UNIVERSITÄT OSNABRÜCK

POPULATION STRUCTURE AND GENETIC DIVERSITY OF WORTHEN'S SPARROW (*Spizella wortheni*) IN NORTHEASTERN MEXICO

Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

am Fachbereich Biologie/Chemie der Universitat Osnabrück

vorgelegt von Julio Cesar Canales Delgadillo

Osnabrück 2011

Promotionsgesuch eingereicht am: 15.07.2011

Die Arbeit wurde angeleitet von Prof. Dr. Judith Korb

Prüfungsausschuß	Vorsitzender:	Prof. Dr. Judith Korb
	1 Prüfer:	Dr. Laura Scott Morales
	2 Prüfer:	PD. Dr. Heike Feldhaar
	3 Prüfer:	Dr. Rebecca Schulte-Iserlohe

Contents

CONTENTS

General Introduction

Conservation of wild animal and plant populations has become a key issue around the world. For several decades, concerns about biodiversity and loss are rising that lead to numerous studies on conservation of biological resources, including habitat protection (Lees and Peres, 2008; Shipper et al., 2008; Corace et al., 2010; Russell et al., 2010), population dynamics (Smith et al., 2005; Sæther et al., 2005, Larsson and Svensson, 2009) and population genetics (Zinck and Dittman, 1993; Croteau et al., 2007, Rudnick et al., 2008).

The development of genetic tools to study populations at the molecular level has been one of the most important contributions to understand demographic processes in wild populations, especially on species of conservation concern. Conservation genetics addresses the impact of founder effects, genetic drift, inbreeding, loss of genetic diversity and accumulation of deleterious mutations on fragmented and declining populations, which are of special interest for threatened and rare species (Frankham, 1995). Such population genetic studies can contribute enormously to management decisions, for instance, because they can predict the loss of genetic variation (Mills, 2007). Loss of genetic variation within populations can imply increased extinction risks due to effects of inbreeding and a decreased capacity to respond to stochastic environmental changes (Frankham, 1995).

Preservation of genetic diversity at all levels of the phylogenetic hierarchy and the promotion of ecological and evolutionary processes that sustain the maintenance of biological diversity on earth are the objectives of genetic and conservation biology (Avise, 2004). The use of molecular markers as tools to improve conservation strategies is one approach to reach such objectives.

Conservation programs require a solid knowledge of the species to be protected. Understanding the species' biology and the processes affecting their population dynamics will help to make accurate decisions on what to protect and how to protect it, e.g., in the risk's assessment of species reintroduction, translocation of individuals (Sutherland et al. 2009), or for creation of corridors that allow species movement between disjunct habitats affected by farming activities. The motivation of this work is to provide information on the genetic diversity and evolutionary history of the Worthen's sparrow (*Spizella wortheni*), an endemic species affected by habitat fragmentation, that contributes to its conservation in the grassland areas of northeastern Mexico.

Molecular markers

Molecular markers such as microsatellites and mitochondrial DNA sequences (mtDNA) are widely used to estimate migration rates, population sizes, bottlenecks, kinship relatedness, and structure of wild populations (Lowe et al., 2004; Freeland, 2005; Selkoe, 2006). As research tools they are complementary, the former can uncover kinship relatedness and also population structure between conspecifics, while the latter is better suited to reveal more ancient relationships at the specific or population level.

Sequencing mtDNA allows researchers infer the evolutionary relationships between individuals, and it is particular useful to track these relations through long time periods. It also allows estimating divergence time between lineages and it is one of the most widely used markers in the study of phylogenetic relationships among taxa (Avise, 2004; Freeland, 2005). Features of mtDNA are well known and its gene arrangement is evolutionarily conserved. In higher animals the mtDNA is a small circular molecule of

16-20 kb (kilobase) that consist of 37 genes encoding 13 messenger RNAs, 2 ribosomal RNAs, and 22 transfer RNAs. A very desirable property of mtDNA is that introns, repetitive DNA and pseudogenes are practically absent (Avise, et al. 1987, Avise 2004). High mutation rates due to inefficient mutation repair mechanisms allow a relatively rapid evolution of mtDNA (compared to coding sequences of nuclear genome) at the sequence level, a characteristic that makes it useful in studies of population structure. Furthermore, due to the maternal inheritance of mtDNA, mutations do not recombine during sexual reproduction.

When the goals are to understand the recent demographic events, estimate genetic flow, present or past population sizes; existence of meta-populations, isolation by distance between geographically separated groups or kin relatedness between individuals, microsatellites are currently one of the most popular and powerful tools (Avise, 2004; Selkoe and Toonen, 2006). Also known as short tandem repeats (STRs), microsatellites are a PCR (polymerase chain reaction)-based approach for identification of polymorphic co-dominant Mendelian markers, namely the STRs. Microsatellites are abundantly scattered through the nuclear genome (Avise, 2004; Selkoe and Toonen, 2006). They typically consist of short sequences of di-, tri- or tetra-nucleotides of varying length (normally 5 – 40 base pairs, bp) that are tandemly arrayed in the chromosomes (Avise, 2004). The co-dominant nature of these markers allows an easy distinction between heterozygous and homozygous loci. Faster mutation rates than in mtDNA, and the possibility to identify more than one allele make microsatellites a powerful tool for identification of genetic relationships between con-specifics and for inferring relatively recent population genetic events (Freeland 2005).

The Worthen's Sparrow

Mexico harbors about 10% of all bird species in the world (Llorente-Bousquets and Ocegueda, 2008). This high diversity is possible because the heterogeneity of landscape and topography allow a high degree of endemism. According to the IUCN (www.iucn.org), about 8% of the birds species recorded in Mexico are endemic and 25% of them are listed by IUCN (2010) as threatened. On the other hand, sea surrounding the western and eastern borders of the country, further adds to the high diversity of habitats that together with regimes ranging from tropical to desert climates allow many species to coexist. This high bird diversity is further expanded through the occurrence of migratory bird species from South and North America caused by the convergence of the Neotropical and the Holoarctic regions.

To protect resident and migratory birds in key habitats across Canada, USA and Mexico, a network of important areas for bird's protection called *Important Bird Areas in North America* (IBAs) was created between 1996 and1998. Such areas should protect globally and nationally threatened species, species with restricted ranges or endemics as well as breeding bird congregations (CEC, 1999).

In Mexico, habitat loss and habitat fragmentation due to livestock ranching, agricultural production, forestry, tourism and industrial activities, are the main factors causing declines of birds populations (Arizmendi, 1999). Combined these factors not only reduce populations' sizes but also cause isolation of habitat patches leading to negative effects on the population's demography and genetic diversity (Manzano-Fischer et al., 1999).

One species facing considerable population declines during the last four decades is the Worthen's sparrow (*Spizella wortheni*). It is a Mexican endemic Emberizid

restricted to scrub and grassland habitats in northeastern Mexico (Wege et al. 1993, Berhstock et al. 1997). The Worthen's sparrow is a small passerine (12.5-14 cm total length, 10-13 g); it has been described as a dull sparrow with a distinctive head pattern: it has a grey head with a rufous crown and brownish one post-ocular stripe, a small and relatively deep conic bill, lacking stripes on the chest and with long tail. Although sexes are similar, males could have brighter colors than females (Fig. 1).

Figure 1. Male (a) and female (b) Worthen's Sparrows caught during the breeding season 2008.

In contrast to other species of this genus, the Worthen's sparrow does not make long-distance migration (Wege et al., 1993; Howell and Webb, 1995; Rising, 1996; Sibley 2000) but moves regionally within the currently known distribution localities.

In general, records of this species are scarce. The only one record outside Mexico comes from New Mexico, USA, made in 1884 (Wege et al., 1993). The Worthen's sparrow was formerly reported to occur in eight Mexican states. However between 1950/1960 until now, it was only recorded at a few localities in the states Coahuila, Nuevo León and San Luís Potosí (Fig. 2, Webster y Orr, 1954; Wege et al., 1993; Berhstock et al., 1997; Scott-Morales et al, 2008, Canales-Delgadillo et al., 2008). Due to the small number of breeding adults recorded in the last 40 years and to the

restricted distributional range currently known, the Worthen's Sparrow became a red list species in 1994 (BirdLife International, 2000).

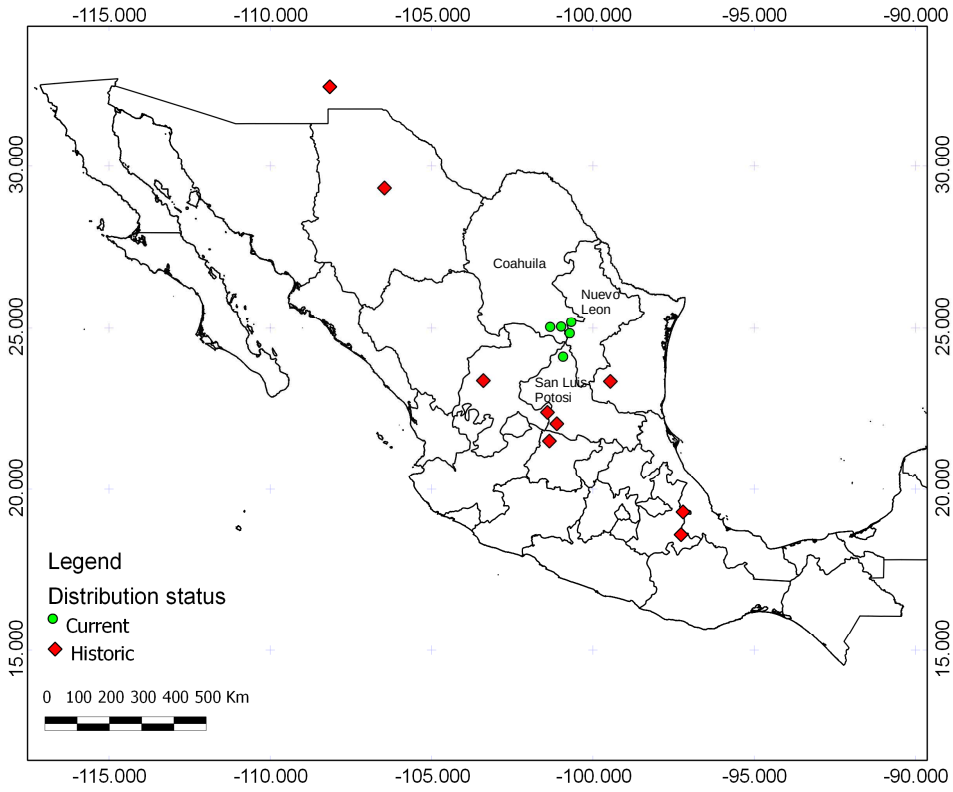

Figure 2. Historic and current distribution range of the Worthen's sparrow. Red diamonds represent locally extinct populations; green circles are the currently known distribution sites. Historical records were mapped after Wege et al. 1993.

Research on the Worthen's Sparrow has focused on its distributions (Webster and Orr 1954, Wege et al. 1993Scott- Morales et al. 2008). A few studies have also studied aspects of its conservation (Wege et al. 1993), reproduction, singing characteristics (Berhstock et al.1997; Garza de Leon 2007), and behavior (Canales-Delgadillo et al. 2008). Population size estimations exist only for one out of seven currently known distribution sites (BirdLife International 2000; Canales-Delgadillo 2006). Development of management and conservation strategies to maximize species'

survival probabilities requires a good knowledge of a species' biology and its life history. Nonetheless, information about reproductive success, mortality and dispersal rates, reproductive behavior and phylogenetic relationships are still necessary.

Aims of this study

The major aims of this study were threefold,

1. Evaluate the population genetic structure to assess at genetic level the effect of fragmentation on the population of *S. wortheni* through estimation of:
    - The genetic diversity of the species,
    - Migration rates between populations,
    - The presence of meta-populations or isolation by distance effects.
2. Determine the taxon relationships of this species with respect to its con-generics.
3. Evaluate the possible causes of rarity of this species.

To achieve the firs aim, I developed nine species-specific microsatellite markers that allowed genotyping samples of the Worthen's sparrow. The development of species-specific markers was necessary after testing of 25 markers which had been developed for other passerine species. From these, only three reliably amplified DNA samples of the Worthen's sparrow. Then, I used a set of twelve microsatellite markers to genotype one hundred DNA Worthen's sparrow samples. Genotypes were used to estimate population genetic parameters such as genetic diversity per sampling locality, levels of inbreeding, isolation by distance and panmixia.

In order to improve our understanding of the phylogenetic position of the Worthen's sparrow, I constructed a molecular phylogeny based on five mitochondrial genes for samples of the Worthen's sparrow and all congeneric species: Black-chinned

sparrow (*Spizella atrogularis*), Chipping sparrow (*Spizella arborea*), Field sparrow (*Spizella pusilla*), Clay-colored sparrow (*Spizella pallida*), Brewer's sparrow (*Spizella breweri*), Timberline sparrow (*Spizella taverneri*) and American-Tree sparrow (*Spizella arborea*). As outgoup I used Black-throated sparrow (*Amphispiza bilineata*) and Dark-eyed junco (*Junco hyemalis*). Applying Bayesian and maximum likelihood approaches, I found conflicting phylogenetic relationships among taxa, these I could resolve with the help of phylogenetic network analysis which allowed me to elucidate the causes of these uncertainties.

Finally, in order to analyze causes of rarity of my study species, I did an extensive literature research to collect geographical records of each species within the genus Spizella. Based on these occurrence records, I constructed a presence/absence matrix on vegetation communities within the region including Mexico, USA and Canada. These data together with the phylogeny study, allowed me to test for phylogenetic attraction/repulsion patterns and whether these patterns were environmentally driven. For each species, I determined marginality (how species niche mostly differs from the conditions in the global area), degree of specialization, niche breadth and niche overlap among species. This was done by means of ecological niche factor analysis.

This work will help to understand the state of the *S. wortheni* population and its relationships with other species and with the environmental factors affecting its prevalence in the shrub-grassland areas of northeastern Mexico.

Chapter 1

Isolation and characterization of nine microsatellite loci in the endangered Worthen's Sparrow (Spizella wortheni)[1]

Julio Canales-Delgadillo, Laura Scott-Morales, Oliver Niehuis & Judith Korb

Abstract

We report the isolation and characterization of nine microsatellite markers from Worthen's Sparrow (*Spizella wortheni*), an endangered bird species endemic to northeastern Mexico. We tested the markers in 32 samples of the species collected near Saltillo Coahuila, Mexico. All markers were polymorphic, with 3-31 alleles per locus, and no difference was found between the observed and expected heterozygosity when applying a sequential Bonferroni correction and a table-wide significance level of 0.05. We found no evidence for linkage disequilibrium between the markers and estimated a null allele frequency of 0.00-0.05. The new markers will allow elucidating the genetic structure and life history of Worthen's Sparrow and guiding conservation efforts for this endangered species.

Keywords: Aves, Emberizidae, New World sparrows, *Spizella wortheni*, microsatellites

The Worthen's Sparrow (*Spizella wortheni*) is an endemic bird of the semi-arid shrub and grassland areas of northeastern Mexico. While historically more widely distributed, its present day range is restricted to only a few sites on the Mexican High Plateau (Wege et al. 1993). Little is known about the life history of this bird and the

[1]Published in: Conservation Genetic Resources 2 (1): 151-153.

extend of genetic exchange between the bird's isolated populations (Behrstock et al. 1997; Canales-Delgadillo et al. 2008; Scott-Morales et al. 2008). Given the tremendous reduction (≈70%) of its former range and the ongoing loss of suitable habitat due to conversion into agricultural land, Worthen's Sparrow is considered endangered (IUCN 2009). To learn more about the structure of and the gene flow between the few remaining populations of this New World sparrow, we developed nine microsatellite markers for use in population genetic studies.

We took blood samples of Worthen’s Sparrows that we caught during their breeding season (May to August) at La Region del Tokio near Saltillo City (Mexico). We collected 0.1 mL blood per individual and mixed it with 0.5 mL of blood preservation buffer (Seutin et al. 1991). DNA was extracted from the blood samples using Gentra Puregene Tissue Kit (Quiagen, Maryland, USA) following the manufacturer's recommendations. Quality and quantity of the extracted DNA was assessed with a NanoDrop ND-1000 Spectrophotometer (PQLab Biotechnology, Erlangen, Germany). We constructed and screened a DNA library enriched for microsatellite sequences from the DNA of a single sparrow female following the protocol described by Glenn and Schable (2005) and applying the specifications given by Niehuis and Korb (2010), except that we digested the genomic DNA only with the restriction enzyme *RsaI* (New England Biolabs, Frankfurt / Main, Germany).

The DNA library sequences were assembled and searched for dinucleotide repeats with BioEdit v7.0.9 (Hall 1999). DNA fragments with dinucleotide repeats were additionally checked for internal *RsaI* restriction sites because they could indicate an artificial ligation of different DNA fragments (Niehuis and Korb 2010). Finally, we

used the program Primer3 v.0.4 (Rozen and Skaletsky 2000) to design locus-specific oligonuclotide primers.

To assess polymorphism of the microsatellite loci, we genotyped samples from 32 Worthen's sparrows. The PCR amplifications were conducted in 20μL volumes with 1× *Taq* incubation buffer, 0.2 mM of each dNTP, 0.5 μM each of forward (5'-labeled with the dye TAM, HEX or FAM) and reverse primer, 1.5 mM $MgCl_2$ (except for locus Sw66 for which we used a $MgCl_2$ concentration of 3.0 mM), 0.25 U *Taq* polymerase (MP Qbiogene, California, USA) and 50-100 ng DNA. The PCR temperature profile started with an initial denaturation step at 94° C for 3 min, followed by 30 cycles of 94° C for 1 min, 55° C for 1 min and 72° C for 1 min, and ended with a final extension step of 72° C for 10 min. All PCR products were separated with an ABI 377 automated sequencer (Applied Biosystems, Foster City, USA) on a 5.25% Page Plus gel (Amresco, Solon, OH, USA) with a 36 cm well-to-read distance. GeneScan 500 ROX Size Standard (Applied Biosystems) was applied as reference to assess the size of the PCR products with the aid of the Peak Scanner Software v1.0 (Applied Biosystems).

Of the 22 microsatellite loci that we tested, nine proved to amplify reliably and were polymorphic, with 3-31 alleles per locus (Table 1). Using Micro-Checker version 2.2.3 (Van Oosterhout et al. 2004), we estimated a null allele frequency of 0.00-0.05 (Table 1). The observed (H_O: 0.43-1.00) and expected (H_E: 0.38-0.95) heterozygosity (Excoffier et al. 2005) did no differ significantly in the nine markers when applying a sequential Bonferroni correction and a table-wide significance level of 0.05 (Rice 1989). We did not find evidence for linkage disequilibrium between the nine markers

(Excoffier et al. 2005; sequential Bonferroni correction at a table-wide significance level of 0.05).

The new microsatellite markers will help significantly to study the elusive life history of the Worthen's Sparrow and probably related passerines. They will in particular allow estimation of critical population genetic parameters, such as genetic diversity and population structure, important for efficient conservation efforts of this endangered bird and other related species.

Acknowledgments

This research was supported by CONACYT research project no. 90647 and Institute of Behavioral Biology at University of Osnabrück. C.D. thanks the CONACYT and the DAAD for a PhD stipend. C.D and S.M are particularly indebted to M. Cotera-Correa for the partial financial support during the field trips. O.N. acknowledges the Alexander von Humboldt Foundation for a Feodor Lynen stipend for postdoctoral researchers. S.M. acknowledges the support of DAAD (A/08/03895), DFG-CONACyT (62051103) and CONACyT (91922) during a visit to the Institute of Behavioral Biology, University of Osnabück.

Table 1. Characterization of nine microsatellite loci isolated from Worthen's Sparrow and genotypes in 32 samples of this species.

Clone/ Locus	Motif	Primer sequence 5' - 3'	Size (bp)	T_a	N	Allelic size range (bp)	Null allele freq	H_O	H_E	EMBL accession no.
Sw09	$(GT)_{15}$	F: CGCAGATGTGTTTGTGTGTG R: TGCCAGCCTAACTTTGAGTGT	133	55	7	109-134	0.05	0.687	0.788	FN652305
Sw22	$(GT)_3(TG)_7$	F: CCGTGTATGCGTGTATTTGG R: AGTGCACCGATCACACAATG	127	55	3	129-135	0.00	0.5	0.388	FN652306
Sw54	$(GT)_6GG(GT)_5AT(GT)_2$	F: CGGACACTGTATCGGGAGTT R: ATAGCGACCTGGTGTCCAGA	151	55	5	133-160	0.00	0.437	0.438	FN652307
Sw58	$(CT)_{10}$	F: TGGCCCTGTCATTACATTCA R: CTGGGCTGGGATAAGAACAA	265	55	8	178-273	0.00	0.468	0.452	FN652308
Sw62	$(GT)_{22}$	F: CTTTGGGCTGAATGCACTTT R: GAGCGCACACGCAATATAGA	132	55	13	109-140	0.00	0.875	0.877	FN652309
Sw65	$(GT)_{21}$	F: GCTTTCAAACACATGCCAGA R: GCCTTCCCAGCTCTTACCTT	126	55	31	108-157	0.04	0.937	0.95	FN652310
Sw66	$(AC)_{25}$	F: CGGCTTTTGCTTTAGGATTG R: TGTCTGGTCTCCTCACACGA	289	55	11	255-291	0.00	0.718	0.661	FN652311
Sw71	$(AC)_{20}GG(CA)_4$	F: AGTGTTGGAGGCCTTTCCTT R: TGTGCTGGATCTGCAGGAAT	128	55	18	124-172	0.00	1.000	0.935	FN652312
Sw75	$(CT)_{12}GT(CT)_7$	F: GATCAAGAGAGCTGGGATGC R: AGCAGACAGTTTCCCTCACC	273	55	8	246-288	0.01	0.687	0.723	FN652313

F: forward primer; R: reverse primer; T_a: annealing temperature in ° C; N: number of alleles; H_O: observed heterozygosity; H_E: expected heterozygosity.

Chapter 2

The influence of habitat fragmentation on genetic diversity of a rare bird species that commonly faces environmental fluctuations[2]

J. Canales-Delgadillo, L. Scott-Morales & J. Korb

Abstract

Habitat fragmentation is one of the main causes of biodiversity loss. Rare species are generally thought to be more sensitive to habitat fragmentation than common ones as small populations become even smaller. We did a population genetic study on a rare bird, the Worthen's sparrow (*Spizella wortheni*) which is endemic to semi-arid and arid regions of Northeast Mexico. Its population numbers suffer greatly from the transformation of grassland into farmland that leads to a patchy distribution with locally small population sizes. Our data show that its genetic diversity is nevertheless high, few to no differentiation between study localities was found, and gene flow was high. Although we can not exclude that is too early to see an impact on the genetic level, we think that these results might be explained by the species' biology: like many others birds living in arid areas, the Worthen's sparrow has a nomadic life style; depending on local conditions individuals flexibly move between areas. This behavior could enhance their ability to find suitable habitat patches in a fragmented landscape. Our results imply that nomadic behavior, which is an adaptation to high temporal variability in environmental conditions, may make species more resilient to

[2]Submitted to: The Journal of Avian Biology

spatial variability caused by habitat fragmentation. This insight contributes to identifying common factors such as nomadism that predict a species' sensitivity to habitat fragmentation.

Introduction

Rare species, as a rule, have higher extinctions risk that their common counterparts (Fagan et al. 2002) and this is why they are in the focus of conservation studies. It is generally accepted that the fewer occurrences a species has or the more fragmented its distribution is, the higher is its extinction risk (e.g. Frankham 1995; but see also Gaston 1994, Mace and Kershaw 1997). One common factor associated with extinction risk in rare bird species is habitat fragmentation, especially in habitat-specialists or small sized birds (Owens and Bennett 2000).

Habitat fragmentation affects the distribution of habitats across the landscape with direct impact on natural populations. It transforms formerly continuous habitats into subdivided units with varying degrees of connectivity (Freeland 2005). During the last fifty years increased habitat fragmentation and transformation of natural habitats into farmland became the principal cause of bird populations' decline (Peterjohn 2003). This often increases species' extinction risks through declining genetic diversity, especially when fragments are isolated (e.g.: Donovan et al. 1995, Díaz et al. 2006). In the long term, habitat fragmentation causes reduced effective population sizes as inbreeding within the isolated subpopulations and genetic drift increases (Croteau et al. 2007). But this is not always the case: migrant and highly mobile bird species apparently can cope with the negative effects of habitat fragmentation without presenting loss of genetic diversity (Lindsay et al. 2008). This strongly contrasts with

non-migrant sedentary, dispersal-restricted habitat specialists (Segelbacher et al. 2008). Nomadic bird species from arid and semi-arid regions are known to deal with patchily distributed resources by moving continuously around (Dean 2004). Yet little is known whether this affects their ability to cope with anthropogenic fragmentation.

We did a population genetic study based on the analysis of microsatellite markers using as model a rare endemic species, the Worthen's sparrow (*Spizella wortheni*). The Worthen's sparrow is a non-migrant emberizid inhabiting grassland-desert shrub habitats in the semi-arid region of Northeast Mexico (Behrstock et al. 1997, Wege et al. 1993). A decline of its population range has been reported with local extinctions in five of eight Mexican states (Wege et al. 1993) and habitat fragmentation is occurring at an alarming rate: from 2001 to 2005 farmland areas increased from about 12 300 ha to 22 000 ha (Canales-Delgadillo, unpubl. data). Correspondingly, the species natural habitat, xerophilous shrub and grasslands, declined by more than 30 % (SEMARNAT 2008). This is why the Worthen's sparrow became a red list species of the International Union for Nature Conservation in 1994 (BirdLife International 2008). Population genetic studies are, however, still missing that elucidate the actual impact of habitat fragmentation on the remnant Worthen's sparrow populations. This species also shows nomadic behavior which allows us to test the impact of habitat fragmentation on genetic parameters in a small bird population restricted to arid and semi-arid regions. Our specific objectives were (i) to determine its genetic diversity and variability in seven localities, (ii) to estimate migration rates between localities, and (iii) to test for isolation by distance effects.

Methods

Study area and blood sampling

The Worthen's sparrow is restricted to open arid shrub-grasslands within the Mexican High Plateau. Its record sites are intermountain valley grassland associated with mesquite/juniper or pine/oak vegetation at 1200 to 2500 m above sea level in Mexico's subtropics (Scott-Morales et al. 2008).

We searched for Worthen's sparrow at seven sampling localities: south of Saltillo Coahuila (Angeles, Perforadora, Frayle and Guerrero), west of Galeana Nuevo León (Esperanzas, Soledad) and north of San Luis Potosí (Manantial) during June to late August in 2007 and 2008 (Fig. 1). In all localities the sampling effort was identical and the number of caught birds is an indicator of population size. Pair-wise geographic distances between localities were calculated with QUANTUM GIS 1.4.0. They ranged from 10 km (Angeles – Perforadora) to 121 km (Guerrero – Manatial) (Fig. 1). Details of the characteristics of these localities, except for Manantial and Frayle, are found in Scott-Morales et al. (2008).

Manantial is located in an arid region while our remaining localities are situated in semi-arid areas. In Manantial we searched for Worthen's sparrow in an area of approximately 1400 ha which includes short grassland patches surrounded by vegetation characterized by the presence of mixed patches of tarbush (*Fluorensia cernua*) and creosote bush (*Larrea tridentata*). We found several singing males showing territorial behavior and we caught one female with a brood patch.

Frayle is located at the western extreme of our study area. The vegetation consists of short grassland mixed with microphilous desert shrub and dispersed patches

of tarbush where the Worthen's sparrow nests. In Frayle we sampled within an area of approximately 600 ha and found singing and territorial males.

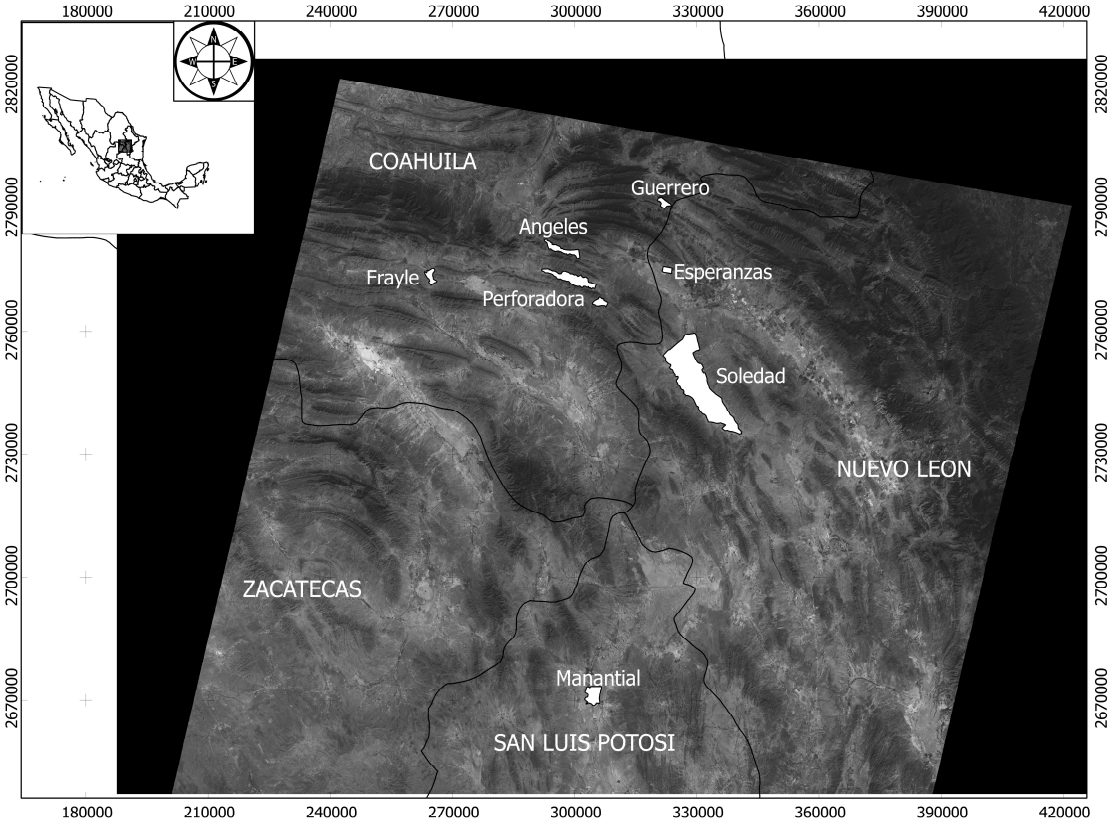

Figure 1. Satellite image of the currently known distribution range of the Worthen's Sparrow indicating our sampling localities. Map units are meters and coordinate system is UTM Zone 14N.

For DNA genotyping, we caught Worthen's sparrows (*N*= 100) by mist netting and took 0.1ml blood from the jugular vein of each individual. Blood samples were mixed with blood preservation buffer (Seutin et al. 1991) and stored until analysis. As recommended for monomorphic passerine birds sampled during the breeding season (Ralph et al. 1993), all caught individuals were sexed by looking for a brood patch and examining the cloacal region. Sample sizes varied between localities, but this reflects natural variation in population size.

DNA extraction

DNA was extracted using the Gentra Puregene Tissue Kit (Qiagen, Maryland, USA) according to manufacturer's instructions. Quality and quantity of the extracted DNA was assessed with a NanoDrop ND-1000 Spectrophotometer (PeqLab Biotechnology, Erlangen, Germany).

Microsatellite analysis

We selected four microsatellite loci developed for other bird species (Table 1) and nine novel species-specific microsatellite loci developed for Worthen's sparrow (Canales-Delgadillo et al. 2010) to genotype our samples. To score polymorphism, all forward primers were dye-labeled at the 5' end with TAMRA, HEX or FAM dyes. Polymerase chain reaction (pcr) was optimized for each primer pair. In 20µl reactions we used 1x pcr buffer (Fermentas, St. Leon-Rot, Germany), 1.5mM $MgCl_2$, 0.2 mM of each dNTP's, 0.5 µM of each primer, 0.25 U *Taq* polymerase, and 50-100 ng DNA template. The cycling protocol consisted of an initial denaturation step at 94°C for 3 min, followed by 30 cycles of denaturation at 94 °C for 1 min, locus-specific annealing temperature (Table 1) for 1 min and extension at 72 °C for 1 min, followed by a final extension step of 72 °C for 5 min. Temperature gradient tests showed that it was possible to use the same annealing temperature for all cross-species primer pairs (52.5°C), however the temperature in the extension step was lower for Gf05 and Gf12 (65°C) than for all other primers. The pcr profiles for the species-specific microsatellite loci were similar but with a longer final extension time (10 min) (see Canales-Delgadillo et al. 2010).

The pcr products were screened on 1.5% agarose gels and visualized under UV-light after staining with ethidium bromide solution (1 mg/ml). Those pcr products showing reliable amplification were separated in an ABI-377 automated sequencer (Applied Biosystems, Foster City, USA) on a 5.25% Page Plus gel (Amresco, Solon, OH, USA) run for three hours with a 36 cm well-to-read distance. GeneScan 500 ROX Size Standard (Applied Biosystems; Carlsbad CA, USA) was used as reference to assess the sizes of the pcr products with Peak Scanner Software v1.0 (Applied Biosystems).

Table 1. Cross-species and species-specific microsatellite primers used in pcr amplification reactions.

Clone/ Locus	Motif	Primer sequence 5' - 3'	Size (bp)	T_a	Accession no.
Mme8†	$(TG)_3TC(TG)_{13}$	F: TCATGGAGATGGGTGAATGCC R: TGAATCAGCAGCACACACAACC	302	52.5	AF127382
Asμ15‡	$(GT)_{19}$	F: AATAGATTCAGGTGCTTTTTC R: TAGCACATGTTGGTTTTTG	195	52.5	AY172993
Gf05§	$(AC)_{14}$	F: AAACACTGGGAGTGAAGTCT R: AACTATTCTGTGATCCTGTTACAC	218	52.5	AF081929
Gf12§	$(AC)_{17}$	F: AATCCTTCTCGTCCCTCTTGG R: TTTGAGTGTGCAGCAGTTGG	193	52.5	AF081936
Sw09ǂ	$(GT)_{15}$	F: CGCAGATGTGTTTGTGTGTG R: TGCCAGCCTAACTTTGAGTGT	133	55.0	FN652305
Sw22ǂ	$(GT)_3(TG)_7$	F: CCGTGTATGCGTGTATTTGG R: AGTGCACCGATCACACAATG	127	55.0	FN652306
Sw54ǂ	$(GT)_6GG(GT)_5AT(GT)_2$	F: CGGACACTGTATCGGGAGTT R: ATAGCGACCTGGTGTCCAGA	151	55.0	FN652307
Sw58ǂ	$(CT)_{10}$	F: TGGCCCTGTCATTACATTCA R: CTGGGCTGGGATAAGAACAA	265	55.0	FN652308
Sw62ǂ	$(GT)_{22}$	F: CTTTGGGCTGAATGCACTTT R: GAGCGCACACGCAATATAGA	132	55.0	FN652309
Sw65ǂ	$(GT)_{21}$	F: GCTTTCAAACACATGCCAGA R: GCCTTCCCAGCTCTTACCTT	126	55.0	FN652310
Sw66ǂ	$(AC)_{25}$	F: CGGCTTTTGCTTTAGGATTG R: TGTCTGGTCTCCTCACACGA	289	55.0	FN652311
Sw71ǂ	$(AC)_{20}GG(CA)_4$	F: AGTGTTGGAGGCCTTTCCTT R: TGTGCTGGATCTGCAGGAAT	128	55.0	FN652312
Sw75ǂ	$(CT)_{12}GT(CT)_7$	F: GATCAAGAGAGCTGGGATGC R: AGCAGACAGTTTCCCTCACC	273	55.0	FN652313

Source: Jeffery et al. 2001 (†); Bulgin et al. 2003 (‡); Petren 1998 (§); Canales-Delgadillo et al. 2010 (ǂ).

Genetic diversity

To test for Hardy-Weinberg Equilibrium (HWE), linkage disequilibrium (LD) and Nei's genetic diversity indices we used Arlequin 3.5 (Excoffier and Lischer, 2010). For each population and locus, we determined the number of alleles and estimated observed (H_o) and expected (H_e) heterozygosity according to Guo and Thomson (1992). We tested for null alleles using Micro-checker 2.2.3 (Van Oosterhout et al. 2004).

Population structure

We determined the partitioning of genetic variance between different hierarchical levels of the population structure by means of analysis of molecular variance (AMOVA). Genetic distances between localities (R_{ST}) were measured using the sum of squared size difference approach (Slatkin 1995) because it is demonstrated that this method is more appropriated for microsatellite loci, which usually show a mutation rate several orders of magnitude higher than the frequency of point mutations derived from other markers (Lowe et al. 2008). Gene flow was computed as the absolute number of migrants exchanged per generation (Nm) between localities; these calculations were carried out using Arlequin 3.5. For identification of immigrants and potential source populations we used IMMANC, a software designed to identify immigrants or individuals having recent immigrant ancestry, by means of the posterior probability density of the allele frequencies (Rannala and Mountain 1997). Isolation by distance (IBD) was computed with GENALEX 6.3 (Peakall and Smouse 2006) using a Mantel's test for correlation of a geographical distance matrix with a matrix of Nei's genetic distances between localities.

Two Bayesian approaches were used to infer the number of independent genetic clusters (*K*) in our data: GENELAND 3.1.5 (Guillot et al. 2005a) and STRUCTURE 2.3.1 (Pritchard et al. 2000). These methods find genetic discontinuities across the landscape between populations that are not clearly separated or if there is no a priori knowledge about population boundaries. The programs assign individuals to genetic groups. One characteristic of GENELAND is that it allows the incorporation of spatial coordinates in the algorithm of inference while STRUCTURE works just using genotypes. Because it was recently demonstrated that clusters found by STRUCTURE can be influenced by variations in sample size (Kalinowsky 2010), resulting in an artificial grouping of individuals from the largest samples, we will concentrate in the main text on the results obtained with GENELAND. Detailed methods and results of STRUCTURE are available as supplementary material.

GENELAND simulates the posterior distribution of individual georeferenced genotypes using MCMC techniques to infer the number of clusters, *K*, in a data set using a spatial statistical model and the allele frequencies (Guillot et al. 2005a, 2005b). For the inference of *K* with GENELAND we selected the correlated frequency model. Statistically, this model takes into account gene flow between populations as a correlation. We assumed correlated allele frequencies between the study sites according to our preliminary gene flow estimations (see above).

In a first step we allowed *K* to vary between 1 and 10. Because some of the genotypes shared the same spatial position, an uncertainty of 70 m was attached to the spatial coordinates. This value was selected because we observed females moving between male territories which were separated on average by this distance. In GENELAND, attaching uncertainty allows assigning individuals sampled in the same

location to different populations in case they were migrants (Guillot et al. 2005a). We used 10^6 iterations in the MCMC and ran ten replications for each value of *K*.

In a second step we performed fifty replicates with *K* fixed at the modal value obtained in the first calculation. For the assignment of individuals to genetic groups a burn-in period of 4×10^5 steps followed by 1.6×10^6 iterations in the MCMC was selected leaving all other parameters unchanged.

Data analyses

After an exploratory analysis, locus Gf05 was excluded because there was significant evidence for linkage with locus Sw09 ($p < 0.0001$) after false discovery rate (FDR) adjustment (Benjamini and Hochberg 1995). Spearman's rank correlations were applied to test whether gene diversity or inbreeding indices correlated with sample size per population. We used the MCMC (9×10^5 steps in the chain after 10^5 dememorisations) approach recommended by Excoffier and Slatkin (1998) to test for significant deviations from HWE. Exact tests using a likelihood ratio approach with 10^5 permutations were used to assess LD between pairs of loci. The *p*-values of both HWE and LD tests were adjusted for multiple testing by using FDR approach.

Immigrants were recognized according to the posterior probability ratio test ($\ln\Lambda$) as implemented in IMMANC. Individuals with positive $\ln\Lambda$ values were identified as non-migrants; whereas individuals showing negative $\ln\Lambda$ values were identified as immigrants or descendant from immigrants.

For the inference of *K*, we calculated the posterior probability for each of the 50 replicates with fixed *K* and selected the 10 runs with the highest values (Guillot et al. 2005a). Maps of the posterior probability of population membership were computed in a

spatial domain of a total of 67 km x 126 km (corresponding to a cell size resolution of 200 m) for each of these runs using a burn-in of 1.6×10^5 iterations.

All tests were two-tailed and the significance level was set to 0.05. Mean values ± standard deviations are given, if not noted otherwise.

Results

Genetic diversity

The mean number of alleles per locus over all populations was 6.0 ± 0.40, ranging from 1 to 23 (Table 2). The genetic diversity over all loci ranged from 0.53 in Angeles to 0.72 in Perforadora (Table 2). There was a trend that the observed gene diversity positively correlated with sample size, reflecting population size (Spearman rank correlation: *rho* = 0.75, *p* = 0.06). Three loci were monomorphic for the population Angeles possibly due to the small sample size (*n* = 3 individuals). We found significant deviation from HWE for some populations for both cross-species and species-specific markers (Table 2). Deviations from the HWE may, for instance, be due to small population sizes, non-random mating or presence of null alleles (Freeland 2005, Lowe et al. 2004). To discard the possibility that HWE departures were an artifact of non-amplifying loci, we re-analyzed all data excluding all markers with evidence of null alleles in at least one locality (Mme8, Asμ15, Gf12, Sw09). These analyses did not change our results and HWE departures remained.

Population structure

The results of the AMOVA did not differ qualitatively when we used all 12 loci or when excluding those loci with evidence for null alleles. Thus we present the 12 loci-based

results. Our analysis showed that the most substantial variation stemmed from within-individual variation (77%) and some from differences among individuals within localities (19%). The least amount (4%) is due to differences between sampled localities. In line with this the global fixation index suggested that there is no significant genetic differentiation between localities (R_{ST} = 0.03, p = 0.113) but overall a high degree of inbreeding (R_{IS} = 0.19, $p < 0.001$). However, pairwise comparisons between localities revealed significant differences in R_{ST} values, ranging from 0.04 to 0.21 (Table 3). In our analysis no evidence for IBD was found (Mantel's test: r^2 = 0.11, p = 0.170) indicating continuous genetic flow.

Locality specific R_{IS} values revealed that Manantial, Perforadora and Soledad had the highest and significant levels of inbreeding, while Angeles, Esperanzas and Frayle, the localities with the smallest sample sizes, showed very low, non-significant levels of inbreeding (Table 3). The observed R_{IS} value increased significantly with sample size (Spearman rank correlation: *rho* = 0.82, p = 0.034). Since we invested the same sampling effort in each locality, sample size is a measure of population size and our result likely reflects the real level of inbreeding.

The number of migrants per generation, *Nm*, among all populations was high accounting to 4.7 individuals after correction for different sample sizes (see Barton and Slatkin 1986, Slatkin and Barton 1989). The *Nm* between Esperanzas, Perforadora and Soledad were going to infinity which indicates that these localities should be considered as a single population (Table 3).

Sixteen out of one hundred individuals showed significant evidence for immigration or a significant probability to be a descendant of immigrants (not showed).

Table 2. Genetic characterization of the Worthen's Sparrow populations at twelve microsatellite loci.

	Angeles ([a]*n* = 3)				Esperanzas (*n* = 9)				Frayle (*n* = 6)				Guerrero (*n* = 11)				Manantial (*n* = 16)				Perforadora (*n* = 22)				Soledad (*n* = 33)			
	[b]GD: 0.53 ± 0.330				GD: 0.64 ± 0.350				GD: 0.67 ± 0.370				GD: 0.62 ± 0.330				GD: 0.71 ± 0.370				GD: 0.72 ± 0.370				GD: 0.69 ± 0.350			
Locus	[c]N_a	[d]H_o	[e]H_e	[f]p	N_a	H_o	H_e	p	N_a	H_o	H_e	p	N_a	H_o	H_e	p	N_a	H_o	H_e	p	N_a	H_o	H_e	p	N_a	H_o	H_e	p
Mme8	1	monomorphic			3	0.22	0.22	1.00	2	0.17	0.53	0.60	4	0.27	0.52	0.07	5	0.44	0.62	0.01	5	0.27	0.52	0.01	4	0.39	0.41	0.78
Asμ15	4	1.00	0.87	0.46	6	0.67	0.80	0.24	5	0.17	0.80	0.01	8	0.36	0.87	0.001	8	0.31	0.78	0.001	9	0.41	0.74	0.001	9	0.79	0.65	0.001
Gf12	1	monomorphic			4	0.67	0.76	0.24	3	0.67	0.55	1.00	3	0.64	0.50	1.00	6	0.38	0.73	0.01	6	0.41	0.69	0.001	5	0.64	0.71	0.001
Sw09	4	1.00	0.87	1.00	5	0.89	0.78	1.00	4	0.33	0.77	0.38	8	0.64	0.80	0.46	4	0.44	0.72	0.06	7	0.59	0.78	0.33	9	0.70	0.81	0.21
Sw22	2	0.67	0.53	1.00	2	0.67	0.47	0.78	2	0.50	0.41	1.00	2	0.55	0.42	0.86	2	0.19	0.18	1.00	4	0.59	0.45	0.51	2	0.21	0.24	0.65
Sw54	1	monomorphic			3	0.11	0.22	0.23	3	0.67	0.62	1.00	2	0.09	0.09	1.00	6	1.00	0.71	0.01	5	0.77	0.60	0.51	5	0.67	0.52	0.50
Sw58	1	monomorphic			2	0.00	0.21	0.23	4	1.00	0.73	1.00	4	0.64	0.52	1.00	7	1.00	0.79	0.14	8	0.73	0.78	0.01	7	0.79	0.76	0.01
Sw62	5	1.00	0.93	1.00	8	0.89	0.87	1.00	6	1.00	0.86	1.00	8	0.91	0.85	0.56	10	1.00	0.87	0.39	10	0.91	0.87	0.02	12	0.94	0.83	0.88
Sw65	4	1.00	0.87	1.00	10	0.67	0.85	0.24	6	0.83	0.88	1.00	8	0.55	0.79	0.14	13	1.00	0.89	0.01	9	1.00	0.85	0.33	15	0.97	0.88	0.78
Sw66	2	1.00	0.60	0.40	8	1.00	0.86	0.93	2	0.33	0.30	1.00	5	0.64	0.53	1.00	6	0.94	0.75	0.42	6	0.86	0.81	0.59	9	0.85	0.76	0.01
Sw71	6	1.00	1.00	1.00	11	1.00	0.94	0.84	7	1.00	0.91	1.00	11	1.00	0.93	0.36	12	1.00	0.92	0.58	14	1.00	0.91	0.59	23	1.00	0.94	0.06
Sw75	3	0.67	0.73	1.00	4	0.89	0.73	0.14	4	1.00	0.71	1.00	6	0.91	0.69	1.00	4	0.75	0.66	1.00	7	0.77	0.68	0.45	11	0.79	0.79	0.001

[a]Sample size, [b]Average genetic diversity over loci ± standard deviation, [c]Number of alleles per locus, [d]Observed heterozygosity, [e]Expected heterozygosity, [f]Significance level of deviations from HWE expectations

Two individuals appear to have older immigrant ancestry, since their parents (with positive values of lnΛ) were born in the same population where they were sampled, while their grandparents and more distant ancestors were classified as immigrants from other localities. For the fourteen remaining individuals there was evidence of recent immigrant ancestry, all born from immigrant parents. Angeles was the population from which most immigrants or descendants from immigrants were identified which were going to Perforadora and Soledad. No individuals belonging to the latter localities appear to have descendants within other populations.

Table 3. Pair-wise genetic distances between localities. R_{ST} values are given below diagonal, numbers in bold are significant at $p \leq 0.05$, above diagonal *Nm* exchanged between pairs of localities. R_{IS} values and their significant level are given in the last two columns.

	Angeles	Esperanzas	Frayle	Guerrero	Manantial	Perforadora	Soledad	R_{IS}	*p*-value
Angeles	*	3.75	2.86	2.71	3.87	5.45	5.93	-0.01	0.60
Esperanzas	**0.12**	*	3.39	4.64	9.46	∞	∞	-0.08	0.68
Frayle	0.15	**0.13**	*	30.38	1.85	6.66	4.30	-0.06	0.50
Guerrero	**0.16**	**0.10**	0.02	*	2.90	46.52	8.24	0.13	0.15
Manantial	0.11	**0.05**	**0.21**	**0.15**	*	10.71	27.28	0.15	0.03
Perforadora	0.08	0.00	0.07	0.01	**0.04**	*	∞	0.32	<0.0001
Soledad	0.07	-0.01	**0.10**	**0.06**	0.02	0.00	*	0.23	<0.0001

Inference of K

GENELAND consistently estimated $K = 6$ clusters across the ten best replicates when K was allowed to vary. After re-running our data with fixed $K = 6$, four of these clusters showed a consistent well defined spatial configuration. No individuals were assigned to two of the six clusters; these clusters are ‘ghost populations’ (Guillot el al. 2005a). Such ghost populations

are an artifact that occurs during the tiling process of the spatial domain into a heterogeneous sampling distribution (Guillot el al. 2005a).Three areas appear to be isolated clusters: in the south Manantial, in the northwest Frayle and directly in the north Angeles. One cluster gathered the data from Perforadora, Esperanzas, Guerrero and Soledad as a single unit (Fig. 2).

The posterior probabilities for these clusters in five out of the ten best runs were consistently above 0.55. Four of the spatial cluster configurations in these runs showed an inconsistent pattern with posterior probabilities less than 0.5. We run the data fifty additional times with unchanged parameters but with fixed $K = 4$ to check for consistency in the clustering arrangement.

One of the runs with fixed $K = 4$ showed an identical spatial pattern as those described previously for Figure 2. The results with fixed $K = 4$ were also consistent in that Perforadora/Soledad and Esperanzas/Guerrero, respectively, clustered in single groups, and in the clustering of Angeles and Frayle as isolated groups. A marked difference compared to the first analysis was that Frayle clustered with Angeles, Esperanzas and Guerrero as a single group in 30% of the runs.

With STRUCTURE the LnPr(X|K) values were similar to those obtained with GENELAND ($K = 3$, Fig. S1 Suppl. Mat.). The individuals from Manantial, Esperanzas, Perforadora and Soledad were distributed across all three clusters. Only individuals from Frayle were assigned to a single cluster, while individuals from Guerrero and Angeles were assigned into two of the three groups.

Discussion

The effect of fragmentation

Despite habitat fragmentation, we found little evidence of genetic depletion or impediment of gene flow among our study localities of the Worthen's sparrow. This might be due to the fact that the impact of fragmentation is still too recent to have any effects detectable at the genetic level, as processes such as genetic drift can be slow. Studying an isolated population of black grouse (*Tetrao tetrix*), Larsson et al. (2008) found difficulties to prove a reduction in population size, which was detected only when genotype of samples from museum were used as reference. They argued that a very small effective population size could mask loss of genetic diversity. Although we cannot exclude this explanation completely, we think it is more likely that our unexpected result is due to the bird's nomadic life style. The connectivity between the remnant localities of the study area is low to moderate (30%-60%; Canales –Delgadillo, unpubl. data). Yet the regular 'travel' distances of the Worthen's sparrow are so high (> 5 km) that they still can move between these localities (Canales-Degadillo, pers obs). So, what seems in absolute terms as disconnected, isolated localities, are patches within the dispersal range of the Worthen's sparrow. These high travel distances are a result of their nomadic behavior, an adaptation to the high temporal and spatial variability in resources availability in arid and semi-arid areas.

Dean (2004) defined nomadic birds of arid and semi-arid areas as organisms whose movements are regionally unpredictable, the number of individuals observed is irregular and who can be present or absent in a locality for long or short periods of time. According to Andersson (1980), nomadism is also associated with the tendency of adults and juveniles to move in the search for food, and to settle and breed where it is locally abundant. Such patterns

are observed in the Worthen's sparrow (Behrstock et al. 1997, Scott-Morales et al. 2008, Wege et al. 1993). Nomadism in birds of arid and semi-arid areas is considered to be an efficient strategy to use highly variable and unpredictable biomass production that depends on environmental conditions, especially patchy rainfalls (Andersson 1980, Dean 1997, Dean 2004). It favors individuals moving around in the search of suitable patches and offers these birds an advantage when habitats become fragmented due to anthropogenic causes. This suggests that nomadism makes species more resilient against habitat fragmentation, even when levels of landscape connectivity are low to moderate.

In fact, the genetic diversity and the genetic differentiation between localities were similar to those reported for common, unthreatened birds (Crochet 2000 and references herein). For instance, our heterozygosity values for the Worthen's sparrow were comparable to those found for its close relative, the Brewer's sparrow (*Spizella breweri*), which is a common species in North America inhabiting natural fragmented habitats (Croteau et al. 2007). That high dispersal abilities can help to buffer against effects of severe habitat fragmentation was also shown for a tropical bird, the white-starred robin (*Pogonocichla stellata*) (Galbusera et al 2004).

Delimitation of subpopulations

The results from STRUCTURE were similar to those from GENELAND; however the degree of admixture between populations was higher in the former than the latter. We think the results of GENELAND are more realistic because: (i) the clustering pattern is in accordance with the geographical connectivity of the localities (Fig 1) and (ii) this model showed higher consistency in its output. Furthermore, calculations of R_{ST} values for the genetic clusters inferred by GENELAND (not shown) corroborated that genotypes from Manantial, Frayle and

Angeles (Fig. 2a, 2b and 2d respectively) are significantly different from the genotypes of the other localities and that they contribute to population structure.

The two ghost populations found with GENELAND could be confirmed as a tilting artifact (see section 3.3): they were located in valleys bordered by steep mountain slopes where the Worthen's sparrow does not occur, these ghost populations have no influence in the biological interpretation of the results as no individuals were assigned to them (Coulon et al. 2006, Fontaine et al. 2007, Guillot et al. 2005a, 2005b).

Inbreeding

Nomadic and gregarious bird species tend to move together across the landscape, and it is well established that moving in flocks increase the chances to find food and escape predators. However, inbreeding might result as a consequence (Dean, 1997). Szulkin and Sheldon (2008) postulated for wild populations of the great tit (*Parus major*) that, when mechanisms of kin-recognition are absent, the likelihood of brother-sister mating increases when siblings disperse in similar directions. The tendency of the Worthen's sparrow to move in gregarious groups (Canales-Delgadillo et al. 2008), might contribute to the high inbreeding levels we found by promoting some degree of non-random mating. An excess of homozygotes derived from the presence of null alleles in our data can be excluded as the inbreeding coefficients for all 12 loci (R_{IS} = 0.198, p = 0.0001) and the 8 loci, when those with evidence for null alleles were excluded, (R_{IS} = 0.265, p <0.0001), were both high and significant.

Implications for Conservation

Our findings suggest that the Worthen's sparrow current population is represented by a single unit composed of four subpopulations. We found no evidence for genetic

impoverishment and the lack of IBD further indicates that mixing commonly occurs. The fact that up to date the Worthen's sparrow seems not to be negatively affected by habitat fragmentation can be attributed to the species' nomadic behavior. Such behavior might allow species to take advantage of moderate levels of landscape connectivity.

More than 30% of the original extension of xerophilus shrub and grassland, the habitat of the Worthen's sparrow, had been lost in Mexico (SEMARNAT 2008). Therefore, conservation efforts should focus on the protection of suitable habitat patches in its current state and it should be ensured that these patches are not too far isolated to secure the viability of the, apparently, single remnant population of this bird. Accordingly, transformation of native grassland desert-shrub areas into farmland, and also other anthropogenic pressures such as cattle grazing, must be controlled. The conservation of the Worthen's sparrow's habitat will not only ensure the conservation of this species and other breeding bird species, but also of other endemic species inhabiting the shrub-grasslands in Northeast Mexico (e.g. Mexican prairie dog, see Scott-Morales et al. 2004).

Conclusion

We did no find a negative impact of habitat fragmentation on the Worthen's Sparrow. This is probably due to bird's nomadic habit which allows it to cross easily from one habitat fragment to another. Such a nomadic life style is an adaptation to the large temporal and spatial variability in suitable environmental conditions in arid and semi-arid regions. We suppose that a nomadic habit might mitigate the effects of anthropogenic habitat fragmentation. Hence, bird species from arid and semi-arid regions which are adapted to cope with variable conditions might be less prone to extinction risks.

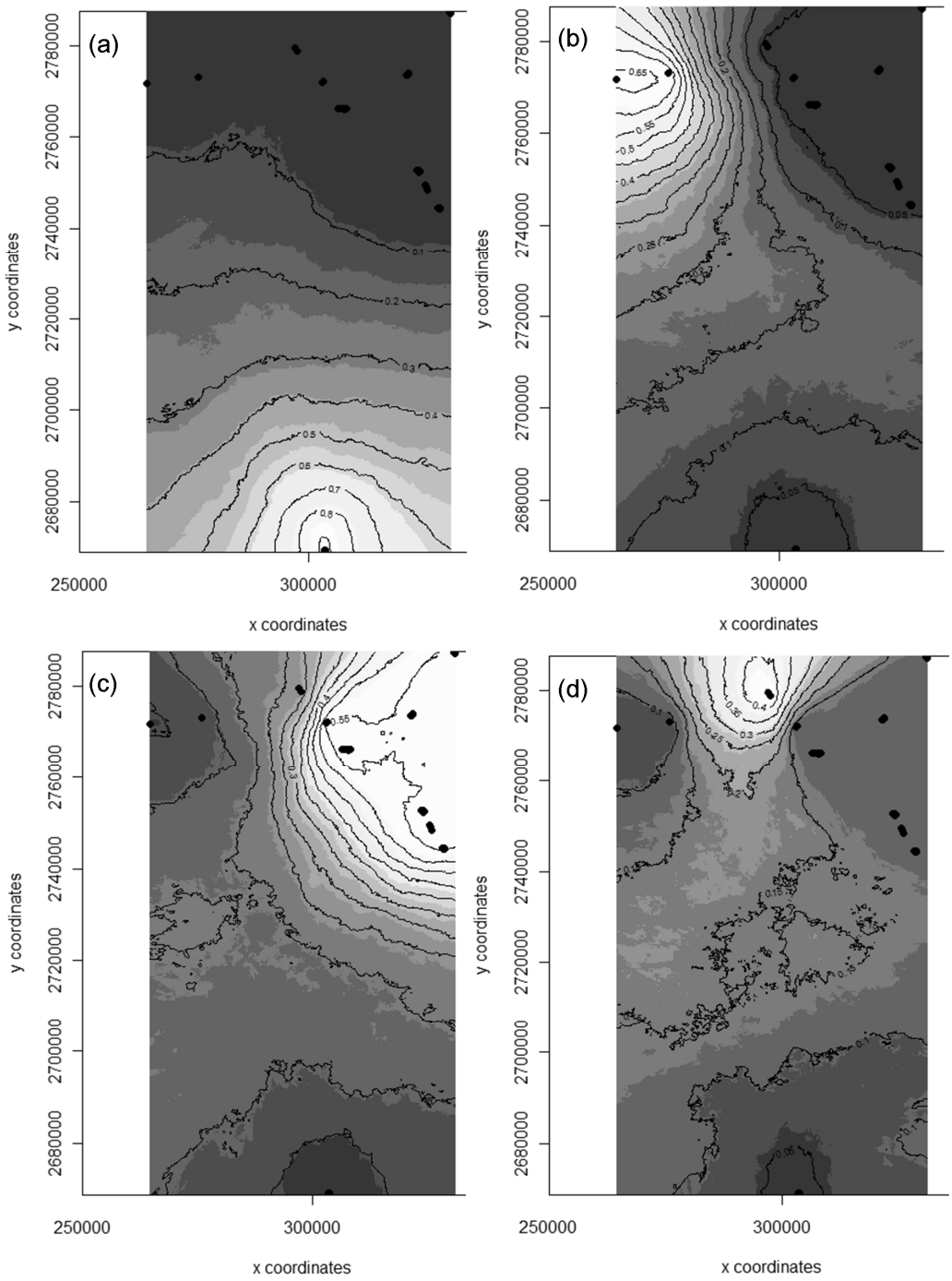

Figure 2. GENELAND's assignment of individuals to clusters (K = 4) in a spatial domain of 67 x 126 km, the brighter the color, the higher probability to belong the same cluster. Black points represent the individuals' location. In (a) individuals sampled at Manatial, (b) individuals from Frayle, (c) individuals from Perforadora, Esperanzas, Soledad and Guerrero, (d) individuals from Angeles. Graph coordinates are meters and coordinate system is UTM Zone 14N.

Chapter 3

From tree to network: conflicting phylogenetic relationships in Spizella sparrows[3]

J. Canales-Delgadillo, L. Scott-Morales & J. Korb.

Abstract

The genus *Spizella* consists of a group of small new world sparrows inhabiting in North America. Several *Spizella* species are intensively studied, yet their phylogenetic relationships are unclear and highly debated. Based on the analysis of five protein-coding genes, we present a well-resolved hypothesis on the phylogenetic relationships within this group. It explains why former studies failed at gaining unambiguous results. We amplified a fragment of 2889 bp containing segments of the mitochondrial genes COI, COIII, Atp6, Atp8 and Cyt b. We conducted Bayesian Inference and maximum likelihood analyses and complemented them with a phylogenetic network approach. Our results showed that seven *Spizella* species form a monophyletic group, yet the position of the species was ambiguous when applying Bayesian and maximum likelihood analyses. The phylogenetic relationships of the *Spizella* species were better represented by a reticulate network as reveled from network analysis. This suggests that rapid radiation and especially hybridization events in past and recent species have obscured phylogenetic relationships. This lead to the continued debates as such relationships cannot be unambiguously resolved with a bifurcating tree approach. This shows how network analyses can help in resolving difficult phylogentic reconstruction.

Keywords: Sparrows, *Spizella* wortheni, phylogenetics, hybridization

[3] Submitted to Molecular Phylogenetics and Evolution.

Introduction

The genus *Spizella* consists of a group of small new world sparrows inhabiting across different ecological regions in North America that range from the arid and semi-arid areas of the Chihuahuan Desert in Mexico through the temperate regions in the USA to boreal forests in Alaska. Several *Spizella* species are intensively studied (e.g. Nelson, 1989; Goodson, 1998; Willson, 1990; Albrech and Oring, 1995; Mahoni et al., 2006; Croteau et al., 2007), however their phylogenetic relationships are still unclear. This hampers comparisons and compromises conservation management of the rare and endangered representatives of this genus, such as *Spizella wortheni*.

During the last fifteen years several authors contributed to clarify the phylogenetic relationship of these sparrows using different molecular techniques such as restriction fragment length polymorphism (RFLP, Zink and Dittmann, 1993) or sequencing of mitochondrial genes (Dodge et al., 1995; Carson and Spicer, 2003; Canales-Del Castillo et al., 2010).

As in other passerine birds (see Jhønsson and Fjeldså; 2006), these studies came to inconsistent results, mainly due to the exclusion of one or more representatives from their data sets. Hence for instance, the species status of *Spizella taverneri* is debated: Some recognize it as an independent species (Mayr and Short, 1970; Sibley and Monroe, 1990; Klicka and Zink, 1999), while others argue that it is a subspecies of *Spizella breweri* (AOU, 1998). This lead to the exclusion of *S. taverneri* from many phylogenetic analyses.

For *S. wortheni*, the full species status is beyond doubt, but its sister taxon relationships is unclear. Because of similarities in plumage and size, *S. wortheni* and *Spizella*

pusilla were supposed to be closely related (Wege et al., 1993), and earlier work even postulated that *S. wortheni* might be a subspecies of *S. pusilla* (Burleigh and Lowery, 1942). During a bird survey in Zacatecas, Mexico, Webster and Orr (1954) conducted a revision and examined several specimens of *S. wortheni*; their data supported the hypothesis of full species status and a close relationship with *S. pusilla*. Yet *S. wortheni* was thought to be closely related to *S. pusilla* as well as to *Spizella passerina* when song characteristics were compared (Webster and Orr, 1954; Wege et al., 1993).

Recently, a mtDNA-based phylogenetic study (Canales-Del Castillo et al., 2010) included *S. wortheni* for the first time in a molecular analysis and this suggested that, despite the external morphological similarities, *S. wortheni* is more closely related to *S. breweri* than to *S. pusilla*.

Here we present an alternative hypothesis on the phylogenetic relationships within the *Spizella* group. Our study is based on the analysis of five protein-coding genes widely used for phylogenetic studies at species level in birds (Zink and Dittmann, 1993; Carson and Spicer, 2003; Klicka et al., 2007). We conducted Bayesian inference (BI) and maximum likelihood (ML) analyses and complemented them with a phylogenetic network approach. This latter is a useful tool to reveal complex, non-tree-like relationships between taxa (Huson and Bryant, 2006.). Our goal was to determine the phylogenetic relationships of *Spizella* species and specifically test for the sister taxon relationships of *S. wortheni* and *S. pusilla*. Our data set comprises for the first time all representatives of this genus, including *S. taverneri* whose species status is still debated.

Methods

Taxon sampling

Our analysis included all eight species belonging to the genus *Spizella*. We caught *S. wortheni* individuals in La Region del Tokio, located in Northeast Mexico (262782 E, 2792794 N and 339244 E, 2667572 N) between Coahulila, Nuevo Leon and San Luis Potosí in order to obtain blood samples for DNA extraction. From each specimen we took 0.1 ml fresh blood from the jugular vein and stored it in blood preservation buffer (Seutin et al. 1991) until analysis in the laboratory. Additionally, two samples of skeletal muscle of *S. taverneri* and three of *S. arborea* (voucher UAM-6669, UAM-6670, UAM-15244, UAM-10807 and LSUMZ B-41447) were obtained, respectively, from the University of Alaska Museum and from the Museum of Natural Science of the Louisiana State University.

Laboratory protocols

DNA extraction and amplification

Mitochondrial DNA was extracted from five blood samples of *S. wortheni* and from the samples of *S. taverneri* and *S. arborea* using the DNeasy Tissue kit (Qiagen, Hilden, Germany) following manufacturer's instructions. Extracted DNA was re-hydrated in 50 µl ultra pure water and the quantity and quality were assessed with a Nano-Drop 1000 spectrophotometer (PeqLab Biotechnologie, Erlangen, Germany).

For each sample, we amplified a fragment of 2889 base pairs (bp) in length containing fragments of the mitochondrial genes cytochrome oxidase subunit 1 and 3 (COI and COIII respectively), adenosine triphosphatase unit 6 and 8 (Atp6 and Atp8) and cytochrome b (Cyt-

b). Primer pairs for amplification and sequencing are given in Table A.1 (see Appendix A in supplementary materials). Due to the large size of the fragment containing the genes COIII, Atp6 and Atp8, polymerase chain reaction (PCR) amplification with internal primers specific for each gene, was carried out after an initial amplification of the large fragment following Carson and Spicer (2003) with some modifications: PCR reactions were conducted in 20 µl volumes with 1x PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM of each dNTPs, 0.5 µM of each oligonucleotide primer, 0.25 U *Taq* Polymerase and ~100 ng of DNA template. The PCR protocol consisted of a denaturation step at 94° C for one min, followed for 30 cycles at 94° C for 45 sec, 52° C for 45 sec, 72° C for 2 min, and a final extension step at 72° C for 5 min.

For Cyt-b amplification we decreased the concentration of oligonucleotide primers to 0.2 mM, of dNTP´s to 0.1 mM each, and used 2 U *Taq* Plolymerase per reaction; all other parameters remained unchanged. For COI, amplification was conducted with 1x PCR buffer, 1.3 mM $MgCl_2$ 0.1 mM of each dNTP's , 0.1 µM of oligonucleotide primer and 1 U *Taq* Polymerase per 20 µl reaction. Cycling conditions for Cyt-b and COI involved touchdown PCR , starting with 5° C above the optimum annealing temperature of the primers with decrements of 0.5° C every cycle; when the optimum temperature was reached, 30 additional cycles were done ending with a final extension step at 72° C for 5 min. For Cyt-b the initial denaturation step was changed to 2.5 min at 94° C. PCR products were screened on 1.5% agarose gel and visualized under UV light after ethidium bromide (1mg/ml) staining.

Sequencing

PCR products were cleaned with NucleoSpin Extract II Purification Kit (Macherey-Nagel) prior to cycle sequencing. For cycle sequencing, we used 10 µl reactions composed as follows: 50-100ng DNA template, 0.5-1 µl primer and 1-2 µl Big Dye Terminator v3.1

(Applied Biosystems). The cycle sequencing protocol included an initial step at 96° for one min, followed for 30 cycles of 96° for 30 sec, 50° for 15 sec and 60° for 4 min. Excess of dye terminator was removed using ethanol precipitation and samples were dried in a desiccator. DNA samples were resuspended in 12 µl formamide and sequenced in an ABI-3500 automated sequencer (Applied Biosystems).

Additionally, we obtained DNA sequences of *Spizella arborea, Spizella atrogularis, Spizella pallida, S. breweri, S. taverneri, S. passerina, S. pusilla, Junco hyemalis* and *Amphispiza bilineata* from GenBank (Table A.2, Appendix A).

Phylogenetic analysis

Several analyses including ten terminal taxa were carried out. The ingroup consisted of all *Spizella* species. As outgroup we used sequences from *J. hyemalis* which is closely related to *S. passerina*, a putative member of the *Spizella* group (Carson and Spicer, 2003); and *A. bilineata* which has been postulated to have a close relation with the members of the clade *Spizella* (Carson and Spicer, 2003).

Sequences were aligned independently for each gene and were also concatenated to form a segment of 2889 bp. Sequences alignment, calculation of base composition, base composition bias, corrected (Kimura two-parameters) and uncorrected pairwise distances among sequences were carried out using MEGA4 (Tamura et al., 2007). To test the probability that sequences evolve with the same substitution pattern, a disparity index test (Kumar and Gadakar, 2001) was conducted and significance levels were computed using a Monte Carlo test with 2500 replicates as implemented in MEGA4 (Tamura et al., 2007).

The quality of the genetic signal of the sequences was assessed with a substitution saturation test (Iss) according to Xia et al. (2003) and Xia and Lemey (2009) with DAMBE v.5.2.34 (Xia and Xie, 2001).

The phylogenetic relationships among taxa were determined with the Bayesian inference (BI) and maximum likelihood (ML) approach. jModeltest (Posada, 2008; Guindon and Gascuel, 2003) was run under the Akaike information criterion (AIC) for concatenated sequences and for each individual gene to determine the best substitution models for our data. After models testing, analyses were carried out using the general time reversible nucleotide substitution model (GTR) with among-site rate variation and with a proportion of invariable sites.

Bayesian analyses were performed with MrBayes 3.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). According to jModeltest results, data were partitioned to allow rate variation across sites for COI, ATP8 and Cyt-b sequences, while invariable rate partition was set for ATP6 and COIII sequences. A run with non-partitioned data was also conducted. Following the recommendations of Huelsenbeck and Ronquist (2001) and Ronquist and Huelsenbeck (2003), we generated a sample of $3x10^4$ trees and discarded the first 25% to ensure stability in the posterior probabilities estimated by means of the Markov Chain Monte Carlo procedure (MCMC). Then, a 50% majority-rule consensus tree was obtained using a sample size of $2.25x10^4$ trees.

The Maximum likelihood analysis was performed with PhyML v.3.0 (Guindon and Gascuel, 2003). Tree search was carried out through the subtree pruning and regrafting (SPR) approach and initiated from a BioNj tree. The nucleotide substitution models were implemented in the same way as in the BI. All parameters in the models were inferred from

the data set. To test the robustness of the branches in the final ML consensus tree, we used approximate likelihood ratio test with a Chi squared-based procedure (aLRT-Chi2). According to Guindon et al. (2010), this algorithm is as accurate as conventional bootstrapping and less time consuming.

We calibrated a molecular clock to estimate the divergence time between taxa. As recommended for Avian phylogenetic reconstruction, we assumed a relative evolutionary rate of 2% per one million years (Myr) (Klicka and Zinck, 1997). To test the molecular clock for equality in evolutionary rate between sequences, we performed a likelihood ratio test (LTR) between the -LnL scores with and without the assumption of the clock, using *A. bilineata* as outgroup.

Phylogenetic networks are widely used to show the evolutionary relationships of organisms in cases in which the data give incompatible or ambiguous genetic signals and cannot be drawn as a bifurcating tree (Huson and Bryant, 2006). Using Splits Tree 4 (Huson and Bryant 2006), we computed a split decomposition network to detect incompatible signals between sequences. Additionally, we computed a reticulate network which was rooted with *A. bilineata* as outgroup. Different to split decomposition networks, a reticulate network is able to represent organisms' evolutionary relationships when hybridization, horizontal gene transfer or recombination between species occur (Huson and Bryant, 2006). The fit value of the reticulate network was computed using the least squares method (LsFit) as implemented in Splits Tree 4, by comparing the pairwise distances in the generated graphs versus the pairwise distances in the data matrix. A bootstrapping test using 1000 replicates was conducted to determine the support for the network's edges.

Results

Sequence analysis

Genes fragments sequenced for *S. wortheni*, *S. taverneri* and *S arborea* were stored at EMBL-Nucleotide Sequence Database under accession numbers FR847849 to FR847858 (Appendix A, Table A.2). For the concatenated sequence, 689 sites were variable (23.84%) and 349 sites (12.08%) were phylogenetically informative. We did no find statistical differences between pairwise nucleotide distances calculated with corrected and uncorrected approaches (*t* test, $t_{1,39} = 0.95$, $p = 0.347$, Fig. 1), therefore we used the uncorrected values in further calculations. The proportion of invariant sites was 0.67. Base composition, base composition bias and average uncorrected distances between species are shown in Table 1.

The estimated average divergence between *Spizella* species was 6.7%. Pairwise divergence distances between species for the concatenated sequence ranged from 2.5% (*S. breweri* vs. *S. passerina*) to 7.1% (*S. passerina* vs. *S. wortheni*). There was no evidence that the different gene fragments or the concatenated fragment differed in substitution patterns (disparity index test: $p > 0.05$). No significant evidence of saturation for sequences of individual genes was found (observed Iss range = 0.100 - 0.149 < calculated Issc range = 0.606 – 0.754; $p < 0.0001$).

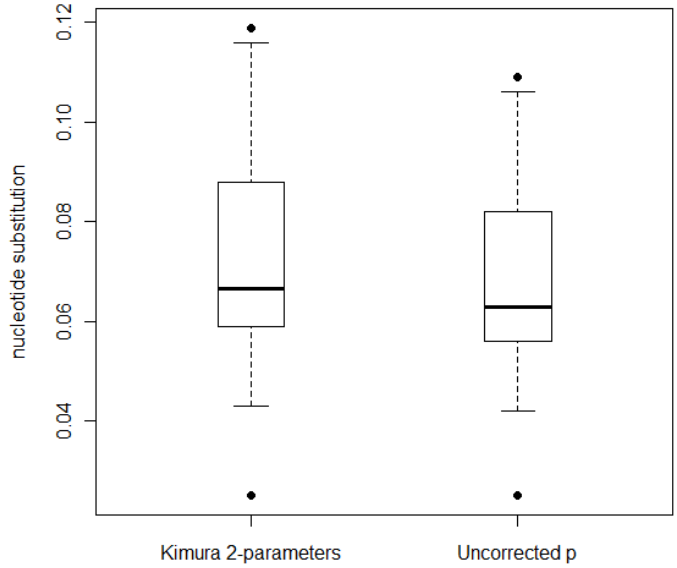

Figure 1. Base substitutions per site between sequences of *Spizella* species calculated for the concatenated sequence. The graph shows Kimura two-parameters corrected and uncorrected-p values. Shown are boxplots with median and quartile ranges, filled circles indicate minimum and maximum values.

Table 1. Number of sequenced sites per gene, base composition, base composition bias (*R*) and average uncorrected distances (*p*-) for the genus *Spizella* among haplotypes at five mitochondrial genes.

	Sites	C	A	G	T	*R*	*p*-
ATP6	402	0.381	0.286	0.109	0.224	3.418	0.030
ATP8	168	0.385	0.304	0.076	0.236	4.915	0.097
CO1	714	0.338	0.266	0.161	0.235	4.628	0.052
CO3	573	0.331	0.271	0.148	0.250	3.655	0.055
Cytb	1032	0.333	0.267	0.166	0.234	2.410	0.061
All	2889	0.346	0.274	0.142	0.238	3.308	0.073

Phylogenetic relationships.

Bayesian and maximum likelihood analyses showed that the seven *Spizella* species form a monophyletic group (Fig. 2). As previously demonstrated (Zink and Dittmann, 1993), *S. arborea* is paraphyletic to this group and seems more closely related to other passerine species.

With the Bayesian analysis, using the whole sequence and the GTR+I+G model, we obtained a well resolved topology for most nodes (99-100 %, Fig. 2a). However, one terminal node (*S. pusilla* + *S. wortheni*) and one central node showed low branch support (Fig. 2a). A second analysis, with the data partitioned and using a different evolutionary model for each gene, successfully resolved the problem of low branch support, but a three-way polytomy arose between the clade *S. pusilla* + *S. wortheni* and the complex *S. taverneri*, *S. breweri* + *S. passerina* (Figure 2b). Note that both runs were congruent in several aspects: (i) the placement of *S. pallida* and *S. atrogularis* at the most basal positions of the tree, although the topology of central nodes changed between runs, (ii) the close relationship between *S. wortheni* and *S. pusilla* which was well supported in the second run; (iii) a close relationship between *S. taverneri* and the clade formed by *S. breweri* and *S. passerina*, which were placed with robust support as sister taxa.

The maximum likelihood analysis was congruent with BI in the placement of *S. taverneri* and *S. breweri* + *S. passerina* in a well supported monophyletic clade and it also assigned *S. pallida* and *S. atrogularis* to the basal positions of the tree. As with the BI approach, *S. wortheni* and *S. pusilla* were placed as sister taxa. The ML tree also showed one unsatisfactory resolved terminal node and a relatively low support for one central node (Figure 2c).

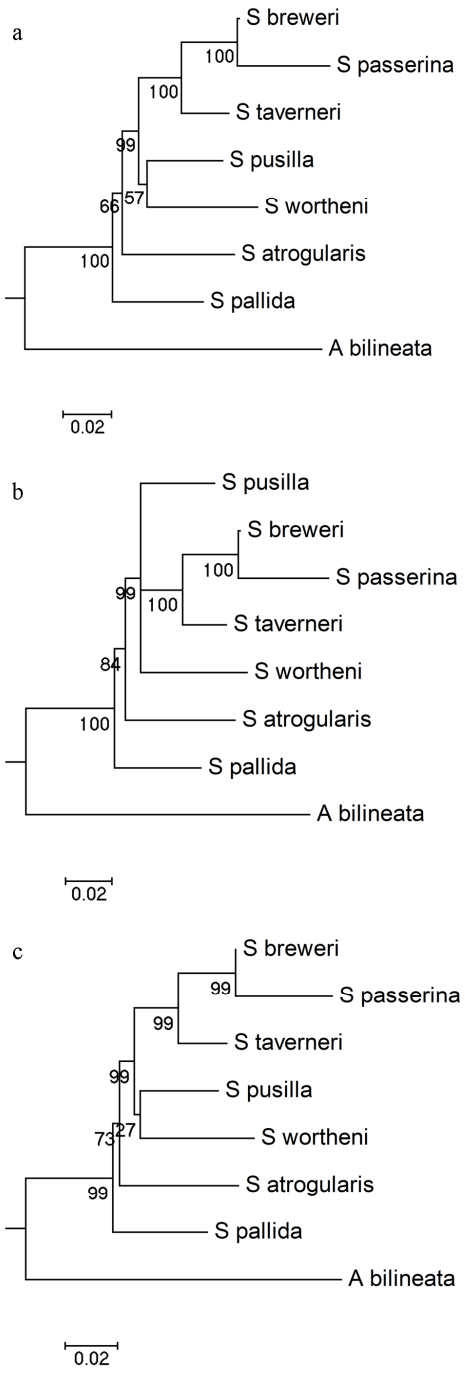

Figure 2. Topology of the phylogenetic relationships based on five mitochondrial genes. (a) Bayesian phylogram based on the concatenated sequence (nucleotide substitution model: GTR+I+G), (b) Bayesian phylogram based on partitioned data using specific nucleotide evolution models for each gene, (c) Maximum likelihood phylogram (nucleotide substitution model: GTR+I+G). Numbers at the nodes give statistical support: posterior probabilities for the Bayesian trees and Chi-squared aLRT for the Maximum likelihood tree.

Results of the molecular clock calibration suggest that *Spizella* lineages diverged approximately 2.5 - 3.3 millions years ago. However this estimation might be unreliable as the evolutionary rates of the different sequences differed significantly (Table 2) and therefore the use of a molecular clock was rejected.

Table 2. Calibration of molecular clock using maximum likelihood scores under different substitution models.

Model:	GTR	TN93	HKY85
LnL with no clock	-8327.980	-8354.345	-8357.691
LnL with clock	-8360.689	-8386.910	-8390.432
χ^2 Likelihood ratio	65.417	65.130	65.482
Df	8	8	8
P	<0.0001	<0.0001	<0.0001

Trying to resolve the inconsistencies, we added two fragments of publicly available DNA sequences of the genes NADH subunit 2 (ND2; 1030 bp) and mitochondrial control region domain I (CRI; 308 bp) (Appendix A, Table A.2). We re-run the analyses under the same conditions as above, except that we had to excluded *S. taverneri* due to a lack of molecular data for these genes. The additional sequences did not improve our results. Both the BI and the ML approach showed that *S. passerina* and *S. breweri* were more closely related to each other than to the other species; however they were not assigned to the same branch as in our previous analysis (Appendix A, Fig. A.1). Neither BI nor ML were able to resolve conclusively the phylogenetic relationships between *Spizella* members. Trees for individual genes also showed inconsistency in their topology, polytomies (ATP6, COIII and CRI) and low bootstrap support (Appendix A, Fig. A.2).

The split decomposition network reveled that most conflicting genetic signals were present between sequences of *S. taverneri*, *S. breweri* and *S. passerina* (Fig. 3a). It also found several incompatibilities at the central nodes of the network indicating that it is difficult to depict the relationships between the species in a tree-like form (Moulton and Huber, 2009). Since these results showed only the complexity of the data but did not reveal the phylogenetic relationship between taxa, we computed a reticulate network. The reticulate network supported all the pendant branches with high bootstrap values (Figure 3b).

In reticulate networks, pattern of descent are described by a putative evolutionary history in which internal nodes represent ancestral species and nodes with more than two parents correspond to reticulate events such as hybridization (Huson and Bryant, 2006).

In our analysis three cyclic reticulations existed (LsFit = 99.05) from which four extant taxa descended: (i) *S. atrogularis* and *S. pusilla* with a direct common ancestor, (ii) *S. pallida* and (iii) *S. breweri* whose immediate ancestor hybridized with the lineages of *S. taverneri* and *S. passerina* (Figure 3b); leading to a similar patter as in the phylogenetic trees (Fig. 1). Hence, inline with the network topology, hybridization events within the *Spizella* group seem have occurred among ancestral lineages.

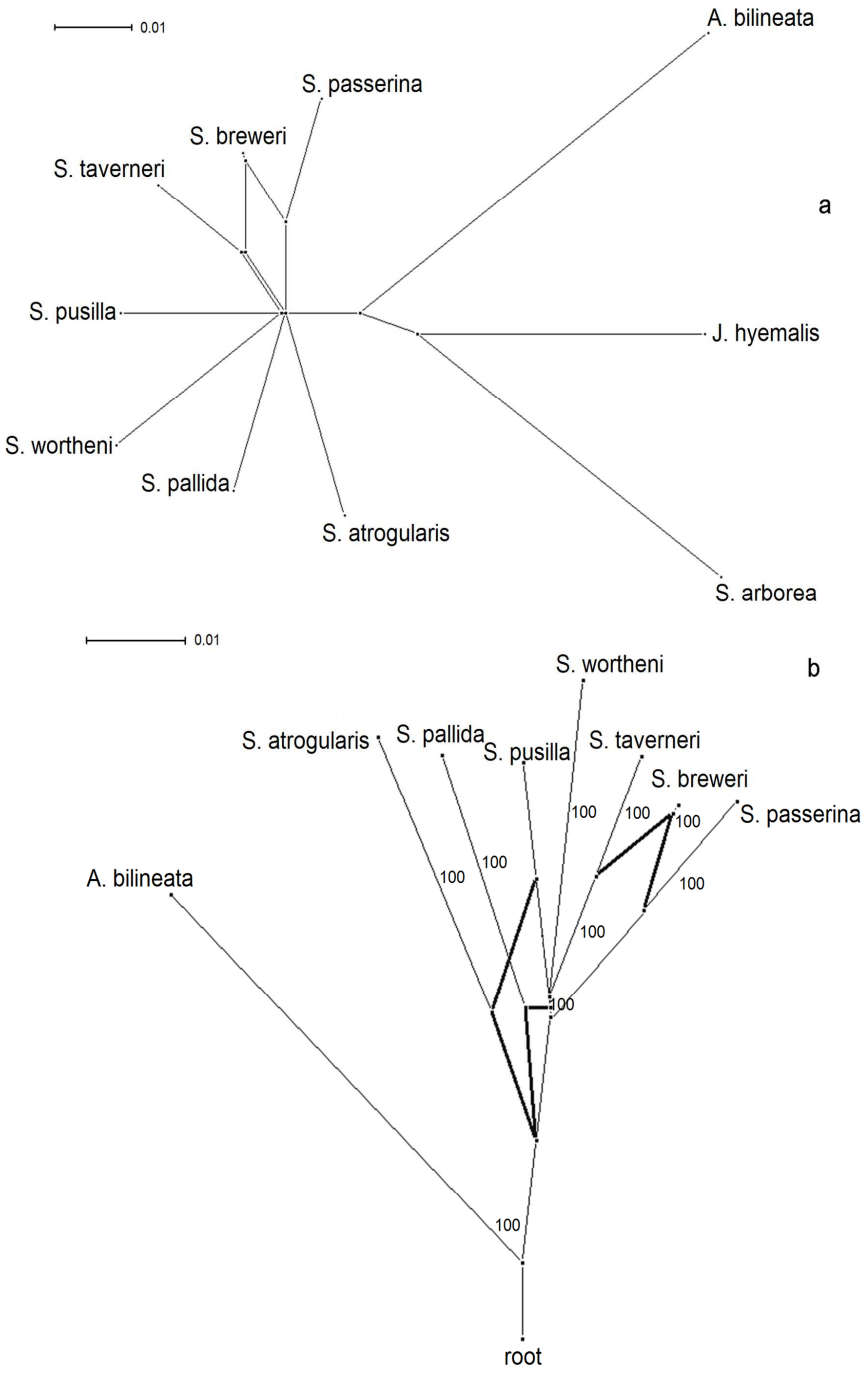

Figure 3. (a) Split decomposition network showing incompatible genetic signals (parallel lines) between central and terminal nodes. (b) Reticulate network with bootstrap values on the branches. Three reticulation events (length = 3) are indicated with thick lines.

Discussion

The phylogeny of the passerine genus *Spizella* is highly debated (Sibley and Monroe 1990; AOU, 1998; Klicka et al, 1999; Canales-Del Castillo et al, 2010). Our analyses suggest that the difficulty to resolve the phylogenetic relationships unambiguously results from inappropriate representation as a bifurcating tree. A well-resolved network representation implies that there was gene flow by effects of hybridization between several species thus obscuring their phylogenetic history.

Early studies often excluded some species from their analyses and this opened the possibility for inaccurate inferences on phylogenetic relationships. Our results support the early hypothesis based on phenotypic characters (Mayr and Short, 1970) that *S. wortheni* and *S. pusilla* are closely related species. It also strengthens the sister taxon relationship of *S. passerina* and *S. breweri* which was found by others (Zink and Dittmann, 1993; Dodge et al., 1995; Carson and Spicer, 2003; Jhønsson and Fjeldså, 2006).

Most recently, Canales-Del Castillo et al. (2010) presented a molecular phylogeny that was well supported at most nodes, but not fully resolved. One of their main conclusions was that *S. wortheni* is more closely related to *S. breweri* than to any other species. Our findings do not support this hypothesis. Yet for the following reason we are quite confident in our results: First, similar to us, Canales-Del Castillo et al. (2010) found evidence for polytomies which resulted in one unsatisfactorily resolved central node. Second, the exclusion of *S. taverneri* from their work might have influenced their findings. Although, Klicka et al. (1999) found only small differences in the Cyt-b sequences between the former species and *S. breweri*, they proposed the recognition of *S. taverneri* as a separated taxon rather than to be a subspecies of *S. breweri*. Based on uncorrected divergence distances we found that these

species show enough differences at least in three individual genes to be included in our analysis (ATP8 = 5.95%, COIII = 7.02%, ATP6 = 8.70%, COI = 0.0%, Cytb = 0.3%).

We recognize that Canales-Del Castillo et al. (2010) used a longer sequence alignment in their study (3571 bp, compared to 2889 bp in our work) and that increased information might enhance results. However, when we added fragments of ND2 and CRI sequences (for a total length of 4233 bp), we did not receive a better resolved phylogeny (Appendix A, Fig. A.1): persistently, polytomies and low supported nodes were present in both BI and ML tests. Despite these unsatisfactory results, however, in both analyses the clade *S. wortheni* + *S. pusilla* was well supported (BI = 99%, ML = 93%). Moreover, by excluding *S. taverneri* (as in Canales-Del Castillo et al. 2010) from our analysis of the concatenated sequence (2889 bp), the persistence of polytomies disappeared but internal nodal support remained poor (Appendix A, Fig. A.3).

Split decomposition and reticulate networks revealed that the lack of clarity in the former analyses might be due to the possibility of hybridization events among ancestral lineages in the past. In general, puzzling phylogenetic relationships between passerines can be caused by the rapid radiation of most of lineages in the early Tertiary, resulting in a lack of synapomorphies (Feducchia, 1995). Our reticulate network results allow two possible interpretations: First, the failure to find fixed differences between taxa could reflect very recent speciation events within the last few thousands of years (Mallet, 2008). According to our analyses, *Spizella* lineages seem to have diverged within a relative short period of time that began approximately 2.5 - 3.3 million years ago, within the late Pliocene. Our estimation must be taken with caution as the LRT results did not support a consistent molecular clock. Yet it is similar to the divergence time calculated by Canales-Del Castillo et al. (2010) and it

also coincides with the beginning of a period thought to have an explosive radiation in passerine birds (Feducchia, 1995; Poe and Chubb, 2004; Weir and Schluter, 2007). Second, the inconsistency in the genetic signal might be caused by hybridization between ancestral lineages or among recent species. Although not yet confirmed by molecular studies, several reports for contemporary hybridization among *Spizella* species exist, for instance: *S. passerina* x *S. pallida* (Parkes, 1990), *S. breweri* x *S. taverneri* (MacCArthy, 2006 and references therein) and *S. pallida* x *S.pusilla* (Hoag, 1999). Moreover, hybrids of *Spizella* birds with other passerines such as *Passer montanus* or *Pooecetes gramineus* also occur (MacCArthy, 2006 and references therein). Hybridization in natural avian populations is especially common when birds have low chances to find mates (e.g. at the edges of its distributional ranges or when they are disorientated during migration) or when male birds are imprinted on an incorrect song (Hoag, 1999; MacCArthy, 2006). As the distributional ranges of most of the *Spizella* species overlap and some species have very similar songs (Webster and Orr, 1954; Behrstock et al., 1997), we think that hybridization account to a high degree for the conflicting phylogenetic relationships of this genus.

In summary, our analyses support the hypothesis of sister relationships between respectively, *S. wortheni* - *S. pusilla* and *S. breweri* - *S. passerina*. For at least five mitochondrial genes, the phylogenetic relationships between lineages could not be completely resolved in a bifurcating tree, but are better represented in a reticulate network. Our study shows how the application of network analyses can resolve long-lasting debates of phylogenetic relationships; this might be especially important in phylogenetic reconstructions on passerine birds which have a history of rapid radiations combined with common hybridization events.

Acknowledgements

We thank to the Museum of the North of University of Alaska and to LSU Museum of Natural Science for facilitate us with tissue samples used for sequencing in this work. CD thanks to CONACyT-DAAD for financial support as fellowship.

Chapter 4

Why are some species rare: A case study on passerine Spizella birds[4]

J. Canales-Delgadillo, L. Scott-Morales & J. Korb.

Abstract

Causes of species rarity are diverse and are commonly attributed to reduction of suitable habitats and habitat fragmentation. However, species may also become rare through evolutionary and ecological processes that result in high specialization or a narrow range of adaptability during times of environmental change. We did a study on passerine birds of the genus *Spizella* to improve our understanding why some species are rare compared to their congeneric counterparts. The Worthen's sparrow (*Spizella wortheni*) is an endemic rare species restricted to the shrub-grassland habitats in the arid and semi-arid regions of northeast Mexico, in contrast to its congenerics, it is not a long-distance migrant but behaves as a nomadic bird. We used multivariate analyses to search for patterns of habitat filtering among *Spizella* sparrows and estimated their bioclimatic marginality, specialization, niche breadth and niche overlap. *Spizella* species occurred in 15 out of 21 vegetation types found in the North American region, principally in areas of shrub-cropland-tree or grassland mosaics. Our results suggest that communities of *Spizella* sparrows are phylogenetically clustered. We found no evidence of similarity in the response to environmental variation among species. Overall, *Spizella* species showed marginality values > 1 and high niche overlap. The high specialization and low tolerance estimated for *S. wortheni* suggest that its rarity is linked to its patterns of dispersal behavior.

[4] Manuscript

Keywords: *Spizella*, rarity, phylogenetics, community structure.

Introduction

In the ecosystems not all species are equally common, some are abundant others common and the most, rare. Biological rarity is defined as the inverse of the magnitude of abundance, range size or a combination of the two (Gaston 1994). Causes of rarity are diverse. Nowadays, they can often be attributed to a reduction of suitable habitats and habitat fragmentation, which negatively affect population sizes and breeding success. However, species may also become rare through evolutionary processes resulting for instance from a narrow niche with strict habitat requirement, limited dispersal or low per capita growth rates. Low population densities and restriction of geographical ranges might also be a consequence of recent speciation (Kuning and Gaston, 1993 and references herein).

We present a bird study to improve our understanding why some species are rare compared with their congeneric counterparts. The genus *Spizella* consists of eight species of small passerine sparrows native to North America where they occur from the tropical and subtropical regions of Mexico to the boreal forests in Alaska and Canada. Winter migration for most of these species is well known; they span from Canada and the USA to northern-central Mexico. Breeding grounds for most species occur in all three countries (e.g.: Naugler 1993, Tenney 1997, Rotenberry et al. 1999, Middleton 1998, Carey et al. 2008, Winter et al. 2006, Scott-Morales et al. 2008, Canales-Delgadillo et al. 2008). Only the Worthen's sparrow (*Spizella wortheni*), probably the least studied species in the genus, is an endemic rare species. Its distribution is restricted to the shrub-grassland habitats in the arid and semi-arid regions of

northeastern Mexico, after populations in the southern part of its historical distribution range went extinct (Wege et al. 1993, Scott-Morales et al. 2008). In contrast to its congeners, *S. wortheni* is not a long-distance migrant but behaves as a highly nomadic bird (Canales-Delgadillo et al. unpub. manuscript). Recently, a microsatellite-based genetic study demonstrated that despite its low population sizes in one of the remaining populations (BirdLife International 2000: 120 birds, Canales-Delgadillo 2006: 111 birds), no genetic depletion could be detected and gene flow among localities was high (Canales-Delgadillo et al. unpub. manuscript).

In the current study, we used multivariate analyses on the phylogenetic structure of *Spizella* communities to investigate whether we can detect patterns of phylogenetic attraction or repulsion between species at large geographical scale. Helmus et al. (2007a) defined phylogenetic attraction as closer phylogenetic relatedness among species in a community than expected by chance, while lower than expected relatedness indicates phylogenetic repulsion. As phylogenetic attraction/repulsion can be caused by environmental factors, we investigated whether factors such as altitude, temperature or rainfall might influence the presence/absence of species across vegetation communities. We also estimated for each species its bioclimatic marginality (the difference of the species niche from the conditions in the global area with respect to its mean), specialization (the difference of the species niche from the conditions in the global area with respect to its standard deviation), niche breadth (the range, as determined by the bioclimatic variables, within which species' niche occur) and niche overlap (the region of niche space shared by two or more species).

Methods

Data sampling

We collected data on the geographical distribution of all eight species of the genus *Spizella* for the entire region of North America (Canada, the USA and Mexico), this included *Spizella taverneri* (Sybley and Monroe 1990), which is considered a subspecies of *Spizella breweri* by some authorities (AOU, 1998), while other consider it an independent species (our recent phylogeny confirms the latter perspective, Canales-Degadillo et al. in prep.). Occurrence data for all species were obtained from the Global Biodiversity Information Facility database (www. gbif.org). As *S. wortheni* is largely understudied, we supplemented this data set by our own records for this species collected between 2003 and 2008.

A set of environmental data, consisting of eleven bioclimatic variables, altitude and vegetation cover was also collected (Table 1). Bioclimatic variables are frequently used in ecological niche modeling and represent temperature and rainfall annual ranges as well as extremes (for details see Table 1). Data for these variables and the altitudes were obtained from WorldClim database 1.4 (Hijmans et al. 2005). A cover layer with 21 vegetation categories was obtained from the Global Land Cover dataset (Latifovic et al. 2003). Due to the large geographical scale of this study, each vegetation category was treated as a community (hereafter called vegetation community).

All used data are publicly available for download in a gis-usable format with a resolution of about 1 km^2. As the data were portioned by country, each environmental variable was merged into a single geo-raster covering the entire region of North America and clipped when necessary to have raster layers with exactly the same resolution and geographical extent. All geographical data were handled using DIVA-GIS v.7.0 and QGIS v.1.6.

Presence/absence data across vegetation communities for all species as collected from the geographical records were summarized in a binary matrix. A phylogenetic tree, including all *Spizella* species was used to compute a pairwise phylogenetic variance-covariance matrix among taxa. Finally, a matrix was constructed that contained the mean values of the 12 environmental variables for the 21 vegetation communities.

Phylogenetic community structure

Following Helmus et al. (2007a), we first analyzed the phylogenetic community structure of *Spizella* sparrows for each of the 21 vegetation communities (excluding environmental variables) using phylogenetic species variability (PSV) metric. This summarizes the degree to which species in a community are phylogenetically related (Helmus et al. 2007b). PSV values close or equal to one indicate that the species in a community are unrelated. Thus, that community shows maximum phylogenetic variability, i.e. overdispersed community structure. As the PSV approaches zero the phylogenetic relatedness between species increases and the variability within the community decreases, indicating a clustered community structure (Helmus et al. 2007a, Helmus et al. 2007b). We used two null hypotheses to test our data: first, that communities consist of random draws of species from the whole species pool of eight species and second, that communities were assembled by selecting species from the species pool in proportion to their prevalence (presence/absence) among vegetation communities (Helmus et al. 2007a). To test the observed data against the null hypotheses, two null distributions were generated using 10 000 permutations. These analyses were restricted to those vegetation communities with at least two species.

Phylogenetic attraction and repulsion

In a next step, we tested the influence of environmental factors for explaining phylogenetic community structure to reveal repulsion/attraction effects. To do this, we fitted multiple logistic regressions on the correlations between species co-occurrences and the correlations of phylogenetic relatedness among species, once with and once without including environmental variables as filtering species factors. To avoid overfitting of the model we used maximum penalized likelihood approach (Kosmidis 2010). We followed the logistic regression model described by Helmus et al. (2007a) which had the form:

$\text{logit}(\theta(x_k)) = \mathbf{b}_o + \mathbf{b}_{1x1,k} + \ldots + \mathbf{b}_{jxj,k}$

where $\theta(x_k)$ is a 8 x 1 vector of the probabilities that each of the 8 species occurs in a vegetation community $k(k = 1,\ldots, K)$; $x_{j,k}$ is the value of the jth environmental variable for each vegetation community $k(j = 1,\ldots, 12)$ and $\mathbf{b}_j$ is a 8 x 1vector of regression coefficients for each of the 8 species in response to environmental variable j. Significance of the correlations was tested using 10 000 replicates in a permutation test.

To determine which environmental factors filter species, $\mathbf{b}_j$ were treated as species traits in order to estimate the phylogenetic signal (K), as a measure of the similarity in which species respond to environmental variation (Helmus et al. 2007a, Ives et al. 2007, Blomberg 2003). A K close or larger than 1 for an environmental variable j, implies that species resemble ecologically each other more than expected under Brownian motion evolution, meaning that species respond in similar way to that environmental variable (Blomberg 2003, Helmus et al. 2007a). To test the significance of K under the null hypothesis of no similarity in the species' response to environmental variation, p values were computed based on variance of phylogenetic independent contrasts relative to tip shuffling randomization, using

5 000 replicates (Blomberg 2003, Kembel et al. 2010). All these analyses were carried out as implemented in the R package Picante (Kembel et al. 2010).

Bioclimatic marginality and specialization

In terms of ecological niche, species have an optimum within a range of environmental conditions in which they can exist. Thus, habitats laying out this optimum imply some level of marginality (i.e. species mean differs from the global mean) or specialization (species variance is lower than the global variance) that might restrict species' distribution (Hirzel et al. 2002). Both marginality and specialization can be estimated by means of ecological niche factor analysis (ENFA) using only-presence data (Hirzel et al. 2002). Similar to a principal component analysis, ENFA summarizes several environmental predictors into a few uncorrelated factors that explain most of the ecological information (Hirzel and Perrin 2007). The first factor obtained from ENFA represents those variables from which the species niche differs mostly from the conditions of the global area, e. i. marginality (*M*). Although *M* normally ranges between zero and one, it might exceed the unity (Hirzel et al. 2002). This is evidence that the focal species had very particular habitat requirements. The other factors in the ENFA, represent the degree of specialization (*S*), which measure how restricted the species niche is as compared with the total available habitat (Hirzel et al. 2002, Reutter et al. 2003). Specialization coefficient varies between zero and infinite. Tolerance can be estimated as the inverse of *S* and ranges between zero and one. Tolerance values close or equal to one indicate that species inhabit wider niche than species showing values close to zero.

To estimate the degree of bioclimatic marginality and specialization, thirteen ecogeographical maps (EGV's) were constructed from the bioclimatic variables, altitude and vegetation communities. Prior to analysis, all EGV's were Box-Cox normalized. Species

maps consisted of additional raster layers, one per species, with the same geographical extension and resolution as the EGV's but containing a value of 1 only for those cells where the focal species is present. Estimation of *M*, *S* and tolerance, as well as each species niche breadth and niche overlap between species pairs was carried out as implemented in Biomapper 4.0 (Hirzel and Perrin 2007). Here, we refer to the species niche as the realized niche, which is defined as the region of its niche that a species is able to occupy in the presence of interspecific competition and natural enemies (Emerson and Gillespie 2008). From the several indices of niche breadth and niche overlap returned by Biomapper, we present here the Hulbert's indices because different to others, it assumes that individuals are equivalent in their behavior (Hulbert 1978).

Results

Spizella species occurred in 15 out of the 21 vegetation types found in North America. They were absent from open broad deciduous and needle-leaved forest, but occurred in evergreen forests or when forest were closed. Records of these birds were also absent from regularly flooded and bare areas as well as from cropland-tree cover mosaics. By contrast, cropland-shrub-tree or grassland mosaics accounted for most records.

Community structure

The average PSV value was small (PSV_{obs}= 0.0578), indicating that communities of co-occurring *Spizella* sparrows are composed by closely related species. According with the tree's topology, *S. wortheni* and *Spizella pusilla* are sister taxa, while the sister species *S. breweri* and *Spizella passerina* form a three-species complex with *S. taverneri* (Fig. 1). The prevalence of *S. wortheni* across communities was notably lower than the prevalence of its

sister taxa *S. pusilla* (Fig. 1), which highlights the restricted distribution of the former. For the three-species complex and for most of the remaining species in the phylogeny, their level of prevalence among communities was similar (Fig. 1). Under null hypothesis 1 (communities consist of random arrangements of the whole species pool) communities of *Spizella* sparrows seem to be phylogenetically clustered (PSV_{obs}= 0.0578, $PSV_{null\ 1}$ = 0.0532, p = 0.01). That closely related species (e.g., *S. taverneri*, *S. breweri* and *S. passerina*) tend to occur in the same communities was also supported through rejection of null hypothesis 2 which accounted for the species' prevalence ($PSV_{null\ 2}$ = 0.0568, $p < 0.001$).

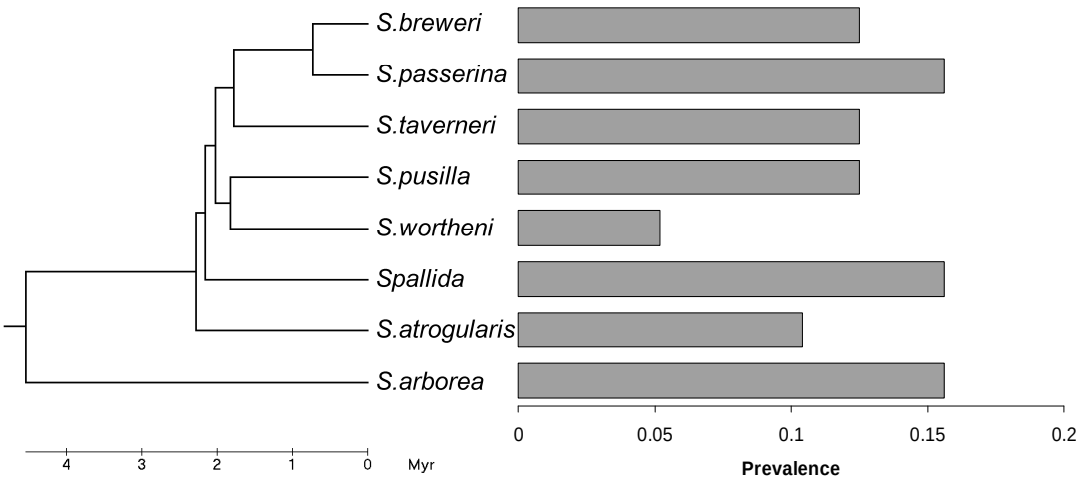

Figure 1. Phylogeny (scaled in million years) and prevalence of *Spizella* species across 15 vegetation communities in North America.

Phylogenetic attraction and repulsion

We found no evidence of phylogenetic attraction or repulsion after the logistic regression analysis. No statistical relation between the correlations of species co-occurrence and their phylogenetic relatedness was found neither when environmental variables were

ignored nor when included (Fig. 2a, 2b respectively). Also, no significant difference in the co-occurrence correlation of closely related species was observed with respect to the co-occurrence of the more distantly related species (Fig. 2c). These results suggest that environmental variables do not add anything in explaining the fact that phylogenetically closely related species often co-occur. Accordingly, we found no evidence that species respond in similar ways to environmental variation. The estimated *K* values were similar for all variables and always < 0.5 (Table 1). No significant difference between the observed data and the random variance was found ($p > 0.05$, Table 1). Thus, the environmental variables do not describe adequately the presence/absence of the species across vegetation communities according to the *Spizella* phylogeny.

Table 1. Phylogenetic signal of eleven environmental variables on the presence/absence of species across communities. Variables are sorted by decreasing value of *K*.

Environmental variable	K	PIC_{obs}	PIC_{rnd}	PIC_{p}
Precipitation seasonality	0.498	0.365	0.315	0.827
Precipitation of driest quarter	0.482	0.679	0.554	0.863
Precipitation of coldest quarter	0.480	4.159	3.405	0.861
Mean temperature of warmest quarter	0.478	1.555	1.236	0.779
Mean annual precipitation	0.477	2.650	2.150	0.866
Precipitation of wettest quarter	0.477	7.358	5.929	0.839
Precipitation of warmest quarter	0.477	9.304	7.524	0.852
Altitude	0.472	0.001	0.001	0.853
Mean temperature of driest quarter	0.469	2.919	2.304	0.858
Mean temperature of coldest quarter	0.469	12.639	10.121	0.855
Mean annual temperature	0.466	7.059	5.466	0.873
Mean temperature of wettest quarter	0.454	0.523	0.397	0.856

PIC_{obs} and PIC_{rand}: Observed and expected variance in phylogenetic inertia correlation, PIC*p*: p-value of observed vs. random variance.

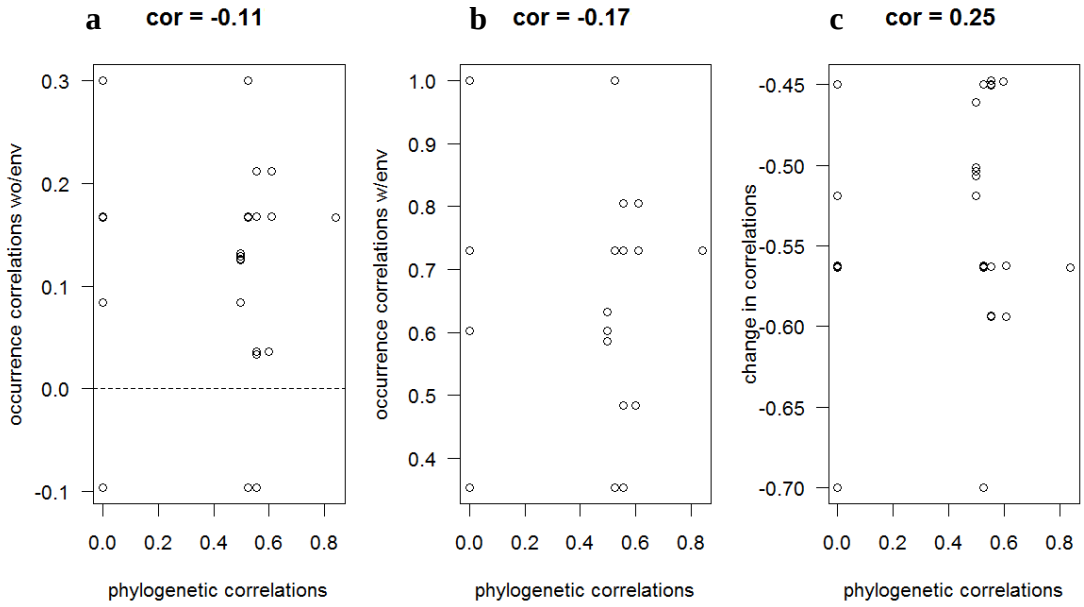

Figure 2. Pairwise phylogenetic correlations against pairwise species co-occurrence correlations estimated without (a) and with (b) environmental variables included. No significant association between the species co-occurrence and their phylogenetic relatedness that indicate patterns of attraction or repulsion was found ($p > 0.05$). Plot (c) shows the change in the correlations co-occurrence once environmental variables are included.

Bioclimatic Marginality and Specialization

We found no correlation between *M* and *S* (Pearson's product-moment correlation: $r = 0.39$, $t_6 = 1.05$, $p = 0.333$), but significant negative correlations were found between *M* and tolerance ($r = -0.75$, $t_6 = -2.83$, $p = 0.030$) and *S* and tolerance ($r = -0.80$, $t_6 = -3.24$, $p = 0.017$). Overall, *M* values among *Spizella* species were larger than 1 (Table 2). Coefficients of *M* showed were similar between species for seasonal extremes of temperature: mean temperature of warmest quarter (0.30 - 0.45), mean temperature of coldest quarter (0.26 - 0.44) and mean temperature of driest quarter (0.18 - 0.46). Contrasting, *M* coefficients for rainfall variables (e.g. annual mean precipitation) differed largely between species (-0.09 - 0.33). *S. wortheni* had the highest degree of specialization, more than one order of magnitude larger than all other species, which had similar values (Table 2). *S. wortheni* also had the lowest level of

tolerance, implying that this species is highly restricted by its habitat requirements. According to the ENFA analysis, habitat marginality of *S. wortheni* is related to altitude and the seasonal minimal and maximal temperatures (Table 3).

Overall, the niche overlap between most species pairs was close to or larger than 1, indicating that *Spizella* birds use some resources (e.g. shrub patches, grasslands) more intensively than others (e.g. tree cover, cropland) and that this use tends to coincide between species (Hulbert 1978) (Table 4). All *Spizella* species had similar niches breadths, except *S. wortheni* and *S. atrogularis* which showed smaller values (Table 4).

Table 2. ENFA estimation of the overall marginality and specialization of the *Spizella* species in North America.

Species	Marginality	Specialization	Tolerance
S. arborea	1.578	2.633	0.380
S. atrogularis	1.735	3.574	0.280
S. breweri	1.428	2.539	0.394
S. pallida	1.073	1.690	0.592
S. passerina	1.472	1.763	0.567
S. pusilla	1.927	3.453	0.290
S. taverneri	1.582	1.970	0.508
S. wortheni	1.823	71.657	0.014

Table 3. Coefficient values for the ecological variables used in the estimation of marginality and specialization of *S. wortheni*. Explained variance for the first six factors is indicated in brackets.

Ecogeographical variable	Marginality (80%)	Spec. 1 (12%)	Spec. 2 (4%)	Spec. 3 (1%)	Spec. 4 (1%)	Spec. 5 (1%)
Altitude	0.48	0.24	0.02	0.03	-0.06	0.00
Mean temperature of coldest quarter	0.45	0.01	-0.50	-0.61	-0.29	0.56
Mean annual temperature	0.42	-0.21	0.68	0.77	0.87	-0.77
Mean temperature of wettest quarter	0.38	-0.09	0.24	-0.15	-0.13	-0.02
Mean temperature of driest quarter	0.36	-0.15	-0.12	0.03	-0.29	-0.01
Mean temperature of warmest quarter	0.32	0.16	-39.00	-0.07	-0.18	0.26
Precipitation seasonality	0.08	0.04	0.01	-0.02	-0.03	-0.03
Precipitation of warmest quarter	0.03	0.16	0.05	0.05	-0.02	0.00
Vegetation type	-0.02	0.00	0.00	0.00	0.00	0.00
Mean annual precipitation	-0.02	-0.58	0.14	0.03	-0.10	0.07
Precipitation of wettest quarter	-0.02	0.05	-0.15	-0.07	0.11	-0.02
Precipitation of driest quarter	-0.03	-0.40	0.10	0.03	-6.00	0.04
Precipitation of coldest quarter	-0.04	0.55	-0.07	-0.07	-3.00	-0.14

Table 4. Estimation of Hulbert's niche overlap *L* (below diagonal) and niche breadth B' indices (in bold on the diagonal) among *Spizella* species based on presence/absence on vegetation communities of North America.

Species	*S. arborea*	*S. atrogularis*	*S. breweri*	*S. pallida*	*S. passerina*	*S. pusilla*	*S. taverneri*	*S. wortheni*
S. arborea	**0.520**							
S. atrogularis	2.20	**0.181**						
S. breweri	2.21	3.07	**0.360**					
S. pallida	0.93	1.84	1.82	**0.601**				
S. passerina	1.82	1.89	2.02	1.08	**0.537**			
S. pusilla	1.96	1.48	1.71	0.94	1.96	**0.440**		
S. taverneri	1.34	2.26	1.99	1.06	1.70	1.42	**0.580**	
S. wortheni	0.19	2.50	0.15	1.04	1.18	0.04	0.04	**0.035**

Discussion

Our results demonstrate that ecological processes such as specialization imply habitat restriction, which might contribute to species rarity. The coexistence of *Spizella* sparrows across communities appears not to be affected by competition since large realized niche overlap exists among species. *S. wortheni* differs from its congenerics in its small niche breath and, as mentioned above, in its nomadic behavior while all other species are migrants. These traits could explain the rarity of this species.

Community structure and coexistence

Our PSV analysis showed that *Spizella* species are phylogenetically clustered across vegetation communities. This might be explained by their general pattern of habitat use: habitats with mean temperatures ranging from -10°C to 20°C and with low to moderate

rainfall regimes (30 mm to 195 mm) are preferred for *Spizella* sparrows. Furthermore, shrubs seems to be crucial for both breeding and non-breeding activities for all species (e.g.: Best 1978, Reynolds and Knapton 1984, Biermann et al. 1987, Carey 1990).

A phylogenetically clumped distribution of taxa indicates that habitat filtering is acting on the species pool in the communities (Emerson and Gillespie 2008). Thus, our results suggest that niche among *Spizella* species might be a phylogenetically conserved trait. These results together with the high niche overlap observed, suggest that the coexistence of *Spizella* species is not limited for processes such as competitive exclusion. Additionally, behavioral traits such as flocking in mixed-species flocks frequently with congenerics during the non-breeding season and during migration trips (see Naugler 1993, Knapton 1994, Tenney 1997, Rotenberry et al. 1999, Middleton 1998, Carey et al. 2008, Winter et al. 2006, Canales-Delgadillo et al. 2008), might explain the pattern of community structure we found.

Assemblage community rules (Diamond 1975) postulate that communities are composed of non-random combinations of species which avoid niche overlap of closely phylogenetically or ecologically related organisms. This concept was strongly debated by others (see Götzenberg et al. 2011) and it has been demonstrated that phylogenetic relatedness or ecological similarity do not always lead to competitive exclusion (see Wiens and Rotenberry 1979 for a review on the coexistence of ecologically similar grassland and steppe bird species in North America). On the other hand, low food availability might also force birds to coexist by exploration of wider niches (Kober and Bairlein 2009).

Co-occurrence and successful coexistence across communities of cladistically or ecologically very similar species has been also found for other organisms, such as fishes (Winston 1995). For sympatric cyprinid fishes foraging rates, niche breadth and territoriality of co-occurring species do not change with food availability fluctuations (Martin and Genner 2009). For

Spizella species, at geographical large-scale habitat filtering allows a successful coexistence of closely related species without considerable conflicts caused by niche overlap.

The set of environmental variables we used in this study showed that *Spizella* species respond mainly to seasonal extremes of temperature. Lack of significance in the phylogenetic signal of the environmental variables might be due to the low number of taxa in our sample. This is because the few possible permutations in small phylogenetic trees make it difficult to detect phylogenetic signals at $p < 0.05$ (Blomberg et al. 2003)

Rarity of Spizella wortheni

One accepted hypothesis to explain rarity is that recently diverged taxa have not yet occupied all possible niches they might be able to use. On the other hand, it is also argued that oldest taxa tend decline in abundances, which eventually leads to species extinction (Kuning and Gaston, 1993). Neither the former nor the later seems to be the case for *S. wortheni* (Fig. 1). In the phylogeny, basal and the more recently diverged species (Fig. 1) show a wider distribution and higher population numbers than *S. wortheni*. Therefore, we do not think that the evolutionary age can explain the rarity of this species.

Besides its narrow niche breadth and low tolerance (Table 2 and 3), the reason of the rarity of *S. wortheni* might lay in its behavior. Among *Spizella* birds, *S. wortheni* is the only species that does not migrate neither for breeding nor wintering. Different to the other species, *S. wortheni* behaves as a nomadic bird (Canales-Delgadillo, unpubl. manuscript).

Our results are congruent with other studies addressing questions on rarity. Birds with specialized feeding behavior or narrow habitat requirements and non-migratory habits tend to be rarer than those with generalist habits or wider niche breadths (e.g., Goerck 1997, Cofre et al 2007). Cognitive abilities and resources-use specialization might also influence bird

populations due to the flexibility in resource use and the ability to cope with habitat changes (Schultz et al. 2007). A comparative study between *S. wortheni* and congenerics showing contrasting characteristics such as synanthropy (e.g., *S. passerina* or *S. arborea*), might help to improve our understanding of the nature of its specialist behavior and therefore of its rarity.

Conclusion

Our results indicate that neither competitive exclusion nor evolutionary history can explain the rarity of *S. wortheni*. Its rarity might be rather linked to its nomadic behavior and its narrow niche. *Spizella* bird communities are composed by closely related species that co-occur among fifteen different vegetation types across North America. We did not find evidence that species respond in similar ways to environmental variation. The independence of correlations co-occurrence and phylogenetic relatedness among species across vegetation communities suggest that competitive exclusion does not affect the coexistence of these birds.

General Discussion

Pollution, transformation of natural habitat into farmland and other human-related activities have been the main causes of bird populations decline during the last decades, leading to an increasing number of threatened species (Donovan et al. 1995, Peterjohn 2003, Freeland 2005, Díaz et al. 2006). This affects especially endemic species and those with restricted distribution areas. For the characterization of the genetic background, important ecological factors and evolutionary processes influencing species survival, population genetics are a useful tool. They help in identifying the best strategies for diversity conservation at all levels of the biological hierarchy resulting in stable ecological and evolutionary processes that sustain the biological diversity on the earth (Avise 2004).

In chapter 1, I present the first molecular tools developed specifically for the study of *S. wortheni* populations. Since microsatellite markers from other passerine species were only applicable to some extent, it was necessary to develop specific markers. Using a hybridization capture approach (Glenn and Schable 2005), I was able to isolate nine variable novel microsatellite markers which reliably amplified DNA from *S. wortheni*.

By using the markers I isolated, and four additional microsatellites developed for other bird species, I did population genetic analyses for seven remnant populations of *S. wortheni* (Chapter 2). Its populations have been described as small and isolated within its distribution range. Thus, I expected to find low genetic diversity. I found that habitat fragmentation has apparently no strong influence on the genetic diversity and the gene flow among the study localities. Potential reasons may be a very recent habitat fragmentation, since it are known to

mask loss of genetic diversity within populations (see Larsson et al. 2008). Alternatively, and perhaps more likely, this unexpected result may be caused by the nomadic life style of *S. wortheni*, which make it tolerant to habitat modification.

Nomadic birds living in arid and semi-arid areas are known to be adapted to unpredictable environmental conditions characteristic of these places. By moving from one patch to another they are able to find sites where resources for feeding or breeding are available (Dean 2004). This behavior allows the genetic exchange between different groups, which prevents the loss of genetic diversity. The dispersal range of *S. wortheni* is > 5km, which allows it to travel between apparently isolated patches (landscape connectivity in the study area is around 30% - 60%). Accordingly, I hypothesize that nomadic birds might be more resilient against negative effects of habitat fragmentation, but only if levels of landscape connectivity allow constant exchange between patches. My findings are supported by previous research on the "buffer effect", which describes that a high mobility of birds protects populations to some degree from negative consequences of habitat fragmentation (Galbusera et al 2004).

In line with this, no evidence for isolation by distance was found since the seven studied localities form a quasi-panmictic group. A weak but significant structure of the population, given by four genetic clusters, was observed. High mobility of *S. wortheni* was confirmed by high rates of gene exchange between them (Chapter 2). For instance, heterozygosity values for *S. wortheni* were similar to those found for its close relative, the Brewer's Sparrow (*Spizella breweri*), which is a common migratory species in North America (Croteau et al. 2007).

Despite the favorable levels of genetic variability among study localities, I found high and significant levels of inbreeding in three of these localities. One of the principal genetic

threats for wild populations is inbreeding depression (Frankham 1995). Inbred populations show low reproductive success and increased extinction risk. I attribute the high levels of inbreeding to the tendency of *S. wortheni* to move in gregarious groups (Canales-Delgadillo et al. 2008). It is true that nomadism might help to solve some of the difficulties of living in arid and semi-arid habitat; however, a drawback of this behavior is that non-random mating might occur (Dean 1997). When mechanisms of kin-recognition are absent, as for wild populations of the Great Tit (*Parus major*), the likelihood of brother-sister mating increases when siblings disperse in similar directions (Szulkin and Sheldon 2008). Thus, the tendency of *S. wortheni* to move in gregarious groups might contribute to the high levels of inbreeding.

Previously, the knowledge about *S. wortheni*'s biology was limited to basic aspects of its distribution and habitat use. My research has revealed novel aspects on the population genetic parameters of this species that might help to improve conservation efforts. However, an important aspect of this bird, namely the phylogenetic relationships to its congeners remained unclear: Pioneer work on the phylogeny of the genus *Spizella* (Zink and Dittmann 1993, Carson and Spicer 2003, Canales-Del Castillo et al. 2010) was unsatisfactory as representatives were missing in the analyses. This made previous conclusions ambiguous.

In my study on the phylogenetic relationships of the *Spizella* clade (Chapter 3), for the first time all representatives were included, even *Spizella taverneri.* Currently, there is no consensus on the status of *S. taverneri* as an independent species (Sibley and Monroe, 1990; Klicka and Zink, 1999); some times it is considered as a subspecies of *S. breweri* (AOU, 1998).

By using Bayesian Inference and maximum likelihood approaches I analyzed DNA segments from five mitochondrial genes. As for many other avian groups (Jhønsson and Fjeldså; 2006), the construction of bifurcating phylogenetic trees for the *Spizella* clade was problematic. My results suggested a sister taxon relationship between *S. wortheni* - *S. pusilla*, while a close relationship was also found between the clade *S. breweri* - *S. passerina* and *S. taverneri*. These results differed from those by Canales-Del Castillo et al. (2010). Their molecular data, suggest that *S. wortheni* is more related to *S. breweri* than to *S. pusilla*. However, like me, they also had difficulties solving a three-way polytomy and observed low statistical support in a central node of their tree.

Uncertainties in phylogenetic relationships can arise due to rapid species radiation or can be a consequence of hybridization events (Huson and Bryant, 2006.). By means of phylogenetic network analysis I realized that potential hybridization among ancestors of the recent taxa might be obscuring the evolutionary history of this group. A well resolved network showed three cyclic reticulations from which lineages of four contemporary taxa descend. Although not confirmed by molecular analyses, reports of hybrids between different *Spizella* species (*S. passerina* x *S. pallida* (Parkes, 1990), *S. breweri* x *S. taverneri* (MacCArthy, 2006) and *S. pallida* x *S.pusilla* (Hoag, 1999)) support my findings. My results support the hypotheses that *S. wortheni* and *S. pusilla* (Wege et al. 1993) and *S. breweri* and *S. passerina* (Carson and Spicer 2003) are sister taxa.

An intriguing question in the field of ecology linked to my research is why species are rare. In the past it was postulated that rare species are poorly adapted organisms which will inevitably go extinct (Drury 1974). However, many rare species have existed for long time periods and strategies of rare species to avoid extinction have been documented for several

organisms (e.g.: Williams et al. 2009). Yet, the causes of rarity *per se* are not yet well understood. Based on my research with *S. wortheni* - which also historically is a rare species - I tested several hypotheses about the rarity of species using *Spizella* birds as a model system (Chapter 4).

Rarity can be explained by diverse reasons. They are often related to a reduction of suitable habitats and habitat fragmentation, which negatively affect population sizes and breeding success. However, it can also be a consequence of recent speciation (Kuning and Gaston, 1993). Based on the study of the phylogenetic community structure, patterns of phylogenetic attraction and repulsion and estimation of marginality, specialization and niche overlap among species, I investigated the possible causes of rarity in *S. wortheni*.

The phylogenetic community structure analysis suggested that *Spizella* species are clustered across vegetation communities. This could be explained by the general habitat use of these birds: all are highly dependent on shrubby landscape for both breeding and non-breeding activities (e.g. Best 1978, Reynolds and Knapton 1984, Biermann et al. 1987, Carey 1990). Assuming phylogenetically conserved ecological traits, the phylogenetically clumped distribution may indicate habitat filtering among *Spizella* species. This is in line with a high niche overlap between co-occuring species which also suggests that competitive exclusion does not play an important role in structuring *Spizella* communities. Additionally, behavioral traits such as flocking in mixed-species flocks, frequently with congenerics during the non-breeding season and during migration trips (see Naugler 1993, Knapton 1994, Tenney 1997, Rotenberry et al. 1999, Middleton 1998, Carey et al. 2008, Winter et al. 2006, Canales-Delgadillo et al. 2008), might contribute to explain the pattern of community structure I found, and suggest that the rarity of *S. wortheni* is not a matter of interspecific competition.

According to the ecological niche factor analysis, the factors limiting the distribution of *S. wortheni* are altitude and the seasonal minimum and maximum extremes of temperature. This implies that *S. wortheni* is highly restricted by its niche requirements. Recently diverged taxa might not yet occupy all possible niches they might be able to use, thus having a restricted distribution. On the other hand, also the oldest taxa tend to decrease in abundance, thus becoming rare (Kuning and Gaston, 1993). Neither the former nor the later seem to be the case for *S. wortheni*. According to the phylogeny, species older (*Spizella atrogularis*) or younger (*S. breweri*) than *S. wortheni* have wider distribution ranges and higher population sizes. Therefore, evolutionary age can be excluded as the reason of the rarity of *S. wortheni.*

Specialist habits may arise from evolutionary changes that induce a shift to more restricted habitat requirements (Futuyama and Moreno 1988). Compared with its congenerics, *S. wortheni* is the most specialized of all *Spizella* species. This is demonstrated by its low tolerance to variation of climatic conditions as well as by its narrow niche breadth.

Rarity of *S. wortheni* might mirror a potential evolutionary change in its behavior that results in niche specialization. My results are matching other studies on rarity (e.g., Goerck 1997, Cofre et al 2007), that found that birds with specialized feeding behavior, narrow habitat requirements and non-migratory habits tend to be rarer than those with generalist habits or wider niche breadth.

Biological conservation programs require a solid knowledge of the focal species to be protected. Understanding the biology of species and the processes affecting the population dynamics will help to make accurate decisions on what to protect and the best way to protect it. The studies I presented here open new possibilities to understand the current state of the *S. wortheni* populations, its relationships with other species and with the environmental factors

affecting its prevalence in the shrub-grassland areas of northeastern Mexico. Further studies on its life history are needed to prevent this species from going extinct.

Since I found moderate high levels of genetic diversity among the studied populations in a fragmented landscape, I suggest that nomadism makes bird species more resilient to habitat fragmentation. This work is also added to those trying to establish the phylogenetic relationships among *Spizella* birds and to those that could deduce the evolutionary causes of species rarity.

Summary

The development of genetic tools to study populations at the molecular level has been one of the most important contributions to understand demographic processes in wild populations of conservation concern. Habitat loss and habitat fragmentation are the main factors causing declines of birds populations, as in the case of the Worthen's sparrow (*Spizella wortheni*), a Mexican endemic Emberizid restricted to scrub and grassland habitats in northeastern Mexico.

Here, I present the first molecular tools developed specifically for the study of *S. wortheni* populations. Genetic analyses of seven remnant populations of *S. wortheni* showed that no genetic impoverishment is present. This unexpected result may be caused by the nomadic life style of *S. wortheni*, which makes it tolerant to habitat modification. The high levels of inbreeding I found can be attributed to the tendency of *S. wortheni* to move in gregarious groups, since non-random mating might be present.

My analyses based on the study of DNA segments from five mitochondrial genes support the hypotheses that *S. wortheni* and *S. pusilla* (Field Sparrow), as well as *S. breweri* (Brewer's Sparrow) and *S. passerina* (Chipping Sparrow) are sister taxa. A phylogenetic network analysis showed that conflicting relationships among *Spizella* species might be caused by possible hybridization events between ancestors of extant taxa.

Based on the study of the phylogenetic community structure, patterns of phylogenetic attraction and repulsion, and on the estimation of marginality, specialization and niche overlap among species, I investigated the possible causes of rarity in *S. wortheni*. The phylogenetic community structure analysis suggested that *Spizella* species are clustered across vegetation communities. Assuming phylogenetically conserved ecological traits, the

phylogenetically clumped distribution may indicate habitat filtering among *Spizella* species. Additionally, behavioral traits such as flocking in mixed-species flocks might contribute to explain the pattern of community structure I found, and suggest that the rarity of *S. wortheni* is not a matter of interspecific competition. According to the phylogeny, evolutionary age was also discarded as a cause of rarity of this species. Compared with its congenerics, *S. wortheni* is the most specialized of all *Spizella* species. A potential evolutionary change of *S. wortheni* behavior that results in niche specialization is more likely to be the cause of its rarity.

Zusammenfassung

Die Entwicklung von molekular-genetischen Methoden zur Untersuchung von Populationen ist eine der wichtigsten Beiträge, um demografische Prozesse von Wildpopulationen hinsichtlich des Artenschutzes zu verstehen. Verlust und Fragmentation von Lebensräumen sind die wichtigsten Faktoren für den Rückgang von Vogelpopulationen, wie im Fall der Zacatecasammer (Worthen's Sparrow, *Spizella wortheni*, Emberizidae), einer, Ammer, die in Strauch- und Graslandschaften im Nordosten Mexikos endemisch ist.

Hier stelle ich die Ergebnisse der ersten molekular-genetischen Untersuchungen an *S. wortheni* Populationen vor. Grundlage dafür sind neun artspezifische Nukleär Mikrosatelliten, die im Rahmen dieser Arbeit isoliert wurden.

Eine Analyse der sieben letzten vorhandenen *S. wortheni* Populationen konnte keine genetische Verarmung nachweisen. Dieses unerwartete Ergebnis kann durch die nomadische Lebensweise von *S. wortheni* erklärt werden, die sie flexibel auf Habitatveränderungen reagieren lässt. Die Untersuchungen ergaben ein hohes Inzuchtpotential. Dies ist möglicherweise auf die Eigenart von *S. wortheni* zurückzuführen in Schwärmen zu wandern, was eine nicht zufällige Partnerwahl für die Fortpflanzung bedingen könnte.

Meine Analysen, basierend auf der Untersuchung von DNA-Segmenten von fünf mitochondrialen Genen, unterstützen die Hypothese, dass *S. wortheni* und *S. pusilla* (Klapperammer), sowie *S. breweri* (Nevadaammer) und *S. passerina* (Schwirrammer) Schwesterarten sind. Eine Phylogenetische Netzwerk-Analyse hat gezeigt, dass die widersprüchlichen Beziehungen zwischen den *Spizella*-Arten durch mögliche Hybridisierungsreignisse zwischen den Vorfahren der rezenten Taxa verursacht worden sein könnte.

Aufgrund der Untersuchungen zur phylogenetische Struktur der Artengesellschaft, dem Muster der phylogenetischen Anziehung und Abstoßung, sowie der Abschätzung von Marginalität, Spezialisierung und Nischenüberlappung der Arten, habe ich die möglichen Ursachen der Seltenheit von *S. wortheni* untersucht. Die phylogenetische Struktur der Artengesellschaft-Analyse lässt vermuten, dass die *Spizella*-Arten über Vegetationstypen hinweg aggregiert vorkommen. Unter der Annahme von phylogenetisch konservierten ökologischen Eigenschaften, gibt die aggregierte phylogenetische Verteilung möglischerweise einen Hinweis auf Habitatsfilterung der *Spizella*-Arten. Zusätzlich könnten Verhaltensmerkmale, wie die Neigung interspezifische Schwärme zu bilden, das von mir gefundene Muster der Artengesellschaft erklären. Des weiteren legen diese Verhaltensweisen den Schluß nahe, dass die Seltenheit von *S. wortheni* nicht durch interspezifische Konkurrenz hervorgerufen wird. Aufgrund des phyllogenetischen Alters von ca 1.8 Millionen Jahren kann das Alter der Art als Grund für ihre Seltenheit ausgeschlossen werden. Im Bezug auf die Habitatansprüche ist *S. wortheni* die spezialisierteste aller *Spizella*-Arten.. Im Bezug auf die Habitatansprüche ist *S. wortheni* die spezialisierteste aller *Spizella*-Arten. Die hohe Spezialisierung in Bezug auf die Habitatwahl könnte ein wahrscheinlicher Grund für die Seltenheit von *S. wortheni* sein.

References

Albrech, D.J., and Oring L.W., 1995. Song in Chipping sparrow, *Spizella passerina*: structure and function. Animal Behavior 50: 1233-1241.

American Ornithologists' Union (AOU), 1998. Check-list of North American Birds. 7th Edition.

Andersson, M. 1980. Nomadism and site tenacity as alternative reproductive tactics in birds. Journal of Animal Ecology 49: 175-184.

Barton, M. H. and Slatkin, M. 1986. A quasy-equilibrium theory of the distribution of rare alleles in a subdivided population. Heredity 56: 409-415.

Behrstock RA, CW Sexton, GW Lasley, TL Eubanks, JP Gee (1997) First nesting records of Worthen's sparrow *Spizella wortheni* from Nuevo Leon, Mexico, with habitat characterization of the nest site and notes on ecology, voice, additional recent sightings and leg coloration. Cotinga 8:27–33

Benjamini, Y. and Hochberg, Y. 1995. Controling the False Discovery Rate: A practical and powerful approach to multiple testing. -Journal of the Royal Statistical Society Series B (Methodological) 57: 289-300.

Best, L. B. 1978. Field Sparrow reproductive success and nesting ecology. Auk 95:9-22.

Biermann, G.C., W.B. McGillivray, and K.E. Nordin. 1987. The effect of cowbird parasitism on Brewer's Sparrow productivity in Alberta. Journal of Field Ornithology 58(3):350–354.

BirdLife International., 2008. *Spizella wortheni*. In: IUCN 2010. IUCN Red List of Threatened Species. Version 2010.1. Available from <www.iucnredlist.org> (accessed on May 2010).

Blomberg, S. P., Garland, T. J. and Ives, A. R. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution, 57, 717–745.

Bulguin, N. L., Gibbs, H. L., Vickery, P. and Baker, A. J. 2003. Ancestral polymorphisms in genetic markers obscure detection of evolutionary distinct populations in the endangered Florida grasshopper sparrow (*Ammodramus savannarum floridanus)*. Molecular Ecology 12: 831-844.

Burleigh, T.D. and Lowery, G.H., 1942. Notes on the birds of southeastern Coahuila. Occas. Pap. Mus. Zool. La. Univ. 12, 185-21.

Canales-Del Castillo, R., Klicka, J., Favela, S., and Gonzáles-Rojas, J.I., 2010. Molecular phylogenetic analysis of and endangered Mexican Sparrow: *Spizella wortheni*. Molecular Phylogenetics and Evolution 57: 1319-1322.

Canales-Delgadillo J. C. 2006. Uso de hábitat y relaciones sociales de Spizella wortheni. Tesis de Maestría. Facultad en Ciencias Forestales. UANL, México.

Canales-Delgadillo, J. C., Scott Morales, L., Niehuis, O. and Korb, J. 2010. Isolation and characterization of nine microsatellite loci in the endangered Worthen's Sparrow (*Spizella wortheni*). -Conservation Genetics Resources 2: 151-153.

Canales-Delgadillo, J. C., Scott-Morales, L., Cotera, C. M., and Pando, M. 2008. Observations on the flocking behavior of Worthen’s sparrow and occurrence in mixed-species flocks. The Wilson Journal of Ornithology 120: 569-574.

Carey, M. 1990. Effects of brood size and nestling age on parental care by male Field Sparrows (*Spizella pusilla*). Auk 107:580-586.

Carey, M., Burhans, D. E., and Nelson, D. A. 2008. Field Sparrow (Spizella pusilla), The Birds of North America Online (A. Poole, Ed.). Ithaca: Cornell Lab of Ornithology;

Retrieved from the Birds of North America Online: http://bna.birds.cornell.edu/bna/species/103 doi: 10.2173/bna.103

Carson, R.J. and Spicer, G.S. 2003., A phylogenetic analysis of the emberizid sparrows based on three mitochondrial genes. Molecular Phylogenetics and Evolution 29: 43-57.

Cofre, H. L., Böhning-Gaese, K., and Marquet, P. A. 2007. Rarity in Chilean forest birds: which ecological and life-history traits matter? Diversity and Distributions 13: 203-212.

Coulon, A. C., Guillot, G., Cosson, J.F., Angibault, J.M.A., Aulagnier, S., Cargnelutti, B., Galan, M., Hewison, A. J. M., 2006. Genetic structure is influenced by landscape features: empirical evidence from a roe deer population. Molecular Ecology 15: 1669-1679.

Crochet, P. A. 2000. Genetic structure of avian populations – allozymes revisited. -Molecular Ecology 9: 1463-1469.

Croteau, E. K., Lougheed, S. C., Kranitz, P. G., Mahony, N. A., Walker, B. L. and Boag, P. T. 2007. Genetic population structure of the sagebrush Brewer's Sparrow, *Spizella breweri breweri*, in a fragmented landscape at the northern range periphery. Conservation Genetics: 1453-1463.

Dean, W. R. J. 1997. The distribution and biology of nomadic birds in the Karoo, South Africa. -Journal of Biogeography 24: 769-779.

Dean, W. R. J. 2004. *Nomadic desert Birds.*, Berlin, Springer-Verlag.

Diamond, J. M. 1975. The assemblage of species communities. in: Ecology and Evolution of communities. Cody. M. L., and Diamond, J. M. (Editors). Harvard University Press. Cambridge, MA.

Díaz, I. A., Armesto, J. J. and Willson, M. F. 2006. Mating success of the endemic Des Murs'

Wireteil (*Sylviorthorhynchus desmusirii*, *Furnariidae*) in fragmented Chilean rainforest. -Austral Ecology 31: 13-21.

Dodge, A.G., Fry A.J., Blackwell R.C., and Zinck R.M., 1995. A comparison of phylogenies derived from two molecular data sets in the avian genera *Pipilo* and *Spizella*. The Wilson Bulletin. 107: 641-654.

Donovan, T. M., Thomson III, F. R., Faaborg, J. and Probst, J. R. 1995. Reproductive success of migratory birds in habitat sources and sinks. -Conservation Biology 9: 1380-1395.

Ehrich, D. 2006. AFLPdat: a collection of R functions for convenient handling of AFLP data. Molecular Ecology Notes 6: 603-604.

Ehrich, D., Gaudeul, M., Assefa, A., Koch, M., Mummenhoff, K., Nemomissa, S., Consortium, I. and Brochmann, C. 2007. Genetic consequences of Pleistocene range shifts: Contrast between the Arctic, the Alps and the East African mountains. -Molecular Ecology 16: 2542-2559.

Emerson, B. C., and Gillespie, R. G. 2008. Phylogenetic analysis of community assembly and structure over space and time. Trends in Ecology and Evolution, 23 (11): 619-630.

Evanno, G., Regnaut, S. and Goudet, J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. -Molecular Ecology 14: 2611-2620.

Excoffier L, G Laval, S Schneider (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1: 47–50

Excoffier, L. and Lischer, H. E. L. 2010. Arlequin suite version 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10: 564-567.

Excoffier, L. and Slatkin, M. 1998. Incorporating genotypes of relatives into a test of linkage disequilibrium. -American Journal on Human Genetics 62: 161-180.

Fagan, W. F., Unmack, P. J., Burgess, C. and Minckley, W. L. 2002. Rarity, fragmentation and extinction risk in desert fishes. Ecology 83: 3250-3256.

Feduccia, A., 1995. Explosive evolution in tertiary birds and mammals. Science 267: 637-638.

Fontaine, M. C., Baird, S. J. E., Piry, S., Ray, N., Tolley, K. A., Duke, S., Birkun, A. J., Ferreira, M., Janiaux, T., Llavona, A., Öztürk, B., Öztürk, A. A., Ridoux, V., Rogan, E., Sequeira, M., Siebert, U., Kikingsson, G. A., Bouquegneau, J. M. and Michaux, J. R. 2007. Rise of oceanographic barriers in a continuous population of a cetacean: the genetic structure of harbor porpoises in Old World waters. BMC Biology 5: 1-16.

Frankham, R. 1995. Conservation Genetics. -Annual Reviews Genetics 29: 305-327.

Freeland, J. R. 2005. Molecular Ecology. United Kingdom, John Wiley and Sons.

Galbusera P, Githiru M, Lens L. and Matthysen E (2004) Genetic equilibrium despite habitat fragmentation in an Afrotropical bird. Molecular Ecology 13:1409–1421.

Gaston, K. J. 1994. Rarity. Chapman and Hall. UK.

Glenn TC, NA Schable (2005) Isolating microsatellite DNA loci. In: Zimmer, E. A, and Roalson, E. H. Methods in enzymology 395:202–222.

Goerck, J. M. 1997. Patterns of rarity in the birds of the Atlantic forest of Brazil. Conservation Biology, 11(1): 112-118.

Goodson, J.L., 1998. Territorial aggression and down song are modulated by septal vasotocin and vasoactive intestinal polypeptide in the field sparrow (*Spizella pusilla*). Hormones and Behavior. 34, 67-77.

Götzemberg, L., de Bello, F., Bråthen, K. A., Davison, J., Dubuis, A., Guisan, A., Lepš, J., Lindborg R., Moora, M., Pärtel, M., Pellicier, L., Pottier, J., Vittoz, P., Zobel, K., and Zobel, M. 2011. Ecological assembly rules in plant communities—approaches, patterns and prospects. Biological Reviews doi: 10.1111/j.1469-185X.2011.00187.x

Guillot, G., Mortier, F. and Stoup, A. 2005a. Geneland: a computer package for landscape genetics. -Molecular Ecology Notes 5: 708-711.

Guillot, G., Stoup, A., Mortier, F. and COSSON, J. F. 2005b. A spatial statistical model for landscape gentics. -Genetics 170: 1261-1280.

Guindon S., and Gascuel, O., 2003. A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol. 52, 696-704.

Guindon, S., Dufayard J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O., 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst Biol. 59, 307-321.

Guo, S. and Thompson, E. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. -Biometrics 48: 361-372.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41:95-98

Helmus, M. R., Bland T. J., Williams, C. K., and Ives, A. R. 2007b. Phylogenetic measures of Biodiversity. The American Naturalist, 169 (3): 68-83.

Helmus, R. M., Savage, K., Diebel, M. V., Maxted, J. T., and Ives, A. R. 2007a. Separating the determinants of phylogenetic community structure. Ecology Letters, 10: 917-925.

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis, 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965-1978.

Hirzel, A. H., Hausser, J., Chessel, D., and Perrin, N. 2002. Ecological-niche factor analysis: how to compute habitat suitability maps without absence data? Ecology, 83 (7): 2027-2036.

Hirzel, A.H., J. Hausser, & N. Perrin (2007) Biomapper 4.0. Lab. of Conservation Biology, Department of Ecology and Evolution, University of Lausanne, Switzerland. URL: http://www2.unil.ch/biomapper

Hoag, D. J., 1999. Hybridization between Clay-colored Sparrow and Field Sparrow in Northern Vermont. The Wilson Bulletin. 111: 581-584.

Huelsenbeck, J.P. and Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics. 17: 754-755.

Hulbert, S. H. 1978. The measurement of niche overlap and some relatives. Ecology, 59 (1): 67-77.

Huson, D.H., and Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Molecular Biology and Evolution. 23: 254-267.

IUCN (2009) IUCN Red List of Threatened Species version 2009.2. www.iucnredlist.org Accessed 14 January 2010

Ives, A. R., Midford, P. E. and Garland, T. 2007. Within-species variation and measurement error in phylogenetic comparative methods. Syst. Biol., 56: 252–270.

Jackobsson, M. and Rosenberg, N. A. 2007. clumpp: a cluster matching and permutation program with label switching and multimodality in analysis of population structure. Bioinformatics 3: 1801-1806.

Jeffery, K. J., Keller, L. F., Arcese, P. and Bruford, M. W. 2001. The development of microsatellite loci in the song sparrow *Melospiza melodia* (Aves) and genotyping errors associated with good quality DNA. -Molecular Ecology Notes 1: 11-13.

Jhønsson, K., and Fjeldså J., 2006. A phylogenetic supertree of oscine passerine birds (Aves:Passeri). Zoologica Scripta. 35: 149–186.

Kalinowski, S. T. 2010. The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. Heredity: 1-8.

Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P., and Webb, C.O. 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26:1463-1464.

Klicka J., and Zink, R.M., 1997. The importance of recent ice ages in speciation: a failed paradigm. Science 277: 1666–1669.

Klicka, J., and Zink, R.M., 1999. Evidence supporting the recent origin and species status of the Timberline sparrow. Condor 101: 577-588.

Klicka, J., Burns, K., and Spellman, G.M., 2007. Defining a monophyletic Cardinalini: A molecular perspective. Molecular Phylogenetics and Evolution 45: 1014-1032.

Knapton, Richard W. 1994. Clay-colored Sparrow (Spizella pallida), The Birds of North America Online (A. Poole, Ed.). Ithaca: Cornell Lab of Ornithology; Retrieved from the Birds of North America Online: http://bna.birds.cornell.edu/bna/species/120 doi: 10.2173/bna.120

Kober, K., and Bailein, F. 2009. Habitat choice and niche characteristics under poor food conditions. A study on migratory Nearctic shorebirds in the intertidal flats of Brazil. Ardea, 97 (1): 31-42.

Kosmidis, I. 2010. brglm: Bias reduction in binary-response GLMs. http://go.warwick.ac.uk/kosmidis/software

Kumar S., and Gadagkar S.R., 2001. Disparity Index: A simple statistic to measure and test the homogeneity of substitution patterns between molecular sequences. Genetics 158: 1321-1327.

Kunin, W. E., and K. J. Gaston. 1993. The Biology of Rarity: Patterns, Causes and Consequences. TREE 8: 298-301.

Larsson, J. K., Hugh, Jansman H. A. H., Segelbacher, G.,Höglund, J., and Koelewijn, H. P. 2008. Molecular Ecology 17: 1897-1904.

Latifovic, R., Z. Zhu, J. Cihlar, J. Beaubien, and R. Fraser. 2003. The Land Cover Map for North America in the Year 2000. GLC200 database, European Commission, Joint Research Centre. http://www-gem.jrc.it/glc2000.

Lindsay, D. L., Barr, K. R., Lance, R. F., Tweddale, S. A., Hayden, T. J. and L., L. P. 2008. Habitat fragmentation and genetic diversity of an endangered, migratory songbird, the golden-cheeked warbler (Dendroica chrysoparia). -Molecular Ecology 17: 2122-2133.

Lowe, A., Harris, S. and Ashton, P. 2004. Ecological Genetics. Design, analysis and applications. United Kingdom, Blackwell Publishing.

MacCarthy, E.M., 2006. Handbook of avian hybrids of the world. Oxford University Press. NY, USA.

Mace, M. G., and Kershaw, M. 1997. Extinction risk and rarity on an ecological time scale. In: Kunin, E and Gaston, K. J. The biology of rarity: causes and consequences of rare-common differences. Chapman and Hall. UK.

Mahoni, N.A., Kranitz P.G., and Martin K., 2006. Seasonal fecundity of sagebrush Brewer's sparrow (*Spizella breweri*) at the northern edge of its breeding range. Auk 123: 512-523.

Mallet, J., 2008. Hybridization, ecological races and the nature of species: empirical evidence of the easy of speciation. Phil Trans R Soc B. 363, 2971-2986.

Martin C. H., and M. J. Genner. 2009. High niche overlap between two successfully coexisting pairs of Lake Malawi cichlid fishes. Canadian Journal of Fisheries and Aquatic Sciences 66: 579-588.

Mayr, E., and Short, L.L. Jr., 1970. Species Taxa of North American Birds: A contribution to comparative systematics. Publications of the Nuttall Ornithological club, no. 9. Cambridge, Massachusetts, USA.

Middleton, Alex L. 1998. Chipping Sparrow (Spizella passerina), The Birds of North America Online (A. Poole, Ed.). Ithaca: Cornell Lab of Ornithology; Retrieved from the Birds of North America Online: http://bna.birds.cornell.edu/bna/species/334 doi: 10.2173/bna.334

Moulton, V., and Huber, K.T., 2009. Split networks. A tool for exploring complex evolutionary relationships in molecular data, in: Lemey, P., Salemi, M., and Vandamme, A.M. (Eds.), The Phylogenetic Handbook: A Practical Approach to DNA and Protein Phylogeny. 2nd edition, Cambridge University Press.

Naugler, Christopher T. 1993. American Tree Sparrow (Spizella arborea), The Birds of North America Online (A. Poole, Ed.). Ithaca: Cornell Lab of Ornithology; Retrieved from the Birds of North America Online: http://bna.birds.cornell.edu/bna/species/037 doi: 10.2173/bna.37

Niehuis O, J Korb (2010) Isolation and characterization of seventeen polymorphic microsatellite markers in the cleptoparasitic cuckoo wasp *Hedychrum nobile* (Hymenoptera: Chrysididae). Conservation Genetic Resources DOI 10.1007/s12686-009-9167-8

Owens, I. P. F. and Bennett, P. M. 2000. Ecological basis of extinction risk in birds: Habitat loss versus human persecution and introduced predators. Proceedings of the National

Academy of Sciences 97: 12144-12148.

Parkes, K.C., 1990. Additional records of birds from the Distrito Federal, Mexico, including a possible hybrid Spizella. Condor 92: 1080-1081.

Peakall, R. and Smouse, P. E. 2006. GENALEX 6: genetic analysis in Excel. Population genetics software for teaching and research. -Molecular Ecology Notes 6: 288-295.

Peterjohn, B. G. 2003. Agricultural Landscapes: Can they support healthy bird populations as well as farm products? Auk 120: 14-19.

Petren, K. 1998. Microsatellite primers from *Geospiza fortis* and cross-species amplification in Darwin's finches. -Molecular Ecology 7: 1782-1784.

Poe, S., and Chubb, A.L., 2004. Birds in a bush: five genes indicate explosive evolution of avian orders. Evolution 58: 404-415.

Posada D., 2008. jModelTest: Phylogenetic Model Averaging. Molecular Biology and Evolution 25: 1253-1256.

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

Ralph, C. J., Geupel, G. R., Pyle, P., Martin, T. E. and Desante, D. F. 1993. Handbook of field methods for monitoring land birds. -Gen. Tech. Rep. PSW-GTR-144-www. U.S. Department of Agriculture, California, USA.

Rannala, B. and Mountain, J. L. 1997. Detecting immigration by using multilocus genotypes. -Proceedings of the National Academy of Sciences 94: 9197-9201.

Reutter, B. A., Helfer, V., Hirzel, A. H., and Vogel, P. 2003. Modelling habitat-suitability using museum collections: an example with three sympatric *Apodemus* species from the Alps. Journal of Biogeography 30: 581-590.

Reynolds, J. D. and R. W. Knapton. 1984. Nest-site selection and breeding biology of the Chipping Sparrow. The Wilson Bulletin 96:488-493.

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43:223–225.

Ronquist, F. and Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

Rotenberry, J. T., M. A. Patten and K. L. Preston. 1999. Brewer's Sparrow (Spizella breweri), The Birds of North America Online (A. Poole, Ed.). Ithaca: Cornell Lab of Ornithology; Retrieved from the Birds of North America Online: http://bna.birds.cornell.edu/bna/species/390 doi: 10.2173/bna.390

Rozen S, HJ Skaletsky (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, S Misener (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365–386

Scott-Molares, L., Estrada, E. C., Chávez-Ramírez, F., Cotera, C. M., 2004. Continued decline in geographic distribution of the Mexican Prairie dog (*Cynomys mexicanus*). Journal of Mammalogy 85: 1095-1101.

Scott-Morales, L., Nocedal, J., Cotera, M. and Canales-Delgadillo, J. C. 2008. Worthen's Sparrow (*Spizella wortheni*) in the Northern Mexican Plateau. -Southwestern Naturalist 53: 91-95.

Segelbacher, G., Manel, S. and Tomiuk J. 2008. Temporal and speatioal analysis disclose consequences of habitat fragmentation on the genetic diversity in capercaillie (Tetrao urogallus). Molecular Ecology 17: 2356-2367.

SEMARNAT. 2008. Compendio de Estadísticas Ambientales. Edición 2008. http://app1.semarnat.gob.mx/dgeia/informe_2008/index_informe_2008.html

Seutin, G., White, B. N. and Boag, P. T. 1991. Preservation of avian blood and tissue samples

for DNA analyses. -Canadian Journal of Zoology 69: 82-90.

Shultz, S., Bradbury R. B., Evans, K. L., Gregory, R. D., and Blackburn, T. M. 2007. Brain size and resource specialization predict long-term population trends in British birds. Proceedings of the Royal Society, B 272: 2305-2311.

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

Slatkin, M. and Barton, N. H. 1989. A comparition of three indirect methods for estimating average levels of gene flow. Evolution 43: 1349-1368.

Sybley, C.G., and Monroe, B.L., 1990. Distribution and Taxonomy of Birds of the World. Yale University Press, New Haven and London, USA.

Szulkin, M. and Sheldon, B. C. 2008. Correlates of the occurrence of inbreeding in a wild bird population. Behavioral Ecology DOI: 10.1093/beheco/arn086.

Tamura, K., Dudley, J., Nei, M., and Kumar, S., 2007 MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24: 1596-1599.

Tenney, Chris R. 1997. Black-chinned Sparrow (Spizella atrogularis), The Birds of North America Online (A. Poole, Ed.). Ithaca: Cornell Lab of Ornithology; Retrieved from the Birds of North America Online: http://bna.birds.cornell.edu/bna/species/270 doi: 10.2173/bna.270

Van Oosterhout, C., Hutchinson, W., Willis. P. M. D., and Shipley, P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535-538.

Webster, J.D., and Orr, R.T., 1954. Summering birds of Zacatecas, Mexico, with a description of a new race of Worthen [sic] Sparrow. Condor 56: 155-160.

Wege, D. C., Howell, S. N. G., and Sada A. M. 1993. The distribution and status of Worthen's sparrow *Spizella wortheni*: a review. Bird Conservation International 3: 211-220.

Weir, J.T., and Schluter, D., 2007. The latitudinal gradient in recent speciation and extinction rates of birds and mammals. Science. 315. 1574-1576.

Wiens, J. A., and Rotenberry, J. T. 1979. Diet niche relationships among North american grassland and shrubsteppe birds. Oecologia, 42: 253-292.

Willson, F.E. 1990. Extra ocular control of seasonal reproduction in female tree sparrow (*Spizella arborea*). Gen Comp Endocrinol. 77, 397-402.

Winter, M., Douglas H. J., Shaffer, J. A., Donovan T. M, and Svedarsky, W. D. 2006. Patch size and landscape effects on density and nest success of grassland birds. Journal of Wildlife Management 70(1): 158-172. Jamestown, ND: Northern Prairie Wildlife Research Center Online http://www.npwrc.usgs.gov/resource/birds/patchsiz/index.htm (Version24AUG2006).

Wisnton, M. R. 1995. Co-occurrence of morphologically similar species of stream fishes. The American Naturalist, 145 (4): 527-545.

Xia, X., and Lemey, P., 2009. Assessing substitution saturation with DAMBE, in: Lemey, P., Salemi, M., and Vandamme, A.M. (Eds.), The Phylogenetic Handbook: A Practical Approach to DNA and Protein Phylogeny. 2nd edition, Cambridge University Press. pp. 615-630.

Xia, X., and Xie, Z., 2001. DAMBE: data analysis in molecular biology and evolution. Journal of Heredity 92: 371-371.

Xia, X., Xie, Z., Salemi, M., Chen, L., and Wang, Y., 2003. An index of substitution saturation and its application. Molecular Phylogenetics and Evolution 26: 1-7.

Zink, R.M., and Dittmann, D.L., 1993. Population structure and gene flow in the Chippng Sparrow and a hypothesis for evolution in the genus *Spizella*. The Wilson Bulletin 105: 399-413.

Acknowledgements

I thank to Consejo Nacional de Ciencia y Tecnología (CONACYT) and to the Deutscher Akademischer Austausch Dienst (DAAD) for accept me as scholarship holder and for supporting my stay in Germany.

I would like thank to Prof. Dr. Judith Korb for giving me the opportunity of belonging to her workgroup and learn with her, for all the time she invested in reviewing my work and for all the useful tips she gave me to improve as biologist.

I would also thank to Dr. Laura Scott Morales for help me to improve my work, for encourage me to study in Germany, for the unconditional support before my travel and during my whole stay in this country and for her friendship.

I thank especially to Barbara Hausberger, Katharina Hoffmann and Johannes Gowin for help me to improve my German language skills, for share your time with me and for all these moments of Schnaps and Gutes Essen.

Finally I would like to say that it was a wonderful experience to meet all the people from the Verhaltensbiologie – Heike, Rebecca, Anja, Eva, Henry, Walter – and all others there. I am sure I had a question for each of you, thank you for your answers and your help!

Erklärung über die Eigenständigkeit der erbrachten wissenschaftlichen Leistung.

Ich erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirect übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.
Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise entgeltlich / unentgeltlich geholfen.
1. Prof. Dr. Judith Korb stand mir als Betreuerin in alle Phasen der Arbeit beratend zur Seite und war behilflich bei der vorbereitung der Datenaufnehme, der statistischen Auswertung und dem Verfassen der Manuskripte.
2. Dr. Laura Scott Morales war behilflich bei der vorbereitung der Datenaufnehme in Mexiko und dem Verfassen der Manuskripte.
3. Dr. Oliver Niehuis war behilflich bei dem Verfassen des Kapitel eins und bei der Mikrosatelliten Entwicklung.

Weitere Personen waren an der inhaltlichen materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.
Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Osnabrück den 15.07.2011 _____

Ort. Datum Unterschrift

Supplementary Materials

Supplementary material for chapter 2: Analysis of population structure implemented with STRUCTURE 2.0.

STRUCTURE is a program to recognize genetically differentiated populations, to assign individuals probabilistically to genetic groups, and to identify migrants. It uses a model-based method that accounts for the deviations from Hardy Weinberg equilibrium (HWE) and linkage disequilibrium (LD) by introducing population structure into the whole data set to identify groups that are not in disequilibrium (Pritchard et al. 2000). For the estimation of *K* we performed several runs to test the behavior of our data. We selected the admixture option and the correlated frequency model because we assumed some level of gene flow and hence that populations were related. We allowed *K* to vary from 1 to 10 and 2 x 10^6 iterations on the MCMC including a burn-in period of 8 x 10^5 steps were used, with ten replications for each *K*. To check for consistency of our estimate of the probability of the data (LnPr(X|K)) we performed 20 additional replicates but with *K* ranging from 1 to 5 and leaving all other parameters unchanged. These analyses were carried out on the freely available Bioportal at the University of Oslo (<www.bioportal.uio.no>).

Data Analysis

The results obtained with STRUCTURE were analyzed using Structure-sum.2009 (Ehrich 2006, Ehrich et al. 2007), a script of the R-programming language, which includes a series of functions to analyze results of genetic clustering from STRUCTURE. Using this script we summarized the data of LnPr(X|K) across all runs to determine the true value of *K* (Evanno et al. 2005, Pritchard et al. 2000).

To visualize the genetic structure we used the estimated membership fractions for each individual across the best runs. The coefficient matrices of membership from each program were aligned and averaged with CLUMPP 1.1.2 (Jakobsson and Rosenberg 2007) and drawn with DISTRUCT 1.1 (Rosenberg 2004).

Results

With STRUCTURE the LnPr(X|K) values were similar between K = 1 and K = 3 (Fig. S1a). Thus, we investigated the individuals' assignment matrices for these values of K because a symmetric pattern in the percentage distribution of the individuals' assignment within the clusters might indicate a lack of population structure (Pritchard et al., 2000). For K = 3 there was no symmetric pattern in the percentage of assignments to clusters. The maximization of ΔK method (Ehrich 2006, Ehrich et al. 2007, Evanno et al. 2005) showed that the true number of clusters inferred with STRUCTURE is K = 3 (Fig. S1b). The individuals from Manantial, Esperanzas, Perforadora and Soledad were distributed across all three clusters. Only individuals from Frayle were assigned to a single cluster, while individuals from Guerrero and Angeles were assigned into two of the three groups (Fig. S2).

The Evanno's criterion for optimization of K is an approach that helps to find the true number of genetic clusters when initial analyses with STRUCTURE show no a clear pattern for a given peak of K as in Fig. S1a, where K =1 was similar to K =3. Evanno's et al. (2005) optimization clearly demonstrated that $K > 1$ (Fig. S1b), this result is supported by the findings of GENELAND, and pairwise R_{ST}; it is also according to the geographic distribution of the study localities.

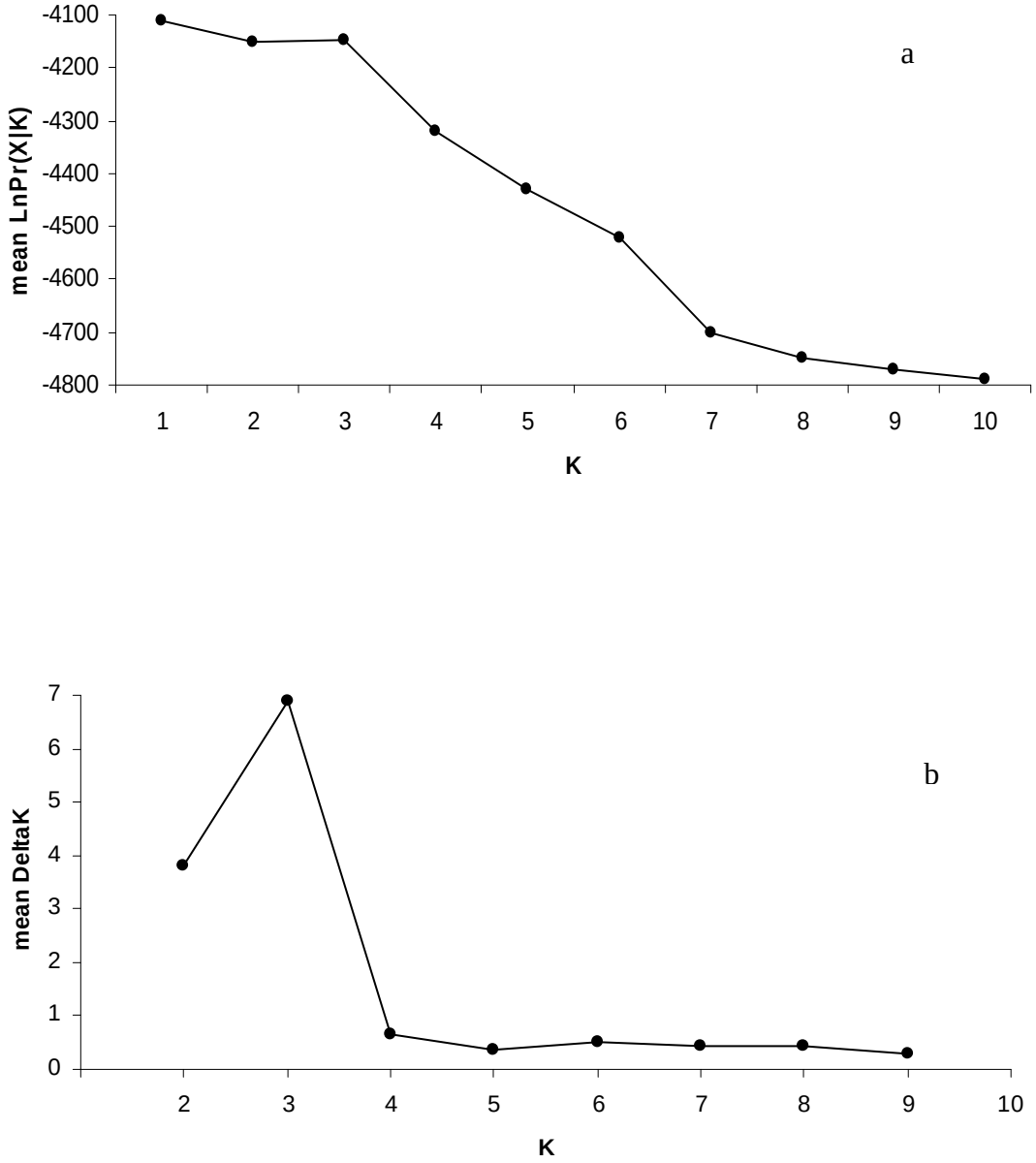

Figure S1. Estimated number of genetic clusters from STRUCTURE. In a), the mean value of LnPr(X|K) shows no clear peak pattern with similar values between $K = 1$ and $K = 3$, b) shows the optimum value of K after application of Evanno's et al. (1995) optimization criterion.

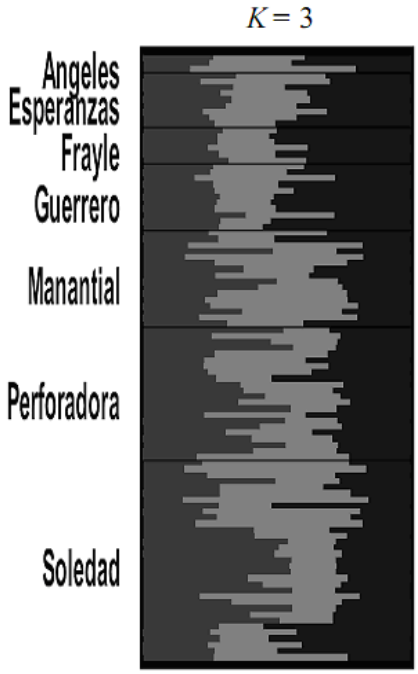

Figure S2. Population structure of Worthen's sparrow estimated with STRUCTURE with fixed K = 3, predefined study localities are shown in blocks and separated by a black horizontal line, for each individual the probability to belong to a cluster K is indicated by a thin colored line.

Suplemental Material for Chapter 3

Appendix A

Table A1. Primer pairs used for DNA amplification and sequencing of S. wortheni and S. taverneri samples. Note the use of internal primers for the amplification of genes ATP6, ATP8 and COIII after an initial amplification with primers tRNA-Lys-corvus and CO3BB-corvus.

Gen	Primer pair	Forward sequence 5'- 3'	Reverse sequence 5'- 3'	Source
COI	BirdF1/BirdR2	TCTCCAACCACAAAGACATTGGCAC	ACTACATGTGAGATGATTCCGAATCCAG	Kerr et al. 2007
Cyt-b	L14995/H-16065	CTCCCAGCCCCATCCAACATCTCAGC ATGATGAAACTTCG	CTAAGAAGGGTGGAGTCTTCAGTTTTTGG TTTACAAGAC	Helbig et al. 1995
ATP6	SPIZ1FD/CO3BB Corvus	GACCCATCCTAGGAGCAGCCGC	GATTGGAAGTCGATTATAAT	Carson and Spicer 2003
ATP6	tRNA-Lys-corvus/SPIZ2BK	CAGCACTAGCCCTTTTAAGCTAG	GTACGAAGACGTAGGCTTGG	Carson and Spicer 2003
ATP8	tRNA-Lys-corvus/SPIZ3BK	CAGCACTAGCCCTTTTAAGCTAG	GATTAGTGCTCATTTGTG	Carson and Spicer 2003
COIII	SPIZ2FD/CO3BB-corvus	GACCCATCCTAGGAGCAGCCGC	GATTGGAAGTCGATTATAAT	Carson and Spicer 2003

Table A2. Accession numbers of sequences obtained from GenBank for phylogenetic reconstruction of genus *Spizella*.

Species	COI	Cyt-b	ATP6 – 8 / COIII	ND2	CR
S. arborea	DQ433190	**FR847858**	AF468627	AY138925	AY138908
S. atrogularis	DQ433192	HQ263417	AF468628	HQ263433	HQ263429
S. breweri	DQ434134	AF118236	AF468629	AF290121	HQ263427
S. pallida	DQ434759	FJ547279	AF468630	FJ547320	AF122949
S. passerina	DQ433193	FJ547278	AF468631	FJ547319	AY862849
S. pusilla	DQ434768	EF529933	AF468632	EF529824	HQ263428
S. taverneri	**FR847850**	AF118239	**FR847855 / FR847856 / FR847857**	—	—
S. wortheni	**FR847849**	**FR847851**	**FR847852 / FR847853 / FR847854**	HQ263432	HQ263426
A. bilineata	DQ433307	FJ547280	AF468635	FJ547321	HQ263430
J. hyemalis	DQ434618	EU325787	AF468624	AF447288	AY995313

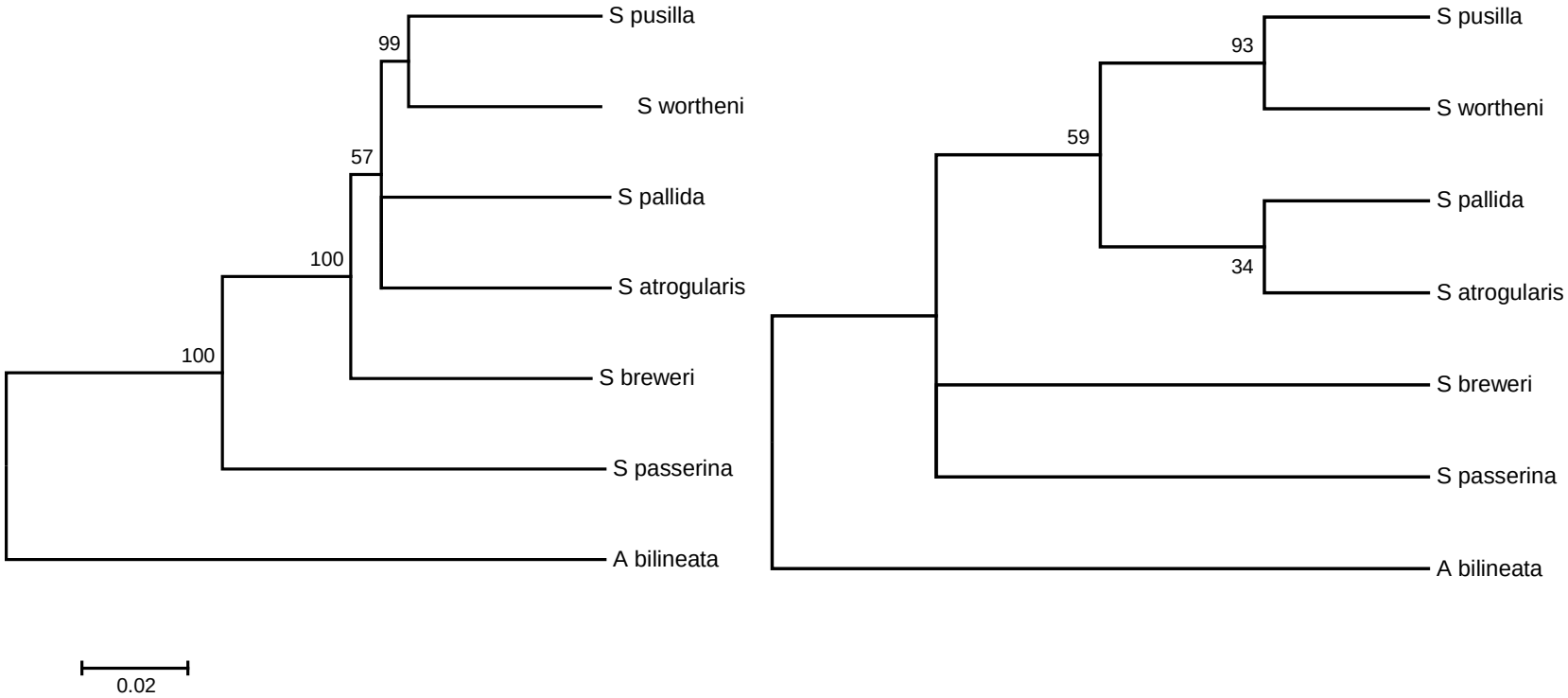

Figure A1. Bayesian (left) and maximum likelihood (right) tree topologies' constructed with a total length fragment of 4233 bp and rooted with *A. bilineata*, numbers are the nodal support values.

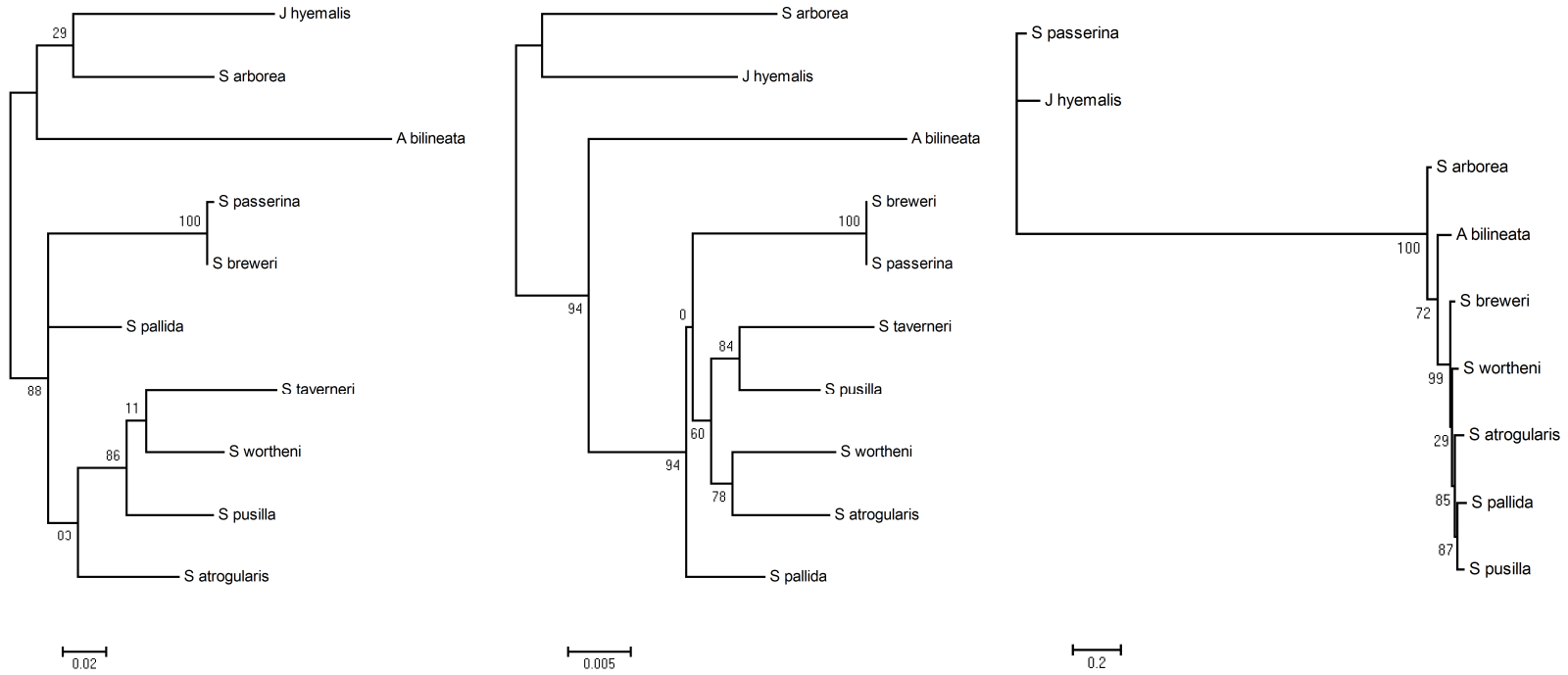

Figure A2. From left to right, tree topologies' for genes ATP6, CO3 and the mitochondrial control region, numbers are nodal support values after a Chi-squared aLRT carried out with PhyML.

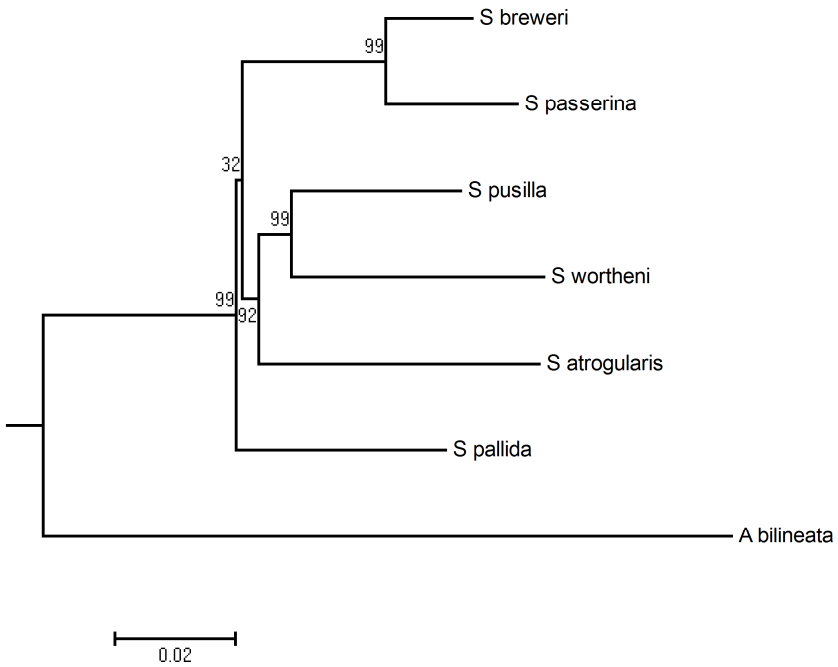

Figure A3. Maximum likelihood phylogram constructed with a fragment length of 2889 bp and excluding S. taverneri from the analysis, numbers are nodal support, note that low aLRT support was persistent in a central node.