacccaaaca 161

cccaggaacttatctgttatgcttaccctccacctagtgaacgaggaatgtccca 216

gtacacctttgcattctcccggtctactgaattcgcccacctcctaggtaatgtt 271

cacggIIccaacagccttcaggaactcccaagccagaggacaaggttatctattg 324

atcgcgcttctcacgagaaccgagctactcaacgtatgagtgatatcgtttattg 379

tccttgagcccataaatcgcctaatcttgctctttggcgctagtggttgtaactt 434

caggaacatacctggttgactccggatcccttgctcttactgatacaagtggtcg 489

gtttgaatagtcctccctactctcattttcccggcataccgacctcctacacttg 544

ttttttttctctctcctttcaataagcccctcaagtgcagagcaggtgttatctt 599

cctcttgacatgtccatcacatgaccgtcgagcatatgaatcccctaccacgcag 654

aatgtcatggtttgatg |---------- 200 bp ----------| 671

IIIaccaaacaacaaacccacattttcctacattgtctagatcatctatcgtcaa 723

ttcatcatcaattaaccttcctctacattttctgctactaaaaacaaaaactaat 778

catcattatttttatcttttacatcatacaaattagccccaaaattactgcccct 833

tcaaaaaccaaacaaaaacacaaaccatgacaataaacaatcaa 3'          877
```

Appendix 2. Distribution of mtDNA haplotypes among the five populations. Refer to Figure 3 for details as to how haplotypes differ.

Haplotype	Florida	Georgia	Ohio	Massachusetts	Arizona	Total
GRSP1	3					3
GRSP2	1				1	2
GRSP3					2	2
GRSP4	8	2	2	1	6	19
GRSP5	1		2	1	1	5
GRSP6	11	1	3	2	2	19
GRSP7	9	1		1	1	12
GRSP8					1	1
GRSP9	2		1			3
GRSP10		1	2			3
GRSP11			1			1
GRSP12	1	1	1			3
GRSP13			1			1
GRSP14	3	4	1	1		9
GRSP15			1		1	2
GRSP16	2		1	1		4
GRSP17			1			1
GRSP18		3				3
GRSP19		1				1

Appendix 2. Cont'd.

Haplotype	Florida	Georgia	Ohio	Massachusetts	Arizona	Total
GRSP 20		1		1		2
GRSP 21	3	2				5
GRSP 22	11	1		3		15
GRSP 23		1				1
GRSP 24				1		1
GRSP 25				1		1
GRSP 26				1		1
GRSP 27				1		1
GRSP 28				1		1
GRSP 29	5			1		6
GRSP 30				1		1
GRSP 31				1		1
GRSP 32				1		1
GRSP 34				1		1
GRSP 35	2					2
GRSP 37	1					1
GRSP 38	3					3
GRSP 39	1					1
GRSP 40	1					1
GRSP 41	7					7

Appendix 2. Cont'd.

Haplotype	Florida	Georgia	Ohio	Massachusetts	Arizona	Total
GRSP 42	1					1
GRSP 43	1					1
GRSP 44	2					2
GRSP 45	1					1
GRSP 46	2					2
GRSP 47	3					3
GRSP 48	1					1
GRSP 49	1					1
GRSP 50	1					1
GRSP 51	1					1
GRSP 52	1					1
GRSP 53	1					1
GRSP 54	2					2
GRSP 55	1					1
GRSP 56	1					1
GRSP 57	1					1
GRSP 59	1					1
GRSP60	1					1
GRSP61					1	1

Appendix 3. Variance in allele size per locus per population for original *rho* calculations, removing alleles at a frequency of less than 5%, and using a Florida sample size of 25.

	Original	Removal of alleles with a frequency < 5%	$n = 25$ for Florida
Florida			
Dpμ16	1.046	0.513	1.248
Maμ23	0.949	0.991	0.857
Caμ02	1.056	1.124	1.312
Asμ09	0.872	0.970	0.388
Asμ15	0.969	1.025	1.043
Asμ18	1.108	0.999	1.025
Georgia			
Dpμ16	0.944	1.292	0.931
Maμ23	0.597	0.716	0.585
Caμ02	0.959	1.094	0.937
Asμ09	1.560	1.187	1.559
Asμ15	0.740	0.432	0.708
Asμ18	1.147	1.041	1.274

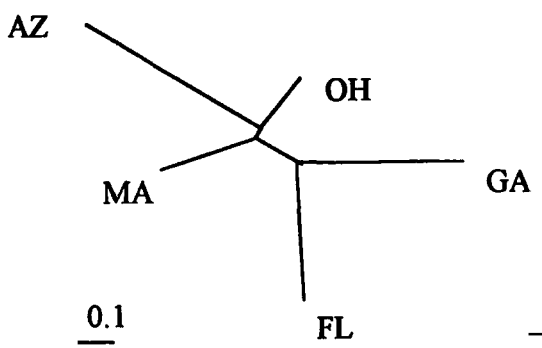
Appendix 3. Cont'd.

	Original	Removal of alleles with a frequency < 5%	n = 25 for Florida
Ohio			
Dpμ16	0.953	0.424	0.941
Maμ23	0.795	0.808	0.779
Caμ02	0.915	0.593	0.894
Asμ15	0.980	0.956	0.938
Asμ18	0.843	0.996	0.937
Massachusetts			
Dpμ16	1.107	0.386	1.092
Maμ23	1.473	1.397	1.443
Caμ02	1.009	0.782	0.986
Asμ09	0.936	0.704	0.935
Asμ15	1.127	1.119	1.079
Asμ18	0.722	0.882	0.803

Appendix 3. Cont'd.

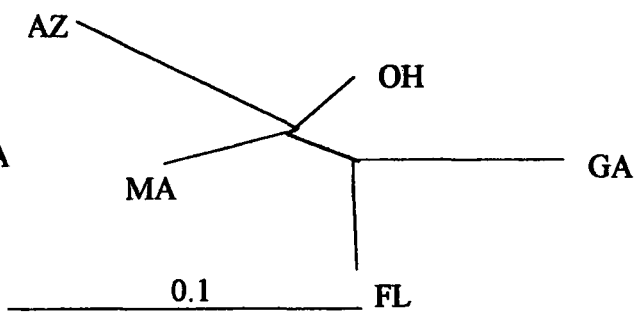
	Original	Removal of alleles with a frequency < 5%	n = 25 for Florida
Arizona			
Dpμ16	0.536	0.036	0.529
Maμ23	1.051	0.940	1.029
Caμ02	0.925	0.976	0.905
Asμ09	0.934	0.787	0.933
Asμ15	1.204	0.991	1.152
Asμ18	0.991	0.987	1.101

Appendix 4. Psuedo-jackknife neighbor-joining trees inferred from microsatellite data using Nei's genetic distance (A-F) and the Cavalli-Sforza and Edwards chord distance (G-L). Branch lengths are proportional to the scale at the bottom of each tree. Locus names below individual trees indicate the locus removed for that particular jackknife simulation.



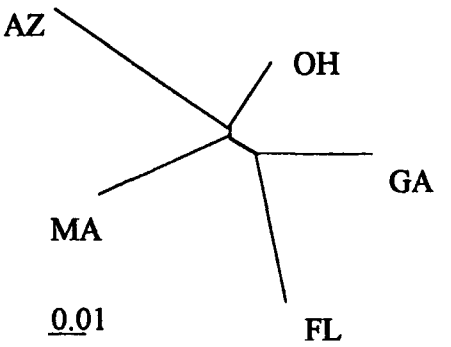
*Dp*μ16

A



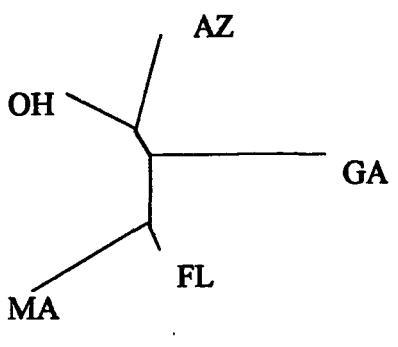
*Ma*μ23

B



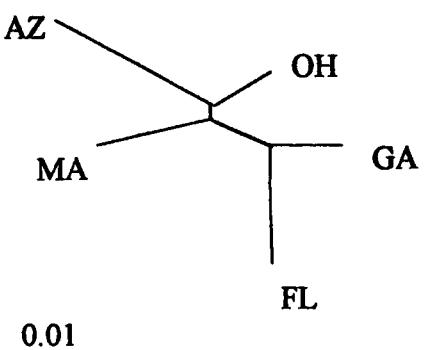
*Ca*μ02

C



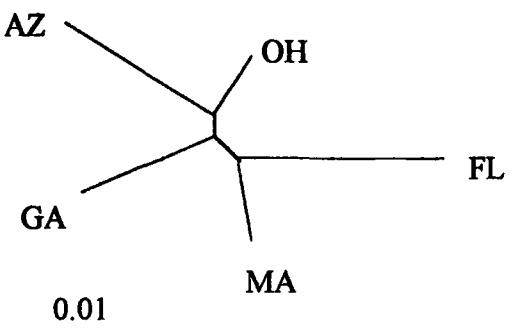
*As*μ09

D



*As*μ15

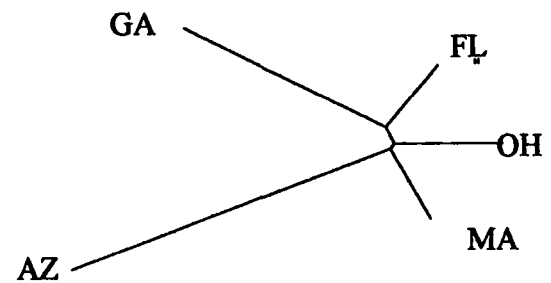
E



*As*μ18

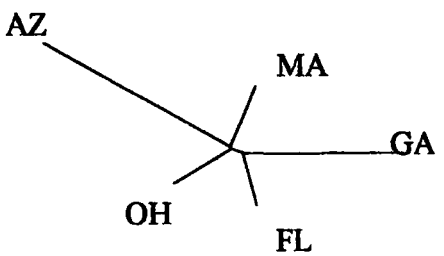
F

Appendix 4. Cont'd.



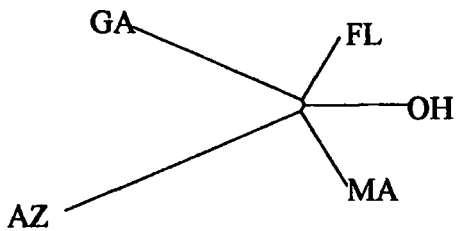
*Dp*μ16

G



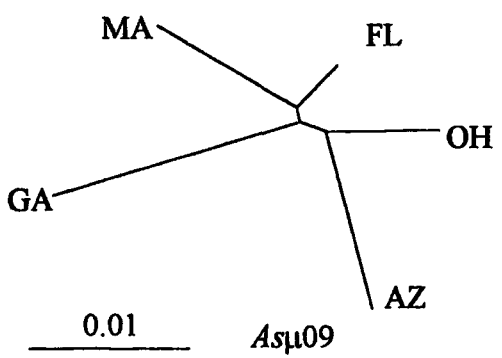
*Ma*μ23

H



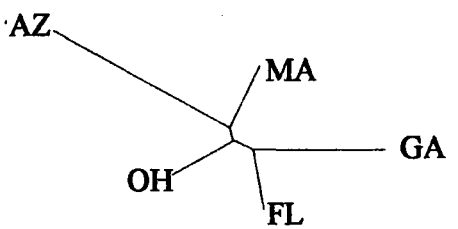
*Ca*μ02

I



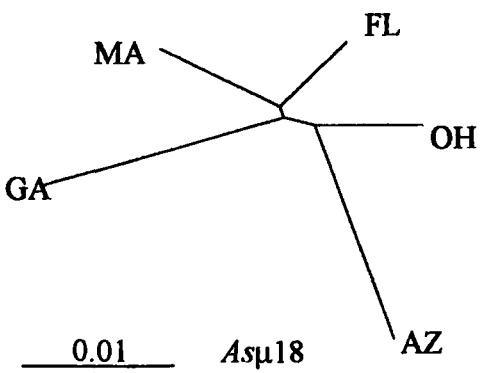
*As*μ09

J



*As*μ15

K



*As*μ18

L

Chapter 3 - Genetic Estimates of Dispersal Among Sub-Populations of the Florida Grasshopper Sparrow (*Ammodramus savannarum floridanus*)

ABSTRACT

Accurate measures of dispersal among heterogeneously distributed bird populations are necessary for assessing the demographic fate of individual sub-populations. In the past, these measures were based entirely on direct observations of banded birds using either recapture methods or radio-telemetry. Recently, genetic assignment tests have provided an alternative indirect method of measuring dispersal. These tests assign individuals to particular populations based on the probability of the individual's genotype originating from each of a set of potential source populations. Here I use such tests in combination with DNA-based genetic data [mitochondrial DNA (mtDNA) control region sequence and six microsatellite DNA loci] to assess population structure and levels of movement among four sub-populations of the endangered Florida Grasshopper Sparrow (*Ammodramus savannarum floridanus*). I observed no significant genetic differentiation among the four sub-populations based on mtDNA variation (F_{ST} = 0.011 to 0.026), or microsatellite variation (θ = -0.001 to 0.012; ρ = 0.003 to 0.033). Despite this low level of differentiation, I was able to use the genetic assignment test program IMMANC to detect moderate levels of movement among the sub-populations. All sub-populations both donated and received immigrants, but varied in the percentage of immigrants found in each sub-population (4 to 15%). However, each sub-population also had a ratio of immigrants to emigrants near one (range: 0.5 to 1.2). This suggests

that the dynamics of these sub-populations best fit a balanced dispersal, as opposed to a source-sink, metapopulation model. From a conservation perspective, the lack of significant genetic structure and moderate levels of movement among the sub-populations both imply that the *floridanus* subspecies should be managed as a single management unit on a subspecies level versus multiple units on a sub-population level.

INTRODUCTION

Natural populations of organisms tend to be heterogeneously distributed in space. An increasingly common conceptual way of describing such spatial distributions is as metapopulations. A metapopulation is defined as a group of geographically disjunct sub-populations that are linked by dispersal (Hanski and Simberloff 1997). Population biologists originally assumed that dispersal between sub-populations was random, and that extinction and colonization of the sub-populations were due to stochastic processes alone (Levins 1970). However, two current models of metapopulation dynamics offer more realistic views of these processes: the source-sink model (Pulliam 1988); and the balanced dispersal model (Diffendorfer *et al.* 1995). Both of these models acknowledge that sub-populations of a metapopulation vary in the level of dispersal among sub-populations, but differ in terms of the ratio of immigrants to emigrants.

In order to persist, a population must at least maintain a birth rate equivalent to the death rate. Otherwise, the population will go extinct at a rate proportional to the ratio of deaths to births per unit time. In a source-sink metapopulation model, breeding success differs among sub-populations (Pulliam 1988). Sub-populations in which natality exceeds mortality are deemed sources. In contrast, sub-populations in which mortality exceeds natality are referred to as sinks. Essentially, a source-sink structured metapopulation's demographic dynamics are driven by emigration from the source and immigration into the sink (Dias 1996) and are believed to be the result of intraspecific competition or density-dependent dispersal (Weatherhead and Forbes 1988). Consequently, sink sub-populations are expected to go extinct without continued

immigration from the more fecund source sub-populations to counteract the high net mortality (Pulliam 1988).

In contrast to the source-sink model, Diffendorfer *et al.* (1995) proposed a balanced dispersal model in which there is an equilibrium state between immigration into and emigration out of each sub-population within a metapopulation. This model was suggested based on a ratio of immigrants to emigrants near one, which was observed among habitat patches occupied by three rodent species regardless of their size. Similarly, Doncaster *et al.* (1997) observed near unity ratios of immigration to emigration among sub-populations of Collared Flycatchers (*Ficedula albicollis*). In this study, the observed balanced dispersal appeared to be independent of sub-population density, and reproductive success was similar among the sub-populations.

Essentially, the main factors defining the two metapopulation models are (i) the ratio of immigrants to emigrants per sub-population and (ii) the annual ratio of births to deaths per sub-population. In a source-sink structured metapopulation, a source exhibits a ratio less than one and net natality, while a sink exhibits a ratio greater than one and net mortality. In contrast, sub-populations of a balanced dispersal metapopulation all have an immigrant to emigrant ratio near one and exhibit no differences in reproductive success. Determination of which model best fits a metapopulation is crucial when considering conservation issues because the viability of source-sink structured metapopulations is dependent on the survival of the sources. Thus, efforts should be focused on managing source sub-populations where metapopulations are known to be source-sink structured. Conversely, in a balanced dispersal metapopulation, conservation priorities should be

evenly balanced among all sub-populations because each one is equally and independently susceptible to extinction.

Bird Metapopulations

Currently, many North American avian species are suffering population declines associated with drastic habitat fragmentation and degradation. Accordingly, it may be more appropriate to describe the demographic patterns of these fragmented populations using a metapopulation approach to compensate for the effects of human-induced heterogeneous distribution. For example, Brawn and Robinson (1996) observed especially high levels of predation and cowbird parasitism associated with small (< 200 ha) woodlots in the Wood Thrush (*Hylocichla mustelina*), the Summer Tanager (*Piranga rubra*), the Scarlet Tanager (*P. olivacea*), and the Red-Eyed Vireo (*Vireo olivaceus*) populations in Illinois. Thus, Brawn and Robinson speculated that the persistence of these predated populations was likely associated with continued immigration from source populations outside the state. Likewise, Robinson *et al.* (1995) observed decreased reproductive success in seven Neotropical passerine species occupying Midwestern United States fragmented forest habitat. From these results, Robinson *et al.* concluded that the entire fragmented forest functioned as a sink habitat maintained on a regional scale by emigration from surrounding larger forest tracts, which served as sources. Computer simulations of source-sink dynamics in a hypothetical migratory bird metapopulation have shown that source-sink interactions are indeed severely affected by habitat fragmentation (Donovan *et al.* 1995). In particular, fragmentation of source habitat decreases overall metapopulation size because the source is no longer able to

provide a sufficient number of immigrants to sustain the sink(s) once a certain threshold of core habitat has been removed. If this is true, conservation efforts should first determine if a metapopulation contains sources and sinks, and if so, focus efforts on managing the source sub-populations in order to preserve diversity on a metapopulation scale.

To date, several studies suggest that avian metapopulations fit either a source-sink or a balanced dispersal model. For instance, Sæther *et al.* (1999) observed a negative population growth rate (λ) in an island population of the House Sparrow (*Passer domesticus*), suggesting that this island is a sink habitat. Matthysen (1999) detected higher numbers of post-summer immigrants in isolated versus continuous patches of European Nuthatches (*Sitta europea*) in northern Belgium. This directional immigration suggested that large forest patches are source habitats that contribute immigrants to isolated patches that serve as sinks. McGowan and Otis (1998) also reported negative population growth rates in two South Carolinian Mourning Dove (*Zenaida macroura*) sub-populations, again suggesting source-sink structure on a regional scale. In contrast, Collared Flycatchers exhibit net zero movement among sub-populations with no difference in reproductive success suggesting that a balanced dispersal model best describes the interactions between sub-populations of these birds (Doncaster *et al.* 1997).

Direct Methods for Estimating Dispersal

To determine which model best describes birds with a metapopulation structure, the magnitude and direction of dispersal among the sub-populations can be assessed directly. Usually such data has been collected using band and recapture techniques

which involve direct physical sampling of the individuals present in each sub-population. The downfall of this method is that the probability of recapture of banded individuals is low and limited by the fact that dispersers are missed if they fall outside the study area (Barrowclough 1980). The advent of radio-telemetry made it possible to track an individual without having to physically recapture it (White and Garrott 1990). Studies comparing the two direct methods have shown a minimum three-fold increase in dispersal distance when using telemetry versus band and recapture methods (reviewed in Koenig *et al.* 1996). For instance, Armitage (1991) found a 282% and 332% increase in dispersal distance of female and male Yellow-Bellied Marmots (*Marmota flaviventris*), respectively, when comparing distances derived from telemetry data to previously estimated distances from band and recapture studies. Similarly, Koenig and Mumme (1987) found a 743% and 504% increase in dispersal distance of female and male Acorn Woodpeckers (*Melanerpes formicivorus*), respectively, when comparing distances derived from telemetry data to previously estimated distances from band and recapture studies. As with band and recapture methods however, the ability to track a tagged bird is limited by the size of the sampling area, while the number of birds that can be monitored is limited by the number of distinct radio frequencies available. In addition, antennas are frequently lost or damaged rendering them useless.

Genetic Methods for Estimating Dispersal

Aside from direct methods of measuring movement within metapopulations, indirect genetic methods offer several advantages recapture and telemetry do not. Most importantly, genetic estimates only require one sampling of an individual; there is no

need to capture or detect a specimen for any purpose subsequent to the initial sampling. To indirectly estimate the level of movement between two populations from genetic data, ornithologists have traditionally inferred the number of migrants per generation from F_{ST}, a measure of population differentiation (Wright 1943). This approach relies on the general assumption that the magnitude of population differentiation is inversely proportional to the magnitude of dispersal between the two sites (Whitlock and McCauley 1999). However, this method has one major limitation in that the estimate infers rates of effective dispersal from long-term effective population sizes averaged over hundreds to thousands of years (Koenig *et al.* 1996) which makes it unacceptable for such purposes. Consequently, migration estimates derived from F_{ST} values are not necessarily accurate reflections of levels of contemporary dispersal.

Recently, a different analytical approach has been developed to measure movement at the level of the individual. Specifically, "assignment tests" use genetic data to determine the source (i.e. donor in this context) population of individual specimens (Waser and Strobeck 1998). Essentially, an individual's genotype serves as a genetic "tag" which is compared to the likelihood distribution of individual genotypes of each population to identify the most likely origin of the individual. The first such assignment test, Doh, was designed by Paetkau *et al.* (1995). This frequency method calculates the likelihood of an individual's genotype in each population based on allele frequency data. An individual is then assigned to the population in which its genotype is most likely. A disadvantage of this method is that it does not provide any information as to the significance of the assignment (Davies *et al.* 1999). As a consequence, measures of

movement could potentially be overestimated because individuals which have a higher likelihood score by random chance alone will be incorrectly interpreted as dispersers.

As an alternative, Rannala and Mountain (1997) designed a similar assignment test, IMMANC. Building on the Doh frequency method, IMMANC also calculates the likelihood of an individual's genotype in each population by multiplying the probabilities of the individual possessing each allele at each locus given the population's allele frequency distribution. However, IMMANC also tests whether the ln likelihood of the individual being from its sampled population minus the ln likelihood of the individual being from an alternative source population is significantly less than zero and corrects for differences in diversity among sub-populations (Davies *et al.* 1999) using the following formula:

$$\ln \Lambda = \ln[\Pr(X_{im} \mid n_i)] - \ln[\Pr(X_{im} \mid n_{i'})] \quad (1)$$

where ln Λ is the difference between the ln likelihood of the genotype originating in the sampled population minus the ln likelihood of the genotype originating in an alternative population, $\Pr(X_{im} \mid n_i)$ is the probability of the genotype in the sampled population and $\Pr(X_{im} \mid n_{i'})$ is the probability of the genotype in an alternative population (Equation 16 from Rannala and Mountain 1997). If the difference is significantly less than zero, the individual is then assigned to the alternative population. Converesly, if the difference is greater than zero, the individual is assigned to the sampled population.

IMMANC also provides a power estimate per comparison related to the reliability of the test to detect an individual from an alternative source population. Generally speaking, the power to detect an immigrant increases with the number of loci surveyed

(Waser and Strobeck 1998). Although conservative in requiring a significantly greater likelihood for alternative assignments, any such assignments can be regarded with high confidence due to a combination of the power and significance values calculated per comparison. Furthermore, Cornuet *et al.* (1999) recommend IMMANC as they found greater assignment accuracy using IMMANC over both frequency and genetic distance methods regardless of the number of loci assayed, the sample size, the mutation model, and the magnitude of F_{ST} values between population pairs.

The Endangered Florida Grasshopper Sparrow

The Florida Grasshopper Sparrow (*A. s. floridanus*) is an endangered subspecies endemic to central grasslands of Florida and is distinct from other subspecies in that it is non-migratory, shows subtle differences in plumage, and exhibits distinct habitat preferences (Dean *et al.* 1998; Vickery 1996; Delany and Linda 1998). Intense agricultural and urban development throughout the state of Florida during the last century has led to substantial decreases in preferred Grasshopper Sparrow habitat and an associated population decline (Vickery 1996). Presently, the entire subspecies exists as four geographically isolated sub-populations separated by 5 to 18 km at the outermost boundaries. Both the sedentary behavior and disjunct distributions of the sub-populations suggest that significant genetic differentiation and minimal movement may exist among the sub-populations. If this is true, conservation efforts will need to focus on individual *floridanus* sub-populations versus the subspecies as a whole.

Behavioral observations of *floridanus* individuals using direct band and recapture techniques or telemetry have suggested that little movement also occurs among the sub-

populations (D. Perkins, pers. comm.). In particular, a study on site-fidelity in a single sub-population re-captured 21 of 25 previously banded individuals within 300 m of their original banding site over two to four years (Delany *et al.* 1995). In addition, an extensive survey of movement patterns by 300 banded adults showed no movement of greater than one kilometer, which is much less than the distance separating sub-populations (D. Perkins pers. comm.).

To understand the metapopulation dynamics of this subspecies, I have used mtDNA control region sequence and six microsatellite loci to examine population structure and to calculate genetically-derived measures of dispersal among sub-populations of *floridanus*. The results of this study will serve four interrelated purposes. First, the results of the assignment test will be used to determine whether Florida Grasshopper Sparrows best fit a source-sink or a balanced dispersal model, and thus could potentially be used to predict the demographic fate of each sub-population. Second, this study will be the first to use a genetic assignment test to support the inference that the presence or absence of significant genetic differentiation is directly related to the direction and magnitude of dispersal among sub-populations. I also use the results to critique the use of direct and indirect measurements of movement. Finally, my findings will be used to complement those of a continental study (see Chapter 1) in designing an overall effective recovery plan for the Florida Grasshopper Sparrow.

METHODS

Sample Collection

A total of 105 adult Grasshopper Sparrows were captured and banded using a combination of play backs and mist nets in four Florida sub-populations: Avon Park (n = 34); Kissimmee Prairie / Latt-Maxcy (n = 10 and 20, respectively); Echo Range (n = 16); and Three Lakes Management Area (n = 25) (Figure 1). The Kissimmee Prairie / Latt-Maxcy sub-population is actually a combination of two smaller populations, the Kissimmee Prairie Sanctuary and the privately owned Latt-Maxcy Wildlife Sanctuary, separated by a minimum distance of 4 km (Delany *et al.* 1999). Although no dispersal information is available between the two sites, the two sub-populations are generally grouped together as one because of geographic proximity (P.D. Vickery pers. comm.), and will herein be referred to as Kissimmee Prairie. Specimens were collected exclusively during the breeding season from 1995 to 1998. Blood samples were collected from each individual, suspended in 1 mL lysis buffer, and then stored at –20 °C. DNA was extracted using DNAzol (Gibco BRL) or standard phenol/chloroform procedure (e.g. Sambrook *et al.* 1989).

Data collection

For details regarding the genetic data collection, please refer to Chapter 1. Briefly, I sequenced 879 bp of mtDNA control region sequence for 98 *floridanus* individuals from the four sub-populations (see Figure 2 of Chapter 1 for details). Unique sequences were assigned haplotype numbers relative to an arbitrarily assigned sequence which I deemed Grasshopper Sparrow (GRSP) haplotype 1 (see Figure 3 of Chapter 1). I

also genotyped 105 *floridanus* samples at six microsatellite loci following the protocols described in Chapter 2.

Population Analysis

Levels of Variation and Population Structure

Analyses used to assess levels of genetic variation in mtDNA haplotypes and microsatellite loci within and across the four sub-populations were performed as described in Chapter 1. Briefly, I used the program Arlequin (Schneider *et al.* 1996) to identify the number of haplotypes, nucleotide composition, the number and type of nucleotide substitutions throughout the mtDNA data set, the average number of pair-wise differences among the haplotypes, and nucleotide diversity (π) both per and across all sub-populations. I also used Arlequin to test for significant genetic differentiation in mtDNA haplotype frequencies and molecular distances among all pairs of sub-populations ($n = 6$).

For the microsatellite data, I tested for significant heterozygote deficiency per locus per sub-population, linkage disequilibrium per locus pair across sub-populations, and the number and frequency of alleles per locus for each sub-population using the program GENEPOP (Raymond and Roussett 1995). I also tested genetic structure between sub-population pairs based on microsatellite variation assuming i) an infinite alleles model (IAM) to generate theta (θ), a modified measure of F_{ST} (Weir and Cockerham 1984) and ii) a step-wise mutation model (SMM) to generate *rho* (ρ), a measure of R_{ST} (Slatkin 1995) to account for the fact that the processes underlying microsatellite mutation are unclear (DiRienzo *et al.* 1994).

Genetic Estimates of Dispersal

I used the program IMMANC (Rannala and Mountain 1997) to examine the magnitude and direction of contemporary movement among the four sub-populations. Combining the mtDNA and microsatellite datasets, I used IMMANC to determine the source of individuals in each sub-population for zero generations in the past. In order to determine if source-sink structure existed among the sub-populations, I calculated the ratio of immigrants to emigrants per sub-population. This method was used by Diffendorfer *et al.* (1995) and Doncaster *et al.* (1997) to define the status of sub-populations as either sources or sinks. Sources will show an immigrant to emigrant ratio of less than one, whereas sinks will show a ratio greater than one. Immigrants were summed across all source populations per sub-population. Likewise, emigrants were summed across all recipient sub-populations per sub-population. Finally, I used a χ^2 test to determine if dispersal was asymmetric between two sub-populations, suggesting source-sink structure within Florida.

RESULTS

Levels of Variation

I identified a total of 37 unique mtDNA sequences for the 98 *floridanus* individuals sampled (see Figure 3 and Appendix 1 of Chapter 1 for details). There were 34 polymorphic sites consisting of 23 transitions, 9 transversions, and 2 indels. Of these 37 haplotypes, 24 (65%) were found in only one sub-population. Haplotypes GRSP6 and 22 were most common across sub-populations, both occurring at an overall frequency of 0.11 (11 individuals each). The 37 haplotypes differed by an average of 8.33 ± 3.71 (SD) pair-wise nucleotide substitutions and had a nucleotide diversity (π) per site of $0.009 \pm$ 0005.

Within each sub-population I observed an average 0.60 distinct haplotypes per individual, ranging from 0.45 (15 for 33) to 0.79 (11 for 14) in Avon Park and Echo Range, respectively (Table 1). Pair-wise differences among haplotypes averaged $4.65 \pm$ 2.37, ranging from 3.80 ± 1.98 in Three Lakes Management Area to 5.60 ± 2.73 in Kissimmee Prairie. Likewise, nucleotide diversity averaged 0.005 ± 0.003, ranging from 0.004 ± 0.003 in Three Lakes Management Area to 0.006 ± 0.003 in Kissimmee Prairie.

For microsatellites, I observed an average of 13 alleles per locus across all Florida samples, ranging from 6 (*Ma*μ23) to 21 (*As*μ18) (Table 2). The number of alleles per locus per individual ranged from 0.15 in Avon Park at *Ma*μ23, to 0.94 in Echo Range at *As*μ18. I did not observe consistent significant heterozygote deficiency ($p > 0.05$) either across loci within a sub-population or across sub-populations at a particular locus (Table 3). This suggests that the individuals comprising each sub-population mate at random

and that there are no null alleles present. Finally, I did not detect significant linkage disequilibrium ($p > 0.05$) for any pair-wise combination of the six microsatellite loci, thereby confirming that each locus segregates independently of the others.

Genetic Differentiation

I did not detect significant genetic differentiation among the sub-populations for either mtDNA or microsatellite DNA variation. Specifically, the sub-populations did not differ significantly in haplotype frequencies (F_{ST} = 0.011 to 0.026; Table 4), or when differences between haplotypes were distance-weighted (F_{ST} = -0.500 to 0.004; Table 4), nor did the θ or ρ values differ significantly from zero in any pair-wise comparison (θ = -0.001 to 0.012; ρ = 0.003 to 0.033; Table 5).

Genetic Measures of Dispersal

The power to detect immigrants approached 1.00, ranging from 0.94 (emigrants of Kissimmee Prairie in Avon Park) to 0.98 (emigrants of Three Lakes Management Area in Echo Range) (Figure 2). I identified a total of 10 immigrants among the 105 individuals (10%), ranging from one (4% Three Lakes Management Area; 6% Echo Range) to five (15% Avon Park) (Table 6 and Figure 3). All alternative assignments had a power greater than 0.94. Although all sub-populations received and donated migrants, Echo Range did not exchange migrants with either of the Kissimmee Prairie or Three Lakes sub-populations.

For determining the presence or absence of source-sink structure among the sub-populations, I calculated the ratio of immigrants to emigrants per sub-population. This ratio was near one (average 0.8) in each of the sub-populations, ranging from 0.5 in Echo

Range and Three Lakes, to 1.5 in Kissimmee Prairie. These ratios suggest that although the number of emigrants is balanced by a near equal number of immigrants in each of the sub-populations, there is evidence that some sub-populations exhibit net emigration (Echo Range and Three Lakes) while others (Avon Park and Kissimmee Prairie) exhibit net immigration. A χ^2 test per sub-population confirmed that the number of immigrants was not significantly different from the number of emigrants ($\chi^2 \leq 0.171$, df = 1, $p >$ 0.05) (Table 6).

DISCUSSION

Population Structure

My major finding was that there was no significant genetic structure among sub-populations of the Florida Grasshopper Sparrow and this was associated with moderately high levels of interpopulation movement detected using the assignment test. These results suggest that the lack of significant structure observed among the sub-populations was in fact due to high levels of contemporary gene flow (Slatkin 1983) and that the Florida sub-populations meet the traditional definition of a metapopulation, characterized as a set of sub-populations linked by dispersal (Hanski and Simberloff 1997).

Other studies of avian species have revealed significant genetic structure on a similar geographical scale. Using microsatellite loci, McDonald *et al.* (1999) observed significant differentiation among populations of the Florida Scrub Jay (*Aphelocoma coerulescens*) separated by a mean distance of 150 km. As a consequence of this study, McDonald *et al.* suggested that all Florida Scrub Jay populations be managed as discrete conservation units. Likewise, Oyler-McCance *et al.* (1999) found significant genetic structure among Small-bodied Sage Grouse (*Centrocercus urophasianus phaios*) populations in Colorado using both mtDNA and microsatellite data. In both cases, the structure observed was explained by restricted interpopulation gene flow supported by Nm values less than one. Rhodes *et al.* (1995) also found significant genetic structure among populations of Rio Grande Wild Turkeys (*Meleagris gallopavo intermedia*) separated by 12.7 to 60 km. Contrary to the previous two studies, Rhodes *et al.* (1995) speculated that the differentiation was a result of polygynous mating. In this system,

substantial genetic variation is lost each generation due to a few selected males gaining the majority of mating opportunities.

In contrast, Van Den Busche *et al.* (1999) did not detect significant differences in allele frequencies among threatened breeding colonies of Wood Storks (*Mycteria americana*) in the southeastern United States. Since it was also believed that these colonies exchanged migrants between breeding seasons, they recommended that the colonies be managed as a single panmictic breeding population. Kvist *et al.* (1999) also detected genetic continuity among local populations of the Blue Tit (*Parus caeruleus*) in Northern Europe. Although these studies imply that significant genetic differentiation is possible among geographically proximate bird populations, it is obvious from the above two examples and my study that genetic discontinuity among heterogeneously distributed avian sub-populations can potentially be prevented with low numbers of migrants.

The lack of significant differentiation between Echo Range and each of Kissimmee Prairie and Three Lakes Management in spite of an absence of movement between these two pairs of sub-populations may have two equally plausible biological explanations. First, movement between the sub-populations may have ceased only recently, hence the sub-populations have had insufficient time since divergence to show significant differentiation. Alternatively, movement may normally occur between these pairs of sub-populations but did not occur during the generations sampled in this study. In order to assess these two possibilities, data of the type described here should be collected over a much longer time period.

Metapopulation Dynamics

It is obvious from the results of the assignment test that the four sub-populations differ in the percent of immigrants (4 to 15%) comprising each sub-population. However, the ratio of immigrants to emigrants was near unity in each sub-population (0.5 to 1.2) and similar in magnitude to that observed in Collared Flycatchers (0.6 to 1.5; Doncaster *et al.* 1997), and some rodent species (Diffendorfer *et al.* 1995). A source-sink structured metapopulation is defined by differences in population growth, and directional net movement among the sub-populations (Pulliam 1988). In contrast, a metapopulation with balanced dispersal exhibits no differences in net movement or reproductive success among the sub-populations (Diffendorfer *et al.* 1995; Doncaster *et al.* 1997). According to the basic criteria of the two metapopulation dynamics models, the near unity ratio observed in each of the sub-populations suggests that the *floridanus* metapopulation may best fit a balanced dispersal model.

The balanced ratio observed in Kissimmee Prairie is particularly striking if one considers the environmental perturbations that this area has experienced in the last five years, especially since Delany and Linda (1998) noted that Florida Grasshopper Sparrows appear unable to adapt to drastic habitat disturbances. During the four breeding seasons that *floridanus* specimens were collected (1995-1998), the Kissimmee Prairie Sanctuary territories experienced severe flooding which left little suitable habitat for nesting (P.N. Gray, pers. comm.). Subsequent to 1995, the observed number of breeding territories in the Kissimmee Prairie Sanctuary decreased from 30 (1995) to 13

(1996), to 6 (1997), to 1.5 (1998), and were completely absent in 1999 (D. Perkins pers. comm.). Fledglings (n = 3) were last present in 1996.

Considering the severe flood damage in the Kissimmee Prairie Sanctuary during the collection period, one prediction is that the sub-population would serve as a sink because mortality exceeded natality at this time. On the contrary, I observed balanced movement in and out of the Kissimmee Prairie sub-population suggesting that natality is at least balancing mortality. However, the Kissimmee Prairie sub-population has gone extinct over this time period. Thus, the patterns of immigration and emigration in Kissimmee Prairie most likely reflect characteristics of the larger Latt-Maxcy sub-population which was unaffected by the flooding (P.N. Gray pers. comm.).

Many studies of avian metapopulations have observed source-sink versus balanced dispersal dynamics (Sæther *et al.* 1999; Brawn and Robinson 1996; McCoy *et al.* 1999). However, in most of these cases the sub-populations have exploited habitats varying in quality, where the sink sub-populations typically exploit sub-optimal habitats and the sources exploit optimal habitats (Dias 1996). Extensive searches for Florida Grasshopper Sparrows and their preferred habitat within Florida have confirmed that the known sub-populations described here represent the entire distribution of the subspecies (P.D. Vikery, pers. comm.). It is also known that Grasshopper Sparrows are habitat specialists, preferring dry prairie-like secondary successional vegetation six months to two years following natural or prescribed burning (Delany *et al.* 1985). Considering the ecological limitations of the species and the general characteristics of source and sink habitats, it seems likely that the Florida metapopulation best fits a balanced dispersal

model in which dispersal occurs between sub-populations occupying long-term habitat patches.

In other studies of source-sink dynamics in birds, Brawn and Robinson (1996) observed population growth rates less than one in six Neotropical migrant species occupying fragmented forest habitats, indicating that these sub-populations were potentially sinks. Conversely, McCoy *et al.* (1999) estimated the necessary fecundity of populations to achieve a population growth rate greater than one to identify source and sink populations of seven passerine species (including the Grasshopper Sparrow) occupying Conservation Reserve Program (CRP) reverted agricultural plots. Using this method, McCoy *et al.* (1999) estimated Grasshopper Sparrow fecundity values greater than that needed for a population growth rate greater than one per and across years (n = 3) and concluded that Missouri CRP plantings function as source populations. Although these estimates can be used to address whether particular populations have net natality or mortality as required for determining source-sink structure, Brawn and Robinson (1996) and McCoy *et al.* (1999) also stated that their conclusions were provisional until direct dispersal data could be collected for these populations. Considering this, it may also be of future interest to apply Pulliam's replacement method to the *floridanus* populations, to support the conclusions interpreted from the assignment tests.

Direct Versus Indirect Measures of Dispersal

Here I have used genetic data to infer patterns of movement usually estimated by recapture or radio-telemetry data. Generally speaking, indirect genetic methods are

superior to direct methods because it is not necessary to sample the individual following initial banding event. However, indirect methods do have a number of disadvantages.

A major problem with genetic assignment tests is that they are indirect. These methods assign individuals to sub-populations based entirely on probability. For example, IMMANC calculates the likelihood of an individual's genotype in each sub-population and then determines whether the likelihood of the individual being from a sub-population other than the one it was sampled in is significantly greater than its likelihood in the sampled sub-population. It assumes that an individual assigned to another population is a dispersing individual and yet without the direct evidence from trapping banded specimens, such assignments could conceivably occur due to chance alone.

To compensate for the underlying possibility of incorrectly assigning an individual to an alternative sub-population, IMMANC also calculates a power value which acts as a reliability index for each assignment. This is an important value to consider since the program does not correct for multiple comparisons. Of the 315 possible alternative assignment calculations I performed, all were over 0.93 per individual per assignment, regardless of whether the assignment was significant or not. This means that the ability to detect an immigrant from an alternative sub-population was a minimum of 93% reliable, and a maximum of 7% unreliable for any given individual from any of the three potential alternative source sub-populations. These results suggest that the program has high power to correctly assign an individual to the right source sub-population. However, because alternative assignments of individuals are only made

when the likelihood of an individual's genotype is significantly greater in an alternative sub-population, it is possible that, also due to chance events, real immigrants may not be detected because they are not sufficiently genetically differentiated. Thus, my estimates of dispersal likely represent minimum estimates of dispersal among these sub-populations.

As an alternative, it is possible to not do any statistical testing of the ln Λ values and instead assign individuals solely on the basis of a negative difference in the likelihood values, as done in the Doh assignment test (Paetkau *et al.* 1995). However, this results in an increased frequency of misassignments, particularly when differentiation between populations is low (Waser and Strobeck 1998; Haig *et al.* 1997). Furthermore, an extensive examination of assignment tests by Cornuet *et al.* (1998) concluded that IMMANC correctly assigned more individuals than frequency- and genetic distance-based methods regardless of sample size, number of loci assayed, microsatellite mutation model, and levels of differentiation.

In the present study I have calculated dispersal estimates across a single generation. However, it would be valuable to average such estimates across multiple generations due to the possibility of annual variation in dispersal rates. For instance, both Diffendorfer *et al.* (1995) and Doncaster *et al.* (1997) based their conclusion of balanced dispersal among sub-populations of rodents and Collared Flycatchers, respectively, on demographic data collected over several consecutive breeding seasons. While annual fluctuations in the ratio of immigrants to emigrants per sub-population did occur, the ratios were near one when averaged across years. Thus, it is possible that my results may

actually reflect a sporadic year of balanced dispersal in an otherwise source-sink structured metapopulation. Accordingly, I suggest that the assignment test be repeated for newly banded adults on an annual basis for a minimum of five breeding seasons to ensure that the metapopulation is in fact exhibiting balanced dispersal.

IMMANC also offers an option to detect immigrants which have entered the population a certain number of generations in the past. Because Grasshopper Sparrow adults live for an average of two years (Delany *et al.* 1995; Vickery 1996), it would be useful to determine if I would detect any second-generation immigrants in each sub-population. However, when I ran a second assignment test with the goal of identifying dispersers in the present generation and one generation in the past, power values for detected dispersers one generation ago were very low (0.62 to 0.72). This decrease in power is most likely a result of the need to estimate what the allele frequency distributions per locus per sub-population for the past generation were from the present distributions. Hence, I can only be confident of inferring movement among the sub-populations if I assume that all dispersal occurred in the present generation. Furthermore, these estimates represent minimal values of movement among the sub-populations because they neglect dispersal over multiple generations.

In this study, I was able to detect individual immigrants with a high degree of power among sub-populations which showed little genetic differentiation. The likely reason for this phenomenon is the different way in which assignment tests and measures of population differentiation are calculated. Individual assignments are based on the calculation of likelihood values. For instance, IMMANC calculates the likelihood of any

given genotype as the product of the probabilities of an individual possessing a particular allele at a particular locus, given the frequency of the alleles in the sub-population (Rannala and Mountain 1997). Therefore, the likelihood of an individual's genotype is greatly affected by the frequency of individual alleles in a sub-population (Waser and Strobeck 1998). In contrast, measures of population differentiation, such as F_{ST}, are less sensitive because they are based on population-wide comparisons of the distributions of alleles for sets of individuals (Whitlock and McCauley 1999; Slatkin 1995). Because of this difference in power, it is possible to detect an individual with high power despite a lack of genetic differentiation among the sub-populations (Waser and Strobeck 1998).

There is some discrepancy between the direct and indirect measures of movement among the four sub-populations. First, D. Perkins (pers. comm.) observed very little movement (< 200 m) of 5 radio-tagged females. In addition, he banded 300 adults and observed no movements greater than one kilometer. Finally, neither D. Perkins nor T.F. Dean (pers. comm.) have ever observed movement between any two sub-populations. However, using indirect genetic methods I have detected immigrants between sub-populations geographically separated by 5 to 18 km. To date, only one known study has compared movement measures calculated from band and recapture methods to those estimated from genetic assignment tests. Favre *et al.* (1997) estimated immigration rates among sub-populations of the Greater White-Toothed Shrew (*Crocida russula*) in Switzerland using both band and recapture methods and the Doh frequency assignment method. Both methods suggested high degrees of natal philoptry and low levels of female-biased juvenile dispersal. Based on this study it is apparent that the two methods

can complement each other. However, mammals are inherently easier to capture and theoretically disperse over shorter distances than birds, which enabled Favre *et al.* to have a considerable sample size (146 total) for both measurements. D. Perkins (pers. comm.) acknowledged this inherent problem in that he was only able to recapture 1 of 50 banded Florida Grasshopper Sparrow nestlings. In addition, Delany *et al.* (1995) were only able to recapture or resight 25 of 48 banded males over a four year period. Considering that the probability of physical recapture in band and recapture or radio-telemetry studies is low in avian species, the most likely explanation for the observed discrepancy is that dispersing Florida Grasshopper Sparrows are missed using direct methods. The results of the present study also suggest that indirect methods of measuring movement within metapopulations may be superior to direct methods in general when recapture rates are low.

Conservation Implications and Conclusions

My results suggest that the lack of genetic structure among the *floridanus* sub-populations is most likely a result of contemporary gene flow on a metapopulation scale. On the basis of these results, I suggest that the *floridanus* subspecies be managed as a single management unit on a regional scale versus multiple management units on a sub-population scale (Moritz 1994). In terms of metapopulation dynamics, my results indicate that there is zero net movement among the four *floridanus* sub-populations.

LITERATURE CITED

Armitage, K.B. 1991. Social and population dynamics of yellow-bellied marmots: results from long-term research. *Ann. Rev. Ecol. Syst.* 22: 379-407.

Barrowclough, G. 1978. Sampling bias in dispersal studies based on finite area. *Bird Banding* 49: 333-341.

Barrowclough, G. 1980. Gene flow, effective population sizes, and genetic variance components in birds. *Evolution* 34: 789-798

Brawn, J.D. and S.K. Robinson. 1996. Source-sink population dynamics may complicate the interpretation of long-term census data. *Ecology* 77: 3-12.

Cornuet, J., S. Piry, G. Luikart, A. Estoup, and M. Solignac. 1999. New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* 153: 1989-2000.

Davies, N., F.X. Villablanca and G.K. Roderick. 1999. Determining the source of individuals: multilocus genotyping in nonequilibrium population Genetics *TREE* 14: 17-21.

Dean, T.F., M.F. Delany, E.W. Chapman, and P.D. Vickery. 1998. Longevity and site

fidelity of Florida Grasshopper Sparrows. *J. Field Ornithol.* 69: 51-54.

Delany, M.F. and S.B. Linda. 1998. Characteristics of Florida Grasshopper Sparrow nests. *Wilson Bulletin* 110: 136-139.

Delany, M.F., C.T. Moore, D.R. Progulske Jr. 1995. Territory size and movements of Florida Grasshopper Sparrows. *J. Field Ornithol.* 66: 305-309.

Delany, M.F., P.B. Walsh, B. Pranty, and D.W. Perkins. 1999. A previously unknown population of Florida grasshopper sparrows on Avon ark Air Force Range. *Florida Field Naturalist* 27: 52-56.

Delany, M.F., H.M. Stevenson, and R. McCracken. 1985. Distribution, abundance, and habitat of the Florida Grasshopper Sparrow. *J. Wildl. Manage.* 49: 626-631.

Dias, P.C. 1996. Sources and sinks in population biology. *TREE* 11: 326-329.

Diffendorfer, J.E., M.S. Gaines, and R.D. Holt. 1995. Habitat fragmentation and movements of three small mammals (*Sigmodon, Microtus,* and *Peromyscus*) *Ecology* 76: 827-839.

DiRienzo, A., A.C. Peterson, J.C. Garza, A.M. Valdes, M. Slatkin, and N.B. Friemer.

1994. Mutational processes of simple-sequence repeat loci in human populations. *Proc. Natl. Acad. Sci. U.S.A.* 91: 3166-3170.

Doncaster, C.P., J. Clobert, B. Doligez, L. Gustafsson, and E. Danchin. 1997. Balanced dispersal between spatially varying local populations: an alternative to the source-sink model. *Am. Nat.* 150: 425-445.

Donovan, T.M., R.H. Lamberson, A. Kimber, F.R. Thompson III, and J. Faaborg. 1995. Modeling the effects of habitat fragmentation on source and sink demography of neotropical migrant birds. *Conservation Biology* 9: 1396-1407.

Favre, L., F. Balloux., Goudet, J., and N. Perrin. 1997. Dispersal patterns in *Crocida russula*. *Proc. R. Soc. Lond. B.* 264: 127-132.

Haig, S.M., C.L. Gratto-Trevor, T.D. Mullins, and M.A. Colwell. 1997. Population identification of western hemisphere shorebirds throughout the annual cycle. Molecular Evolution 6: 413-427.

Hanski, I. and D. Simberloff. 1997. The Metapopulation Approach, Its History, Conceptual Domain, and Application to Conservation. *In*: "Metapopulation Biology: Genetics, Ecology, and Evolution". (I. Hanski and M.E. Gilpin, Eds). pp. 5-26. Academic Press.

Koenig, W.D., D. Van Vuren, P.N. Hooge. 1996. Detectability, philopatry, and the distribution of dispersal distances in vertebrates. *TREE* 11: 514-517.

Kvist, L., M. Ruokonen, J. Lumme, and M. Orell. 1999. Different population structures in northern and southern populations of the european blue tit (*Parus caeruleus*). *J. Evol. Biol.* 12: 798-805.

Lahaye, W.S., R.J. Gutierrez, and H.R. Akcakaya. 1994. Spotted owl metapopulation dynamics in southern california. *Journal of Animal Ecology* 63: 775-785.

Levins, R. 1970. Extinctions. *Lect. Math. Life Sci.* 2: 75-107.

McDonald, D.B., W.K. Potts, J.W. Fitzpatrick, and G.E. Woolfenden. 1999. Contrasting genetic structures in sister species of North American scrub-jays. *Proc. R. Soc. Lond.* B. 266: 1117-1125.

Matthysen, E. 1999. Nuthatches (*Siita europaeae*: Aves) in forest fragments: demography of a patchy population. *Oecologia* 119: 501-509.

McCoy, T.D., M.R. Ryan, E.W. Kurzejeski, and L.W. Burger. 1999. Conservation reserve program: source or sink habitat for grassland birds in Missouri? *J. Wildl. Manage.* 63: 530-537.

McGowan Jr., D.P. and D.L. Otis. 1998. Population demographics of two local South Carolina mourning dove populations. *J. Wildl. Manage.* 62: 1443-1451.

Moore, W.S. and R.A. Dolbeer. 1989. The use of banding recovery data to estimate dispersal rates and gene flow in avian species: case studies in the red-winged blackbird and common grackle. *Condor* 91: 252-253.

Moritz, C. 1994. Applications of Mitochondrial DNA Analysis in Conservation: A Critical Review. *Molecular Evolution* 3: 401-411.

Oyler-McCance, S., N.W. Kahn, K.P. Burnham, C.E. Braun, and T.W. Quinn. 1999. A population genetic comparison of large- and small-bodied sage grouse in Colorado using microsatellite and mitochondrial DNA markers. *Molecular Evolution* 8: 1457-1465.

Paetkau, D., W. Calvert, I. Stirling and C. Strobeck. 1995. Microsatellite analysis of population structure in Canadian polar bears. *Molecular Evolution* 4: 347-354.

Paradis, E., S.R. Baillie, W.J. Sutherland, and R.D. Gregory. 1998. Patterns of natal and breeding dispersal in birds. *Journal of Animal Ecology* 67: 518-536.

Pulliam, H.R. 1988. Sources, sinks, and population regulation. *Am. Nat.* 132: 652-661.

Rannala, B. and J.L. Mountain. 1997. Detecting immigration by using multilocus genotypes. *Proc. Natl. Acad. Sci. USA*. 94: 9197-9201.

Rhodes, O.E., Jr., D.J. Buford, M.S. Miller, and R.S. Lutz. 1995. Genetic structure of reintroduced Rio Grande wild turkeys in Kansas. *J. Wildl. Manage.* 59: 771-775.

Robinson, S.K., F.R. Thompson III, T.M. Donovan, D.R. Whitehead, and J. Faaberg. 1995. Regional forest fragmentation and the nesting success of migratory birds. *Science* 267: 1987-1990.

Sæther, B., T.H. Ringsby, Ø. Bakke, and E. J. Solberg. 1999. Spatial and temporal variation in demography of a house sparrow metapopulation. *Journal of Animal Ecology* 68: 628-637.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular Cloning. Cold Springs Harbor Laboratory, Cold Springs Harbor, New York.

Schneider, S. J.-M. Kueffer, D. Roessli, and L. Excoffier. 1996. Arlequin: A software package for population genetics. Genetics and Biometry Lab, Department of Anthropology, University of Geneva, Geneva, Switzerland.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139: 457-462.

Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. *Science* 236: 787-792.

Van Den Bussche, R.A., S.A. Harmon, R.J. Baker, A.L. Bryan Jr., J.A. Rodgers Jr., M.J. Harris, I.L. Bribin Jr.. 1999. Low levels of genetic variability in North American populations of the wood stork (*Mycteria americana*). *Auk* 116: 1083-1092.

Vickery, P.D. 1996. Grasshopper Sparrow (*Ammodramus savannarum*). *In*: The Birds of North America, no 239 (A. Poole and F. Gills, Eds.). Academy of Natural Sciences, Philadelphia, and American Ornithologists' Union, Washington, D.C.

Waser, P.M. and C. Strobeck. 1998. Genetic signatures of interpopulation dispersal. *TREE* 13: 43-44.

Weatherhead, P.J. and M.R.L. Forbes. 1994. Natal philopatry in passerine birds: genetic or ecological influences? *Behavioral Ecology* 5: 426-433.

Weir, B.S. and C.C. Cockerham. 1984. Estimating *F*-Statistics for the analysis of population structure. *Evolution* 38: 1358-1370.

White, G. C., and R. A. Garrott. 1990. Analysis of wildlife radio-tracking data. Academic Press. 350 pp.

Whitlock, M.C. and D.E. McCauley. 1999. Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$. *Heredity* 82: 117-125.

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

Table 1. Variation in mtDNA among Florida sub-populations. Shown are the number of individuals sampled (N), number of haplotypes observed (No. Haplotypes), number of polymorphic sites, average number of pair-wise differences among haplotypes, and nucleotide diversity per site (π) per population.

	Avon Park	Kissimmee Prairie	Echo Range	Three Lakes
N	33	29	13	24
No. Haplotypes	15	18	10	13
Polymorphic sites	27	29	18	17
Average pair-wise differences	4.48	5.53	4.80	3.80
π	0.005	0.006	0.005	0.004

Table 2. Microsatellite DNA variation in Florida Grasshopper Sparrow sub-populations. Shown are the number of individuals (N), number of alleles (No. alleles), and the frequency of the most common allele.

Population		Locus					
		*Dp*μ16	*Ma*μ23	*Ca*μ02	*As*μ09	*As*μ15	*As*μ18
Avon Park	N	34	34	34	33	34	34
	No. alleles	9	5	8	9	14	17
	Frequency	0.515	0.662	0.353	0.500	0.235	0.250
Kissimmee Prairie	N	30	30	29	30	30	30
	No. alleles	6	5	10	12	16	18
	Frequency	0.617	0.700	0.293	0.367	0.200	0.250
Echo Range	N	16	16	16	14	16	16
	No. alleles	6	5	9	5	5	16
	Frequency	0.647	0.647	0.353	0.433	0.235	0.176
Three Lakes	N	25	25	25	23	25	25
	No. alleles	7	5	11	6	6	12
	Frequency	0.620	0.720	0.380	0.435	0.200	0.360

Table 3. Frequency of observed (H_o) and expected (H_e) heterozygotes per locus per sub-population. Asterisks indicate those loci which exhibit significant heterozygote deficiency as determined using GENEPOP.

		Locus					
		*Dp*μ16	*Ma*μ23	*Ca*μ02	*As*μ09	*As*μ15	*As*μ18
Avon Park	H_o	0.70	0.50	0.61	0.73	0.91	0.94
	H_e	0.70	0.53	0.74	0.70	0.88	0.88
Kissimmee Prairie	H_o	0.53	0.48	0.86	0.83	0.83	0.88*
	H_e	0.60	0.48	0.83	0.83	0.90	0.90
Echo Range	H_o	0.59	0.53	0.71	0.59*	0.82	0.88
	H_e	0.59	0.53	0.82	0.73	0.88	0.94
Three Lakes	H_o	0.56*	0.40	0.80	0.65	0.72*	0.80
	H_e	0.60	0.48	0.80	0.70	0.88	0.80

Table 4. F_{ST}-based values for pair-wise comparisons of mtDNA haplotype frequencies (above diagonal) and weighted for amount of divergence between haplotypes (below diagonal) among the four *floridanus* sub-populations. † $p < 0.05$.

	Avon Park	Kissimmee Prairie	Echo Range	Three Lakes
Avon Park	-	0.019	0.016	0.024†
Kissimmee Prairie	-0.001	-	0.013	0.011
Echo Range	-0.022	-0.008	-	0.026
Three Lakes	-0.006	-0.050	0.002	-

Table 5. Theta (above diagonal) and *rho* (below diagonal) values for pair-wise comparisons among the four *floridanus* sub-populations. † $p < 0.05$.

	Avon Park	Kissimmee Prairie	Echo Range	Three Lakes
Avon Park	-	0.007	0.007	0.012†
Kissimmee Prairie	0.033†	-	-0.001	0.005
Echo Range	0.025	0.020	-	0.002
Three Lakes	0.006	0.003	0.028	-

Table 6. Assignments of individuals to sub-populations within Florida. Rows represent the sample subpopulation; columns represent the source subpopulation. Percent immigrants shows the percentage of the total number of individuals sampled from a subpopulation that were determined to be immigrants. The number of immigrants was calculated per sub-population by summing the number of alternative assignments across source sub-populations (columns). The number of emigrants per sub-population was calculated as the sum of alternative assignments across recipient sub-populations (rows). The ratio of immigrants to emigrants was calculated by dividing the number of immigrants by the number of emigrants per sub-population. All chi-square values are not significant.

	Avon Park	Kissimmee Prairie	Echo Range	Three Lakes	Number of Immigrants	Percent immigrants
Avon Park	29	1	2	2	5	15
Kissimmee Prairie	3	27	0	0	3	10
Echo Range	1	0	15	0	1	6
Three Lakes	0	1	0	24	1	4
Number of Emigrants	4	2	2	2		
Immigrants/Emigrants	1.2	1.5	0.5	0.5		
χ^2	0.056	0.212	0.171	0.171		

Figure 1. Locations of the four *floridanus* sub-populations. APAFR: Avon Park Air Force Range; contains the Avon Park sub-population (2; $n = 34$) and the Echo Range sub-population (3; $n = 16$). Samples were also collected from Bravo Range (1), but were considered part of the Avon Park sub-population due to geographic proximity and habitat considerations. TLWMA: Three Lakes Wildlife Management Area ($n = 25$). KPSP: Kissimmee Prairie State Preserve; contains Latt-Maxcy sub-population ($n = 20$). OPKPS: Ordway-Whittell Kissimmee Prairie Sanctuary from which I obtained the Kissimmee Prairie sub-population samples ($n = 10$). Samples collected from OPKPS and KPSP were also pooled for all population analyses, due to geographic proximity and habitat considerations. This figure was originally published as Figure 1 in Delany *et al.* (1999).

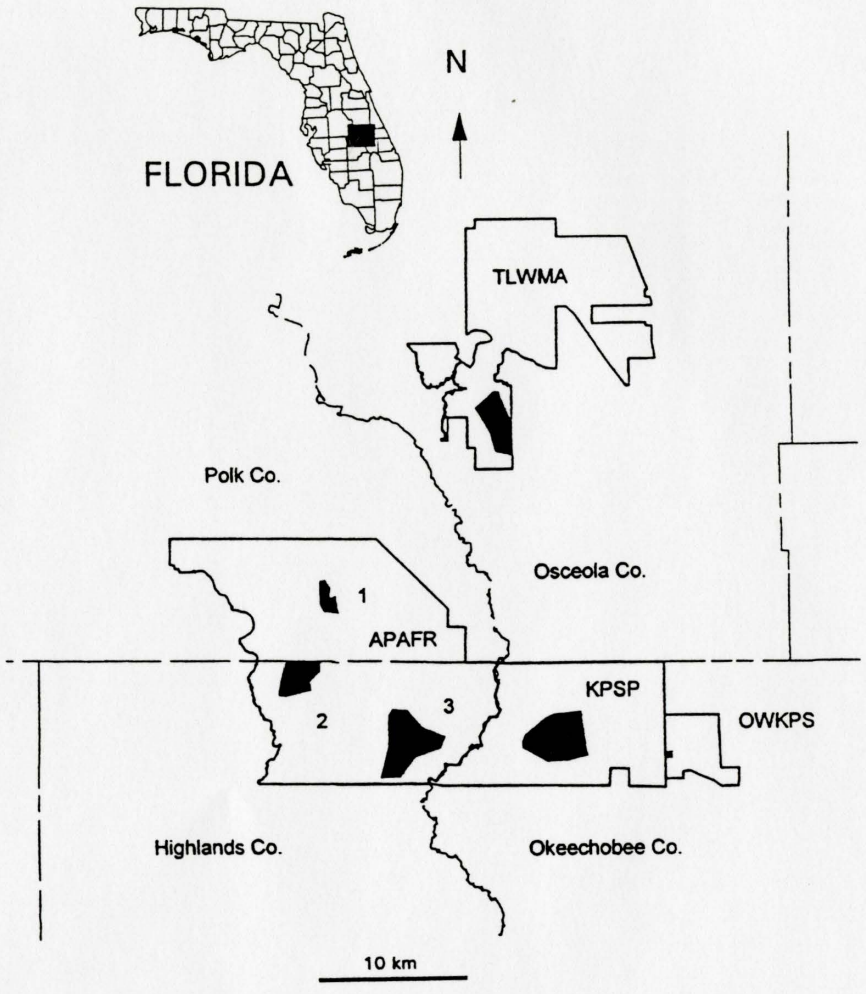

FLORIDA
N
TLWMA
Polk Co.
Osceola Co.
1
APAFR
KPSP
2
3
OWKPS
Highlands Co.
Okeechobee Co.
10 km

Figure 2. Frequency distribution of the power values for IMMANC likelihood assignments. This distribution includes three power values per individual, for a total of 315 comparisons. Power in this application represents the ability to detect an individual in an alternative population.

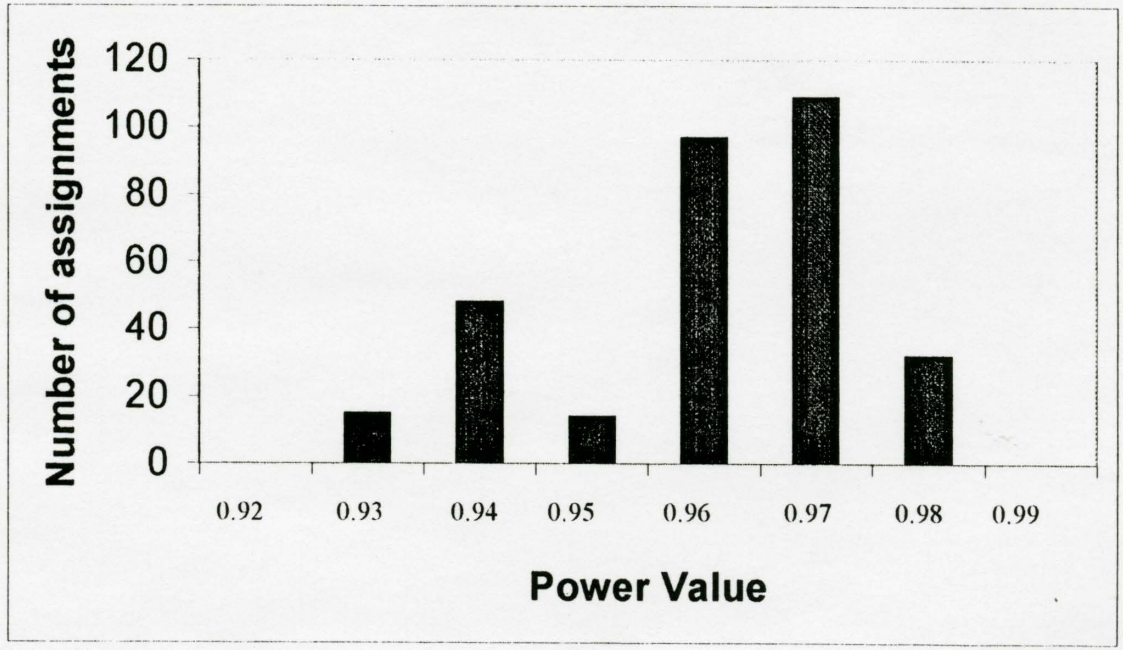
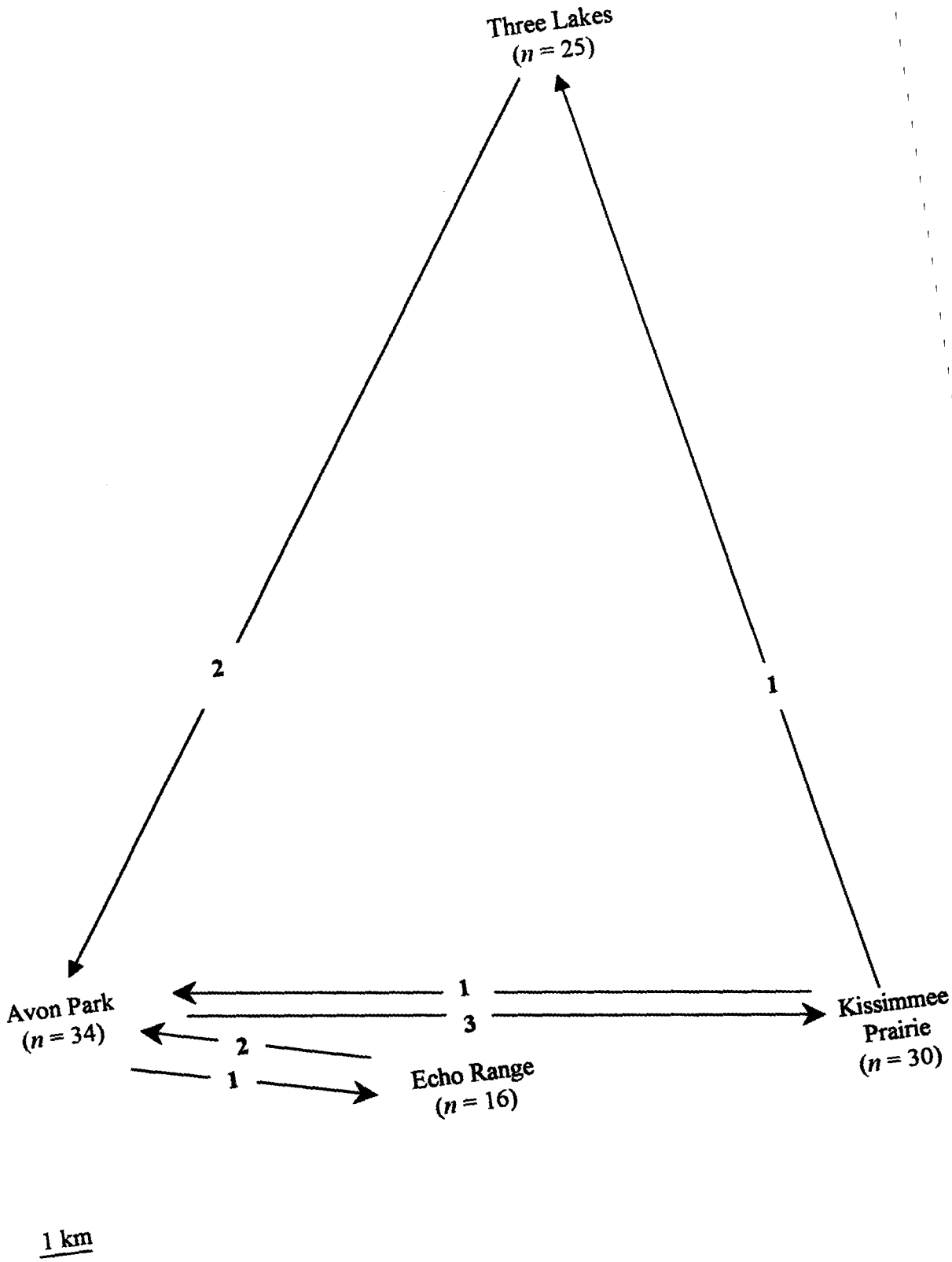

Number of assignments
120
100
80
60
40
20
0
0.92
0.93
0.94
0.95
0.96
0.97
0.98
0.99
Power Value

Figure 3. Movement among the four sub-populations: Avon Park (AP), Kissimmee Prairie (KP), Echo Range (EC), and Three Lakes Management Area (TL). Sample sizes are located in parentheses below sub-population name. Arrows indicate direction of immigration. The absolute number of immigrants is indicated within the arrow between two sub-populations.


Three Lakes
(n = 25)
2
1
Avon Park
(n = 34)
1
3
2
1
Kissimmee
Prairie
(n = 30)
Echo Range
(n = 16)
1 km

Appendix 1. Distribution of *floridanus* haplotypes among the four sub-populations.

Haplotype	Avon Park	Kissimmee Prairie	Echo Range	Three Lakes	Total
GRSP1	3				3
GRSP2				1	1
GRSP4		3		5	8
GRSP5				1	1
GRSP6	7		1	3	11
GRSP7	3	3	2	1	9
GRSP9	1		1		2
GRSP12		1			1
GRSP14	1	2			3
GRSP16			2		2
GRSP21		1	1	1	3
GRSP22	4	3	1	3	11
GRSP29	2	1		2	5
GRSP35			2		2
GRSP37				1	1
GRSP38	2	1			3
GRSP39			1		1
GRSP40			1		1
GRSP41	4	3			7
GRSP42		1			1
GRSP43			1		1

Appendix 1. Cont'd.

Haplotype	Avon Park	Kissimmee Prairie	Echo Range	Three Lakes	Total
GRSP44				2	2
GRSP45		1			1
GRSP46				2	2
GRSP47		3			3
GRSP48	1				1
GRSP49		1			1
GRSP50	1				1
GRSP51				1	1
GRSP52				1	1
GRSP53	1				1
GRSP54	1		1		2
GRSP55		1			1
GRSP56	1				1
GRSP57	1				1
GRSP59		1			1
GRSP60		1			1

Chapter 4 – General Conclusions

For my thesis project, I have used two rapidly evolving DNA markers (i.e. mitochondrial control region sequence and microsatellite loci) to examine historical and contemporary relationships among Grasshopper Sparrow populations in North America, including the endangered Florida Grasshopper Sparrow. The ultimate goals of the study were to identify genetically isolated Grasshopper Sparrow units on both a continental and local geographic scale, and to evaluate local metapopulation dynamics of the Florida subspecies, which will be used to design management plans for the Florida Grasshopper Sparrow.

In Chapter 2, I used these two selectively neutral DNA markers to answer specific questions related to the evolutionary history of five continental populations in North America: one in Florida, representing the *A. s. floridanus* subspecies; Georgia, Massachusetts, and Ohio, all representing *A. s. pratensis*; and one in Arizona, representing *A. s. ammolegus*. These questions were: 1) Is there evidence of genetic isolation among Grasshopper Sparrow populations in North America? 2) Are the populations linked by current migration? and 3) Have any of the populations experienced a bottleneck event at any point during their histories?

Significant genetic differentiation in mitochondrial and microsatellite allele frequencies was observed between the Florida population and each of Georgia, Ohio, Massachusetts, and Arizona, independently, but not between any pair-wise comparisons of the latter four populations. However, the populations exhibited phylogeographic

continuity, suggesting that individuals do move among geographic locales. Migration between Florida and a population of the other four pooled was also strikingly high considering the significant differences in allele frequencies observed among the populations. Acknowledging these contrasting patterns, I evaluated the possibility that retained ancestral polymorphisms were biasing the migration and phylogenetic results. In doing so, I found evidence of ongoing gene flow between the two groups. Thus, it may be possible that ongoing gene flow exists between the two groups, but that the levels of migration are not high enough to prevent significant differentiation. Finally, I did not detect evidence of a bottleneck event in any of the five populations at any point during their evolutionary histories.

On a local geographic scale, I examined the metapopulation structure of four heterogeneously distributed Florida Grasshopper Sparrow sub-populations. Using the same mitochondrial and microsatellite data as in the continental study, I did not detect significant genetic differentiation in any pair-wise comparison of the four sub-populations. Typically, the absence of genetic differentiation is an indication that the populations are connected by current gene flow. I was able to support this interpretation using an independent genetic assignment test which can detect immigrants within a population at the level of the individual. The results of the assignment test showed that each of the four sub-populations both received and donated immigrants, and thus are indeed connected by current gene flow. This study is one of the first to use a genetic assignment test as independent evidence supporting the hypothesis that a lack of significant differentiation is actually a reflection of ongoing gene flow.

The ability to detect individual immigrants within sub-populations also provides a measure for determining which metapopulation model best fits a given system. It is generally accepted that some sub-populations of a source-sink structured metapopulation exhibit net emigration (the source), while others exhibit net immigration (the sink). The results of the assignment test revealed an immigrant to emigrant ratio near one in all four sub-populations, suggesting that the Florida metapopulation is not source-sink structured, but instead best fits a balanced dispersal model.

In conclusion, these two studies have provided valuable information pertaining to the evolutionary and contemporary relationships of Grasshopper Sparrow populations in North America. Taken together, conservation efforts for the Florida Grasshopper Sparrow should focus on managing the subspecies as a single metapopulation consisting of four sub-populations, and separate from continental populations of other North American subspecies. Further genetic studies of interest may be to obtain samples from Central and South American subspecies to determine the global origin of the species. Also, the observed declines of the Florida Grasshopper Sparrow over the last century may essentially have actually had a bottleneck-like effect on the subspecies that is not detectable by the current genetic methods available. Thus, it may be of interest to obtain Florida Grasshopper Sparrow samples from the last century and compare historical to contemporary levels of genetic variability within the subspecies to determine if the human-induced population declines have in fact had a bottleneck effect.