

# The evolutionary ecology of a species ring: a test of alternative models

Jason F. SMYTH<sup>1</sup>, Michael A. PATTEN<sup>2</sup> and Christin L. PRUETT<sup>1\*</sup>

<sup>1</sup> Florida Institute of Technology, Department of Biological Sciences, Melbourne, FL 32901, U.S.A.;  
e-mail: cpruett@fit.edu

<sup>2</sup> University of Oklahoma, Oklahoma Biological Survey and Department of Biology, Norman, OK 73019, U.S.A.;  
e-mail: mpatten@ou.edu

Received 30 December 2014; Accepted 7 July 2015

**Abstract.** The formation of ring species might provide an explanation of how speciation can occur despite ongoing gene flow. However, few species fit all of the criteria of a classic ring species that formed via isolation by distance around a barrier. Population genetic analyses and ecological niche models were used to examine a ring of song sparrow (*Melospiza melodia*) subspecies that surround the Sierra Nevada in North America. Eight models were compared that included both geography-based and ecology-based scenarios of ring formation. Song sparrows do fit some aspects of a classic ring species that formed via expansion around a barrier; however, admixture rather than complete reproductive isolation occurred when populations met at the terminus of the ring in southern California. Niche-models show that variation among subspecies is likely to reflect adaptation to local conditions coupled with current limitations to gene flow across ecotones and that birds are likely to have expanded from a refugium in the southwestern United States. Given that simple isolation-based models often fail to explain many ring species patterns, alternative models that incorporate ecological factors might provide a better explanation of how most ring species formed. Isolation and subsequent partitioning of populations by ecotones can be important drivers of geographic variability in ring species.

**Key words:** ecological niche model, ecotone, *Melospiza melodia*, population genetics, song sparrow, speciation

## Introduction

In a ring species, the ranges of multiple subspecies connect around a geographic barrier. The two subspecies whose ranges meet at the terminal ends of the ring – considered to be the two subspecies that have diverged phenotypically to the greatest extent – are reproductively isolated but other subspecies continue to be connected via gene flow (Fig. 1; Mayr 1942, Irwin & Irwin 2002). The pattern of phenotypic variation in ring species implies that intraspecific variation can be substantial enough to lead to reproductive isolation and speciation (Irwin et al. 2001). There is debate regarding the exact classification of ring species (Patten & Pruett 2009), although there is agreement about two necessary criteria that must be met: 1) a series of morphologically (and presumably genetically) intermediate subspecies must be arranged in a ring, with gene flow only occurring between adjacent populations and not across the geographic barrier, and 2) the subspecies at the terminal ends must behave as biological species and thus exhibit reproductive isolation (Mayr 1942, Irwin & Irwin 2002, Patten & Pruett 2009).

Historically a single model for the evolution of a ring species had been proposed, yet Patten (2010) suggested four possible ways that ring species form (Fig. 1). In the classical I model – the sole model postulated in older texts (e.g. Mayr 1942) – a species expanding its geographic range bifurcates around a physical barrier, diverges as it travels two separate pathways, and reconnects on the far side of the physical barrier as biological species (Jordan 1905). The classical II model is predicated on a similar idea but begins with an expanding species encountering a physical barrier and circling it in one direction. By the time the expanding front reaches the starting point the front and the source have differentiated into biological species (Fig. 1, see Kuchta et al. 2009 for a similar model). Patten (2010) proposed two additional models (the *in situ* and ecological divergence models) given recent evidence for ecological speciation (e.g. Schluter 2009, Nosil 2012). In the *in situ* model, populations diverge across ecotones, with the steepest gradient in habitat between populations at the terminus of the ring that act as biological species. In

\* Corresponding Author

the ecological divergence model populations at the terminus of the ring originate and then diverge across a single sharp ecotone but differences among other subspecies (than those at the terminus) are the result of geographic distance as the expanding fronts move around the physical barrier (Fig. 1).

The primary differences between these ecology-based models and the classical geography-based model(s) are the timing of divergence across the terminus of the ring and the process that causes this divergence (Patten 2010). Hence, under these two classes of models different patterns of isolation by distance and population differentiation are expected as subspecies diverged. In the classical models, the two subspecies that act as biological species would be the most genetically distant despite their physical proximity due to historical isolation. In other words, there would be concordance between phenotypic and genotypic divergence. In the ecological divergence model, the two subspecies at the terminus of the ring would be genetically differentiated indicating isolation across the ecotone. Populations at the top of the ring would be expected to show admixture and ongoing gene flow. In the *in situ* model, the steepness of the ecotones yields different levels of isolation between neighboring populations as a result of natural selection rather than restricted gene flow at a geographic barrier. No subspecies pair is expected to differ genetically more than any other and populations will not show a signal of isolation with distance (Fig. 1). Several ring species have been described on the basis of morphology and geography, but with the advent of genetic techniques and quantitative means of analyzing phenotype and behavior it became apparent that many examples summarized by Mayr (1942) were not ring species. Irwin et al. (2001) reviewed 23 proposed ring species but in many cases there was evidence of gene flow across the alleged barriers or between populations that were thought to be reproductively isolated (Liebers et al. 2004, Packert et al. 2005, Joseph et al. 2008). One ring species proposed by Irwin et al. (2001), which was subsequently supported by additional research, is the greenish warbler (*Phylloscopus trochiloides*) complex (Irwin et al. 2005). However, a recent study found that there was geographical isolation among subspecies in more than one location with subsequent introgression during secondary contact rather than continual gene flow around the ring (Alcaide et al. 2014).

On the basis of phenotypic variation, the spatial arrangement of subspecies, and the presence of two highly differentiated subspecies that behave as biological species where their ranges meet, song

sparrows (*Melospiza melodia*) in the western United States appear to be a ring species (Fig. 2; Patten et al. 2004b, Patten & Pruett 2009, Patten 2010). Song sparrows have progressively intermediate morphology around the Sierra Nevada and Mojave Desert and a sharp morphological and genetic break at what is presumed to be the terminal end of the ring in the Coachella Valley of southern California (Fig. 2, Patten et al. 2004b). Genetic variation has been studied across many of the subspecies involved (Patten et al. 2004b, Pruett et al. 2008a, b, Wilson et al. 2011), but there has been no broad-scale evaluation of the patterns of genetic structure or gene flow around the putative species ring, let alone a test of competing models of how the ring formed. We used genetic data from song sparrows collected around the ring and developed ecological niche models based on documented occurrences to determine the pattern of geographic variation of song sparrow populations. Our goal was to address several important questions about this species and about how ring species evolve. Do song sparrows show the genetic signals expected of a classical ring species? If so, did the ring form as a result of bidirectional or unidirectional range expansion? If not, are ecological models a more parsimonious explanation for the formation of the ring?

## Material and Methods

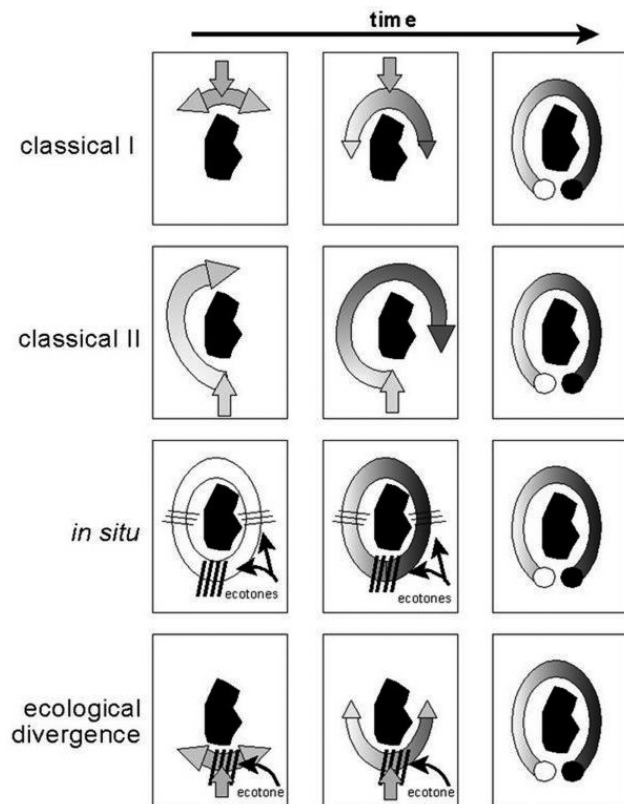
### Population genetics

During the summers of 2009 and 2010, we collected 120 song sparrows from six locations in the western United States in accordance with federal and state permitting requirements and under the supervision of an IACUC agreement (Fig. 3, Appendix I). We extracted DNA from tissues using a Qiagen QIamp DNA mini kit (Qiagen, Valencia, CA) and amplified six polymorphic microsatellite loci for each individual: Mme2, Mme3, Mme7, Mme12 (Jeffery et al. 2001), GF05 (Petren 1998), and EScu1 (Hanotte et al. 1994). Primers were labeled with fluorescent dyes and amplified products were genotyped on an ABI 3130 Genetic Analyzer (Applied Biosystems, Culver City, CA). Specimens of known genotype, that were determined in previous studies (Patten et al. 2004b), were analyzed along with newly genotyped birds to ensure compatibility with previously published datasets. Genotypes of birds from Riverside and the Salton Sea in California were added to our dataset from Patten et al. (2004b, Fig. 3). At least two locations from each of the subspecies surrounding the ring were examined in our analyses (Fig. 2). We used Arlequin (ver. 3.5, Excoffier et al. 2005) to calculate descriptive statistics including average observed heterozygosity

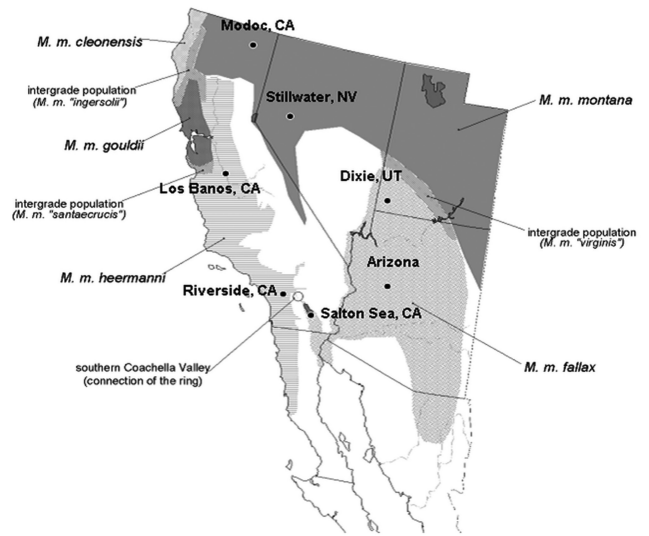
( $H_o$ ), average expected heterozygosity ( $H_e$ ), pairwise  $F_{ST}$ , and to test for Hardy-Weinberg equilibrium and linkage equilibrium ( $\alpha = 0.05$ ; 1000000 permutations, 10000 step burn-in). We corrected for multiple tests using the Benjamini & Hochberg (1995) method in SGOFF+ (Carvajal-Rodriguez & de Uña-Alvarez 2011). We used Fstat to calculate allelic richness (Goudet 1995) and Microchecker (Van Oosterhout et al. 2004) to ensure that there were no null alleles, stutter bands, or large allele dropout in the dataset.

We used GeneClass 2 to conduct assignment tests (Berry et al. 2004, Paetkau et al. 2004, Piry et al. 2004) to provide estimates of recent movement among populations. Individuals were considered immigrants if the probability of exclusion from their population of origin was greater than 0.95. They were then assigned to the population for which they had the highest posterior probability.

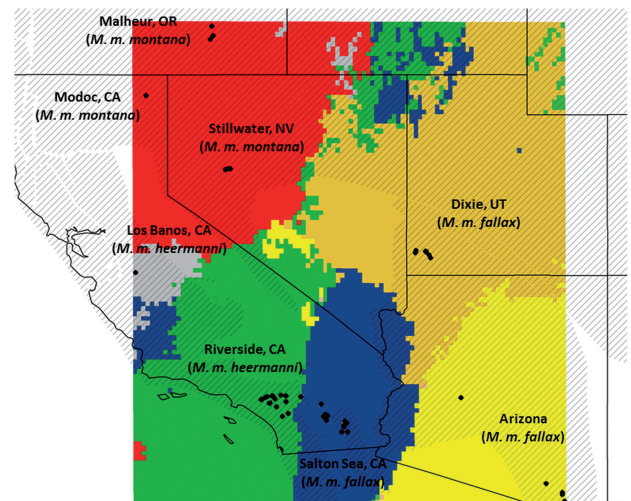
We used two methods to assess population structure. Individuals were clustered into K groups that exhibited Hardy-Weinberg and linkage equilibrium using Structure v. 2.3.4 (Falush et al. 2003). We used the



**Fig. 1.** Possible ring species formation patterns for song sparrows from Patten (2010). In each model the species encounters the geographic barrier, and then follows the arrows around the barrier. Population differentiation is indicated by changes in the shade of the arrows, and ecotones are indicated by series of hash marks. Small black arrows indicate direction of initial colonization of the ring.



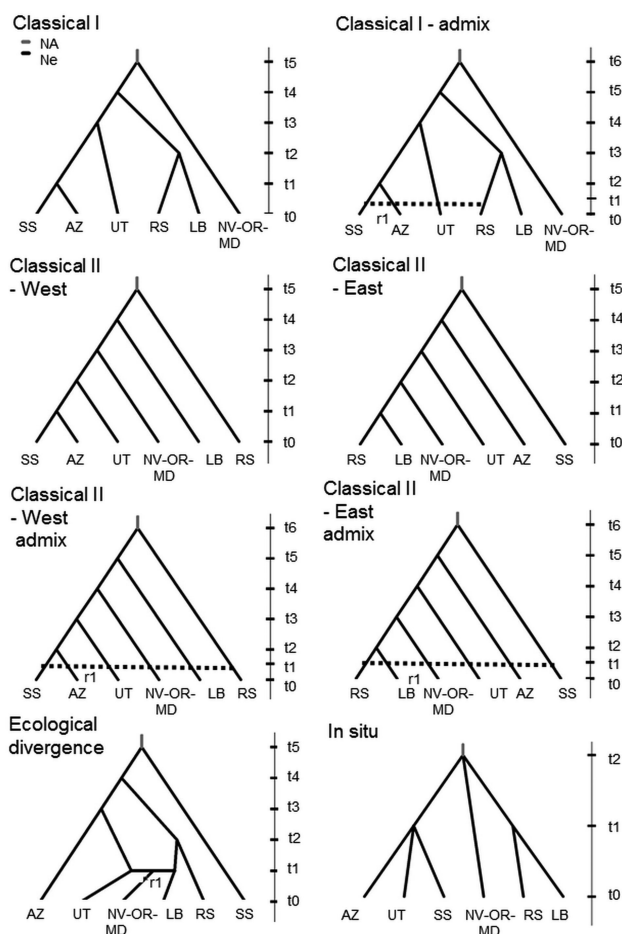
**Fig. 2.** The distribution of subspecies around the song sparrow ring (Patten 2010) including genetic sampling locations. All locations shown except for Malheur which is just north of Modoc in Oregon and is within the distribution of the subspecies *Melospiza melodia montana*.



**Fig. 3.** Map of song sparrow geographic distribution (hatched area) in western United States with Geneland results showing genetic clusters. Each color indicates a different genetic cluster and black dots correspond to individual birds sampled. The location of transitions between colors does not necessarily reflect the exact location of the genetic break between samples.

admixture model with location information as the prior and the correlated allele frequency model. We performed 10 runs of Structure for each  $K = 1-10$  using a 200000 step burn-in and 600000 step MCMC and used the method of Evanno et al. (2005) to determine the most probable number of genetic clusters in Structure Harvester (Earl & vonHoldt 2012). Clump 1.1.2 (Jakobsson & Rosenberg 2007) and Distruct 1.1 (Rosenberg 2004) were used to create graphical displays





**Fig. 4.** Eight models of divergence (Fig. 1) among song sparrows compared using approximate Bayesian computation. Ne corresponds to effective population sizes of each population, NA is effective population size of ancestral population, t1-t5 indicate convergence times for lineages, t6 shows timing of admixture between populations, and r1 indicates admixture rate between populations. Dashed lines indicate admixture between SS and RS. SS – Salton Sea, AZ – Arizona, UT – Dixie, RS – Riverside, LB – Los Banos, NV-OR-MD – grouping of locations at Stillwater, Malheur, and Modoc (Table 1, Fig. 3).

based on the Structure output. We then analyzed data in Geneland 4.0.3 (Guillot et al. 2005, Guillot 2008), which uses a Bayesian MCMC model that clusters individuals into groups that are in Hardy-Weinberg and linkage equilibrium and includes spatial information in the model. We used the correlated allele frequency model and treated spatial coordinates as uncertain (set to 1 km) because some samples were taken from the same location. We initially ran the MCMC at least three times with minimum K of 1 and maximum of 10 using different numbers of replicates and burn-in lengths to discover the length of chains necessary for convergence. We determined that  $10^6$  MCMC replicates with a burn-in of 20000 per run were sufficient.

We chose to test eight competing models of ring formation (Fig. 4), five models based on previous

ring formation scenarios (Fig. 1) and three additional models that included secondary contact across the terminus of the ring in southern California (Figs. 2, 4), a phenomenon suggested as commonly occurring in other ring species (Irwin et al. 2001, Alcaide et al. 2014). We grouped the three sampling locations at the northern-end of the ring based on clusters identified using Geneland (Fig. 3). Competing models include 1) Classical I with locations at the northern-end of the ring (Fig. 3) ancestral to populations on the east and west sides of the ring, 2) Classical I with admixture indicating secondary contact across the terminus of the ring, 3) Classical II – West or Classical II – East with one of the two populations at the terminus of the ring in Southern California ancestral to all other populations with expansion around the west or east side of the ring, 4) Classical II models with admixture that include secondary contact across the terminus of the ring after expansion of the ring is complete, 5) Ecological divergence where populations expand from the terminus around both sides of the ring with admixture at the northern-end of the ring, and 6) *in situ* model with all subspecies diverging at a similar time (Fig. 4). We used a coalescence-based approximate Bayesian computation (ABC) approach in DIYABC v. 2.0 (Cornuet et al. 2014) for these analyses. For each model, we used a series of broad uniform priors to determine the appropriate priors; we used priors of 10-200000 for the effective population size and 10-50000 for ancestral effective population size (NA). Effective sizes for each population were assumed to be similar based on preliminary runs. For the timing of recent divergence events (t1-t3) among clusters we used a uniform prior of 10-20000 generations and for earlier events (t4-t6) increased the prior to 10-100000 which bounds the timing of late Pleistocene glaciations with the generation time of song sparrows (2.56 years, Pruett et al. 2008a). Four models (Fig. 4) exhibited admixture and the prior for the rate of admixture was set to 0.001-0.999. We used a generalized mutation model (Estoup et al. 2002) with a uniformly distributed prior of  $1.00 \times 10^{-4}$ - $1.00 \times 10^{-3}$  substitutions per generation for the mean mutation rate.

We simulated  $8 \times 10^5$  datasets for each model. Six summary statistics were generated for comparisons based on normalized Euclidean distances (Beaumont et al. 2002) between observed and simulated datasets. These statistics include one sample statistics, mean number of alleles and mean genic diversity, and two sample statistics, mean genic diversity,  $F_{ST}$ , classification index, and shared allele distance.

Datasets with the smallest Euclidean distances ( $N = 8000$ ) for each model were retained and used to build posterior parameter distributions. To determine the most probable model, we used the logistic regression approach in DIYABC to estimate posterior probabilities for each model and 90 % credibility intervals. We also estimated Type I and Type II error rates to evaluate confidence in model choice. Posterior distributions of parameters for the most probable model were estimated using a local linear regression estimate of 8000 datasets closest to the actual dataset. We assessed the performance of the ABC method to estimate parameters by simulating 114994 pseudo-observed datasets and computed the average relative bias and the square root of the relative mean square error (Bertorelle et al. 2010).

### Ecological niche models

Core occurrence data for our ecological niche models (ENMs) were locales where we obtained blood samples (Patten et al. 2004b) or specimens (this study) around the species ring. We supplemented our data with breeding season locations in VertNet (<http://www.vertnet.org/>), an online database of museum specimens, song recordings, and census data. We

included no occurrences with duplicate locales but instead included only those with distinct points at least 1 km apart. Our compilation yielded 63 *M. m. fallax*, 73 *M. m. heermanni*, and 99 *M. m. montana* occurrences (Fig. 2, Appendix II). We constructed separate ENMs both to generate predicted geographic ranges of three key subspecies (noted above) around the ring. We established a 30 arc second-grid (slightly  $< 1$  km per side) across California, Nevada, Utah, and Arizona and culled a suite of biotic and abiotic variables: annual precipitation (mm); mean January temperature ( $^{\circ}\text{C}$ ); mean July temperature ( $^{\circ}\text{C}$ ); elevational range (m); the Normalized Difference Vegetation Index (NDVI); and topographic wetness index (Besnard et al. 2013). Variables were obtained from Bioclim (<http://worldclim.org/>, Hijmans et al. 2005) and clipped to the four-state region. We also hindcasted the distributions of each subspecies during the late-Pleistocene ( $\sim 22000$  ybp) based on temperature and precipitation profiles (MPI-ESM-P global circulation model) as there are no NDVI or moisture data from the Pleistocene and there is no elevation model from that period. We also examined models based on mid-Holocene bioclimatic data but models did not differ from present-day niche models.

**Table 1.** Average heterozygosity and allelic richness based on six microsatellite loci for song sparrow populations. N is number of samples,  $H_E$  is expected heterozygosity,  $H_O$  is observed heterozygosity, Ar is allelic richness, all plus or minus standard error.

	N	$H_E$	$H_O$	Ar
Riverside	24	0.83+/-0.08	0.82+/-0.11	8.03+/-2.16
Salton Sea	29	0.83+/-0.06	0.86+/-0.12	7.66+/-1.64
Arizona	15	0.78+/-0.08	0.72+/-0.16	6.49+/-1.09
Stillwater	20	0.79+/-0.20	0.77+/-0.19	8.42+/-3.14
Malheur	20	0.81+/-0.13	0.76+/-0.12	8.13+/-2.00
Modoc	20	0.78+/-0.19	0.73+/-0.15	7.75+/-3.00
Dixie	10	0.85+/-0.10	0.78+/-0.17	7.59+/-3.02
Los Banos	20	0.82+/-0.15	0.77+/-0.19	8.45+/-2.59

**Table 2.** Pairwise  $F_{ST}$  values for each song sparrow population. Bold values are significantly different from zero after multiple test correction (Benjamini & Hochberg 1995).

	Riverside	Salton	Arizona	Stillwater	Malheur	Modoc	Dixie
Salton	0.008						
Arizona	<b>0.039</b>	<b>0.047</b>					
Stillwater	<b>0.024</b>	<b>0.051</b>	<b>0.014</b>				
Malheur	<b>0.021</b>	<b>0.036</b>	<b>0.025</b>	0.007			
Modoc	<b>0.025</b>	<b>0.060</b>	<b>0.050</b>	0.005	0.010		
Dixie	<b>0.029</b>	<b>0.037</b>	<b>0.045</b>	<b>0.053</b>	<b>0.063</b>	<b>0.061</b>	
Los Banos	0.004	<b>0.027</b>	<b>0.031</b>	0.008	<b>0.021</b>	<b>0.023</b>	<b>0.038</b>

**Table 3.** Prior distributions and posterior probabilities of parameters for the Classical II – East with admixture model (Fig. 4) based on approximate Bayesian computation analysis of song sparrows. Time in years are based on a generation time of 2.56 yrs. MRB indicates mean relative bias of mode and RRMSE indicates square root of the relative mean square error of the mode for estimates calculated from simulations.

Parameter	Prior	Mode	5 %	95 %	MRB	RRMSE
Ne	10-200000	141000	79500	184000	-0.04	0.50
NA	10-50000	1380	519	36700	2.70	24.74
t1	10-20000	2637	835	22733	-0.09	5.87
t2	10-20000	9011	3174	36096	0.07	1.91
t3	10-20000	22272	8704	47616	0.28	0.67
t4	10-100000	50176	26368	124928	-0.10	0.91
t5	10-100000	65792	35328	173568	0.04	0.65
t6	10-100000	86016	48128	228864	0.21	0.37
r1	0.001-0.999	0.49	0.12	0.87	0.46	8.77

Pleistocene data are at a coarser scale (2.5' rather than 30''), so we created a grid at the same 30 arc second resolution – essential to hindcast from current to past models – by simply dividing 2.5' cells into a set of 30'' cells. We reported the 90<sup>th</sup> percentile hindcast ranges for each current subspecies' model.

We used the MaxEnt (Maximum Entropy) algorithm (version 3.3.3k) with default settings (see Radosavljevic & Anderson 2014) to create our ENMs. The default settings minimized spatial autocorrelation and overfitting and yielded area under the receiver operating characteristic curve (AUC) that is used to evaluate the model performance (i.e. AUC > 0.75, Elith 2002). We used 5000 maximum iterations to ensure that the model had sufficient convergence time so that it would not under- or over-predict relationships. We created response curves and used the median output for analyses to assess to what extent ecological niche models differed for each subspecies.

## Results

### Population genetics

All loci for each population were in linkage equilibrium and in Hardy-Weinberg equilibrium except for locus Escu1 at Salton Sea and Mme7 at Modoc that both showed an excess of homozygotes; however, there was no evidence of stuttering, large allele dropout, or null alleles. We found high genetic diversity in all populations; average expected heterozygosity varied between 0.78 and 0.85 and average observed heterozygosity varied between 0.72 and 0.86 (Table 1). Average allelic richness across all populations was 7.81 (range 6.49 to 8.45) alleles per locus. These values are similar to those reported in other studies of song sparrows (Chan & Arcese 2003, Patten et al. 2004b, Pruett et al. 2008a, b) using the same markers.

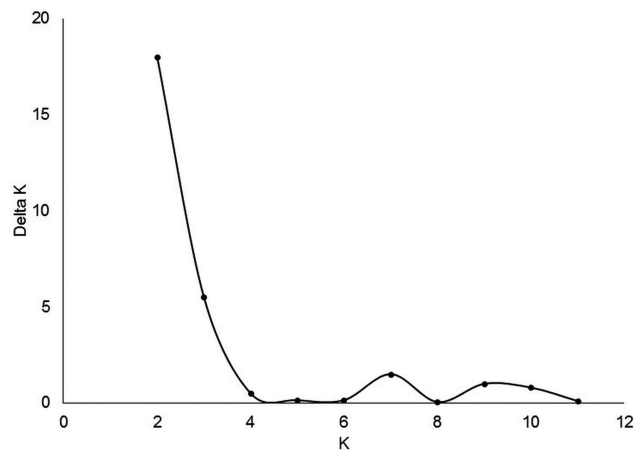
Most pairwise comparisons between sampling locations had  $F_{ST}$  values that differed significantly from zero after multiple test correction (Table 2). Comparisons that did not differ included Riverside – Salton Sea, Riverside – Los Banos, and comparisons among populations at the northern side of the ring, Modoc, Malheur, Stillwater, and Los Banos (Table 2, Fig. 3). The Utah and Arizona locations differed significantly in comparisons with all other locations. All individuals assigned to their population of origin except for one bird from Malheur that did not assign to any population with a high probability (highest probability was 0.097 for Malheur). Two genetic clusters within the dataset were identified using Structure (Fig. 5). Locations at the northern portion of the ring (Stillwater, Malheur, and Modoc) formed one cluster while the Salton Sea population formed the other cluster with intervening populations (Riverside and Los Banos on the western-side of the ring; Arizona and Dixie on the eastern-side of the ring) showing admixture between the two groups (Fig. 6). Six clusters were identified using Geneland (Fig. 3). Each population except for Modoc, Malheur, and Stillwater were identified as being in a separate genetic cluster. Multiple independent runs of the program always supported these groupings except for a few instances when Riverside and Los Banos were in the same cluster and when Los Banos grouped with Malheur, Modoc, and Stillwater in a cluster. In every instance, Riverside and Salton Sea were split into separate groups along the ecotone which currently separates these populations (Fig. 2, Patten et al. 2004a, b). The Classical II – East model with admixture (Fig. 4) received the most support showing posterior probabilities of 0.615 (0.556-0.674); the second most supported model, Classical II – East, had a posterior



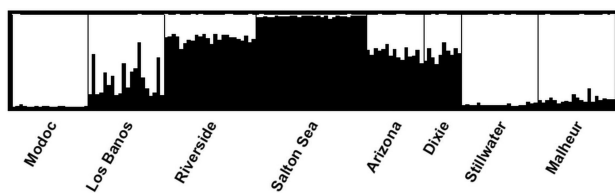
probability of 0.192 (0.099-0.285). Type I and Type II error rates for the Classical II – East model with admixture were, 0.119 for Type I and 0.152-0.204 for Type II. This indicates that there is sufficient power in the dataset to discriminate among models. Posterior probabilities of parameters for the Classical II – East model with admixture indicate that current effective population sizes ( $N_e$ ) are large in comparison with ancestral size ( $N_A$ , Table 3) suggesting an increase in size over time. The divergence times among populations ( $t_2$ - $t_6$ ) are within the Pleistocene (Table 3) with the most recent divergence occurring among populations on the west side of the Sierra Nevada in California ( $t_2$ , mode 9011 ybp) and the oldest among populations on the east side of the ring ( $t_6$ , mode 86016 ybp). However credibility intervals overlap broadly (Table 3), suggesting a relatively uniform expansion around the ring rather than a series of abrupt transitions. The timing of admixture at the terminus of the ring in southern California indicates moderate admixture (Table 3) during the late Pleistocene or Holocene ( $t_1$ , mode 2637 ybp; CI: 835-22733 ybp).

#### Ecological niche models

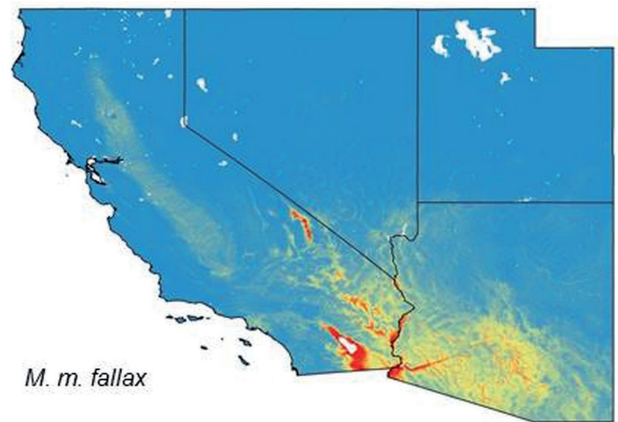
Models created for each subspecies fit the test data well (AUC: *M. m. fallax* = 0.94, *M. m. heermanni* = 0.97,



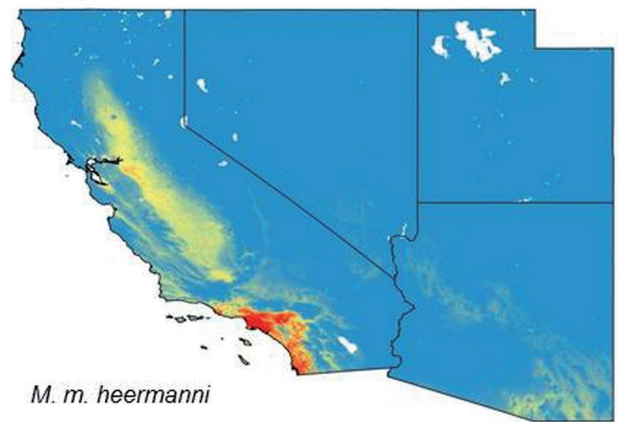
**Fig. 5.** Output from Structure runs indicating K = 2 is the most likely number of genetic clusters using the Evanno et al. (2005) method.



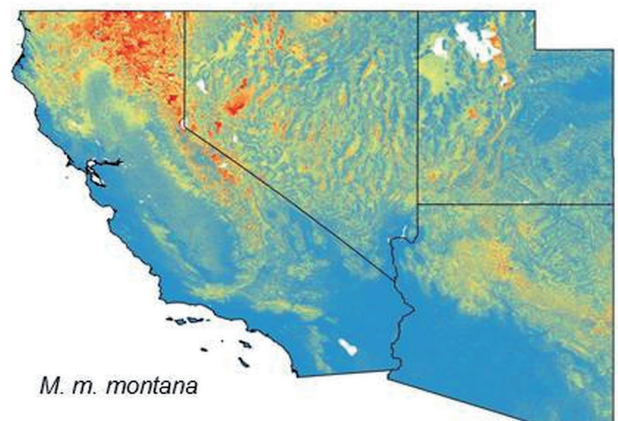
**Fig. 6.** DISTRUCT diagram of Structure output of two genetic clusters for song sparrows. Colors correspond to proportion of membership of each individual in each cluster. Individuals with two colors are likely to be admixed.



*M. m. fallax*



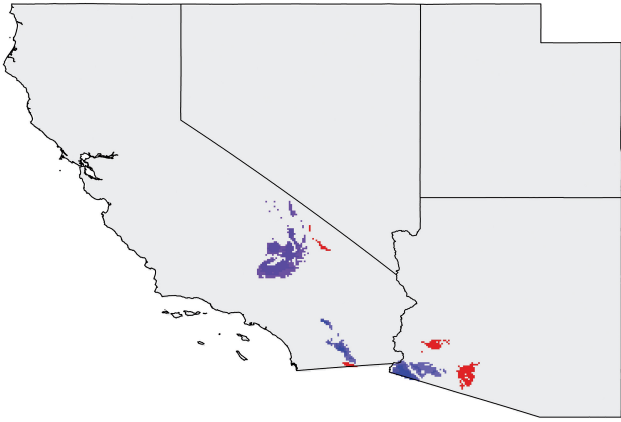
*M. m. heermanni*



*M. m. montana*

**Fig. 7.** Ecological niche models (from MaxEnt) for three subspecies of the song sparrow in western North America. Warm colors correspond to a high probability of predicted occurrence (red is the highest), whereas cool colors correspond to a low probability (blue is lowest). Models imply ecological segregation among the subspecies: rather than known occurrences for a given subspecies predicting a wide geographic range, instead predictions conform well to the known distributions of the subspecies (compare with Fig. 2).

*M. m. montana* = 0.90) and match current geographic distribution of each subspecies (Figs. 2, 7). Mean January temperature ( $^{\circ}\text{C}$ ) was a key environmental correlate for each subspecies, accounting for 43.5 %, 35.5 %, and 37.5 %, respectively, of ENM variation. Elevational



**Fig. 8.** Ecological niche model for the three subspecies of song sparrow in western North America during the late-Pleistocene. Purple areas correspond to the subspecies *Melospiza melodia montana*, blue areas to *M. m. heermanni*, and red areas to *M. m. fallax*.

range (m) accounted for a similarly high amount, 36.7 %, of the ENM for *M. m. heermanni*, whereas the topographic wetness index accounted for 28.4 % for *M. m. montana* and July temperature (°C) accounted for 22.4 % for *M. m. heermanni*. The only other value near 20 % was the 18.1 % for annual precipitation (mm) for *M. m. fallax*, a desert taxon confined to riparian belts and lakes. The niche envelope differs among the three subspecies, such that areas with predicted high suitability for one subspecies do not correspond to predicted high suitability for the other two subspecies (Fig. 7). On the basis of climate, vegetation, and topographic drivers, *M. m. fallax* was predicted to occur chiefly around the Salton Sea in southeastern California, along the lower River Colorado on the border of California and Arizona, and along rivers in central and southern Arizona, whereas *M. m. heermanni* was predicted to occur in coastal southern California as well as that state's Central Valley and *M. m. montana* at lower elevations in the Great Basin and Intermountain West, generally east of the Sierra Nevada (Fig. 6). Pleistocene distributions of all three of the current subspecies were hindcast to occur in what are now the Mojave or Sonoran Deserts of southern California and Arizona (Fig. 8).

## Discussion

### *Ring formation*

The song sparrow ring is unlikely to have formed via a strict geography-based model (Classical I or Classical II) for the evolution of a ring species, a model in which populations diverge at the leading edge of an advancing front around a geographic barrier and are reproductively isolated at secondary contact (Jordan 1905, Mayr 1942). If either Classical model explained the current pattern of subspecies distribution among song sparrows, we

would expect that the two populations at the terminus of the ring, Riverside and Salton Sea (Figs. 2, 3) would be the most genetically divergent. However, the Riverside and Salton Sea (Figs. 2, 3) locations are similar genetically at neutral loci ( $F_{ST} = 0.008$ ) despite differing the most phenotypically (Patten & Pruett 2009), indicating that populations have not experienced long-term reproductive isolation. In addition, we tested the Classical I model with admixture (Fig. 4) on secondary contact between Riverside and Salton Sea and found no support for this model.

The *in situ* model of ring formation (Patten 2010, Fig. 1) is also not supported by the genetic data. Expectations for this model include genetic drift creating a pattern of random genetic divergence caused by a reduction in gene flow among subspecies across ecotones and a close genetic association among populations within subspecies and greater divergence among populations that are different subspecies. However, we found that birds that are morphologically different subspecies (Fig. 2) are genetically similar (Table 2) including the close association between Riverside (*M. m. heermanni*) and Salton Sea (*M. m. fallax*) and between Stillwater (*M. m. montana*) and Los Banos (*M. m. heermanni*). In some instances populations from different subspecies are genetically more similar to populations in other subspecies; for example, Salton Sea is more similar to Riverside than to the Arizona or Dixie populations (Table 2).

The model of ecological divergence as put forward by Patten et al. (2010, Fig. 1) is not supported by the ABC analysis likely due to the close genetic relationship between Salton Sea and Riverside and the expectation that secondary contact would occur at the northern-end of the ring (Modoc, Malheur, and Stillwater) rather than the southern-end of the ring (Salton Sea and Riverside, Figs. 1, 4). Thus, we would expect that populations at the northern-end of the ring would be genetically similar to populations to the east (Dixie) and to the west (Los Banos) as these areas would have served as sources for the admixed populations. Although Los Banos is not significantly differentiated from northern populations, Dixie and Arizona are differentiated from all other locations suggesting that birds in this area have been isolated from other song sparrow populations and admixture has not occurred recently (Table 2).

The most likely model of ring formation is via a process of isolation and subsequent ecological divergence after secondary contact (Fig. 4). In this model, populations gradually diverged during colonization around the Sierra Nevada, as in a Classical II model (Fig. 1), but exhibited admixture during secondary contact at the



terminus of the ring in southern California (Fig. 4). For this model, we expected a close genetic association between the Riverside and Salton Sea populations due to recent admixture with gradual divergence around the ring. Our findings support this model: there is a close genetic relationship between the populations at the terminus of the ring, gradual divergence around the ring (Fig. 6), and a signal of historical admixture with birds likely dispersing from Salton Sea to western Riverside County or *vice versa*. Thus, pronounced morphological differences between birds at the terminus of the ring are probably the result of local selective pressures leading to rapid differentiation among subspecies, as suggested in other studies of song sparrows (Chan & Arcese 2003, Pruett et al. 2008a, Wilson et al. 2011), rather than long-term isolation. Our proposal of local selective pressures is supported by the ecological niche models we created, each of which predicted geographic range of a particular subspecies well to the general exclusion of other subspecies, suggesting strong ecological differentiation. The break between the predicted ranges of *M. m. fallax* and *M. m. heermanni* in the contact zone in southern California is particularly abrupt (Fig. 7) and closely matches the Geneland results (Fig. 3). A concern with using model-based ABC analyses is that none of the models fit the data well and that the best model does not represent the history of the group examined (Templeton 2009). This is why choosing appropriate models based on prior knowledge is important (Csilléry et al. 2010). Thus, we chose models based on previously published ring species models (Patten 2010) and knowledge about the geographical distribution of song sparrow subspecies (Patten & Pruett 2009). Although we cannot be certain that our models encompass all possible scenarios of ring formation, our other genetic analyses ( $F_{ST}$  comparisons, Structure, Geneland, and assignment tests) support the model selected in ABC analyses.

#### *Source of the ring*

Researchers using mitochondrial (mt) DNA-based data discovered high diversity within song sparrow populations throughout the western continental United States and a lack of divergence among populations. On the basis of these data, the song sparrow is thought to have expanded rapidly into their current geographic distribution at the end of the Pleistocene glacial cycles (Zink & Dittmann 1993, Fry & Zink 1998). We also found high genetic diversity within populations (Table 1), large effective population sizes, and low-levels of genetic divergence among populations (Table 2). Thus, formation of the subspecies ring is

probably a Pleistocene-Holocene phenomenon with a large number of birds expanding into their current distributions. Where did these colonists originate?

Several locations in western North America have been postulated as glacial refugia for a variety of terrestrial organisms. Areas near the song sparrow ring include the Baja California peninsula, northern California, and locations in coastal British Columbia and Alaska (Swenson & Howard 2005, Waltari et al. 2007, Topp & Winker 2008). Previous studies have suggested that southern California and coastal British Columbia were glacial refugia for the song sparrow (Fry & Zink 1998, Pruett & Winker 2005, Pruett et al. 2013). Thus, it is possible that more than one location could have served as a source for song sparrow populations that surround the ring. However, we found support for colonization from a single southern refugium in both the genetic and late-Pleistocene ecological niche models. The ABC analysis shows expansion from a southern refugium as the most likely model of ring formation. In addition, hindcasting of the distributions of the three primary subspecies that surround the ring shows that their ancestors likely were found in areas in southern California and Arizona (Fig. 8), which matches other studies that have supported a refugium in this area (Waltari et al. 2007). However, a major assumption of using this method to determine historical range is that these subspecies were adapted to the same conditions historically as they are currently and thus their niches have not changed over time. However, this does not appear to be the case for song sparrows given the relatively rapid expansion of song sparrows into many niches in North America (Fry & Zink 1998, Pruett & Winker 2005, Pruett et al. 2008b).

Fry & Zink (1998) reported that some birds from the Salton Sea area were sister to all other song sparrows on phylogenetic trees of mtDNA sequences. This relationship suggests that western song sparrows originated in the Salton Sea region. However, the same study found that other birds sampled from the Salton Sea possessed haplotypes nested within the phylogenetic tree (Fry & Zink 1998), suggesting a degree of randomness in the placement of haplotypes within the phylogeny that could be due to gene flow or rapid range expansion.

#### *Maintenance of the ring*

All populations appear to be currently isolated from one another. Only a single bird could not be assigned with high likelihood to its population of origin with that bird probably originating in an unsampled location. Our findings suggest a pattern of historically

high gene flow during colonization with subsequent barriers limiting gene flow. Using Geneland, we found six genetic groupings that clustered populations by geography and in some instances by subspecies; e.g. the three populations within the subspecies *M. m. montana* (Fig. 3). Geneland appears to be able to identify closely related yet diagnosably distinct genetic groups better than models that do not use spatial information such as Structure (Guillot et al. 2005). However, caution should be exercised in interpreting the results, especially when genetic clusters do not differ significantly from one another using another method such as pair-wise  $F_{ST}$  and are isolated by distance (Guillot et al. 2009). The Los Banos population formed a single cluster in most analyses but also grouped with another population of the same subspecies (Riverside, *M. m. heermanni*) in a limited number of Geneland runs, and in one instance grouped with the *M. m. montana* group. Thus, the placement of this population is difficult for this analysis as the close relationship between Los Banos and these groups is apparent in comparisons of  $F_{ST}$ . This finding supports the idea that isolation by distance rather than a sharp habitat break is the strongest driver of genetic divergence among populations on the western side of the ring.

On the eastern side of the ring, where the Sierra Nevada and Mojave Desert act as isolating barriers, we found different genetic clusters for each population within *M. m. fallax* (Fig. 3). In addition, the Dixie and Arizona populations were the most differentiated based on  $F_{ST}$  comparisons with all other locations. This pattern of divergence corresponds to the highly fragmented distribution of this subspecies that inhabits narrow riparian corridors that likely restrict gene flow among populations (Patten et al. 2004b, Patten & Pruett 2009, pers. observ.). Nevertheless, on the basis of the Structure result of two genetic clusters, it is apparent that these populations are similar to other song sparrow locations and likely were historically connected.

At the terminus of the ring, in the San Geronio Pass that separates western Riverside County from the Salton Sea, there is a sharp ecotone (Patten et al. 2004a) that appears to be associated with current isolation between *M. m. heermanni* and *M. m. fallax*; subspecies that differ strikingly in phenotype (Patten et al. 2004b, Patten & Pruett 2009). In all Geneland runs, we found a break between these populations that corresponds geographically to the location of the ecotone (Figs. 2, 3). However, these populations are genetically similar at neutral loci, with a pairwise  $F_{ST}$  not significantly differing from zero. Coupled with the assignment tests which indicated a lack of gene flow

and the Bayesian analyses which indicate historical connectivity, we postulate that the populations at the terminus of the ring were historically connected but are currently reproductively isolated from one another as suggested by mate choice and mate signaling experiments (Patten et al. 2004b). A question remains as to why the ecotone did not serve as a barrier during secondary contact formation. One possible explanation is that during secondary contact the ecotone across the San Geronio Pass was not as steep, although myriad lines of evidence (see Patten et al. 2004a) and our niche models imply it is extremely steep now. During the Last Glacial Maximum, the habitat in southern California was markedly wetter and cooler than at present (Van Devender & Spaulding 1979). Thus, the barrier to movement across the ecotone has changed through time and possibly become increasingly more difficult for birds to cross. In addition, birds on each side of the ecotone have adapted to two strikingly different habitat types (Patten et al. 2004b); this is apparent given the marked differences in morphology between the two subspecies. Local adaptation could have reduced the likelihood that migrants or their offspring could persist in different habitats (Postma & van Noordwijk 2005, Cheviron & Brumfield 2009). We found that the song sparrow ring was not formed by a classic, geography-based mechanism alone but by a mechanism that includes ecological divergence across an ecotone. In addition, our data suggest that the evolution and maintenance of the song sparrow ring is likely to be a result of limited current gene flow and the presence of sharp environmental breaks in habitat across the western United States. Additional sampling across all putative ecotones is needed to further clarify the degree of isolation among song sparrow populations along each break. In conclusion, our results support aspects of ecology-based models of ring formation, a departure from Modern Synthesis theory (Mayr 1942). Alternative, ecological-based models of ring formation (or models combining ecology and geography) of ring formation might provide a better explanation of how many ring species patterns form, especially given that most putative ring species do not fit a classical geography-based formation pattern (Irwin et al. 2001).

### Acknowledgements

We thank the Florida Institute of Technology for financial support and archiving of voucher specimens and tissue samples. We thank Malheur National Wildlife Refuge, Modoc National Wildlife Refuge, Los Banos Wildlife Area, and the Hassayampa River Preserve for access to collecting locations. Our thanks to T.D. Fagin and B.D. Smith-Patten for help with environmental data layers and advice on our niche models.

## Literature

- Alcaide M., Scordato E.S.C., Price T.D. & Irwin D.E. 2014: Genomic divergence in a ring species complex. *Nature* 511: 83–85.
- Beaumont M.A., Zhang W. & Balding D.J. 2002: Approximate Bayesian computation in population genetics. *Genetics* 162: 2025–2035.
- Benjamini Y. & Hochberg Y. 1995: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* 57: 289–300.
- Berry O., Tocher M.O. & Sarre S.D. 2004: Can assignment tests measure dispersal? *Mol. Ecol.* 13: 551–561.
- Bertorelle G., Benazzo A. & Mona S. 2010: ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Mol. Ecol.* 19: 2609–2625.
- Besnard A.G., La Jeunesse I., Pays O. & Secondi J. 2013: Topographic wetness index predicts the occurrence of bird species in floodplains. *Divers. Distrib.* 19: 955–963.
- Carvajal-Rodriguez A. & de Uña-Alvarez J. 2011: Assessing significance in high-throughput experiments by sequential goodness of fit and q-value estimation. *PLoS ONE* 6: e24700.
- Chan Y. & Arcese P. 2003: Morphological and microsatellite differentiation at a microgeographic scale. *J. Evol. Biol.* 16: 939–947.
- Cheviron Z.A. & Brumfield R.T. 2009: Migration-selection balance and local adaptation of mitochondrial haplotypes in rufous-collared sparrows (*Zonotrichia capensis*) along an elevational gradient. *Evolution* 63: 1593–1605.
- Cornuet J.M., Pudlo P., Veyssier J., Dehne-Garcia A., Gautier M., Leblois R., Marin J.M. & Estoup A. 2014: DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* 30: 1187–1189.
- Csilléry K., Blum M.G.B., Gaggiotti O.E. & Francois O. 2010: Approximate Bayesian Computation (ABC) in practice. *Trends Ecol. Evol.* 25: 410–418.
- Earl D.A. & vonHoldt B.M. 2012: Structure Harvester: a website and program for visualizing Structure output and implementing the Evanno method. *Conserv. Genet. Resour.* 4: 359–361.
- Elith J. 2002: Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants. In: Ferson S. & Burgman M. (eds.), *Quantitative methods for conservation biology*. Springer, Berlin: 39–58.
- Estoup A., Jarne P. & Cornuet J.M. 2002: Homoplasmy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Mol. Ecol.* 11: 1591–1604.
- Evanno G., Regnaut S. & Goudet J. 2005: Detecting the number of clusters of individuals using the software Structure: a simulation study. *Mol. Ecol.* 14: 2611–2620.
- Excoffier L., Laval G. & Schneider S. 2005: Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1: 47–50.
- Falush D., Stephens M. & Pritchard J.K. 2003: Inference of population structure using multi-locus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587.
- Fry A.J. & Zink R.M. 1998: Geographic analysis of nucleotide diversity and song sparrow (Aves: Emberizidae) population history. *Mol. Ecol.* 7: 1303–1313.
- Goudet J. 1995: Fstat (Version 1.2): a computer program to calculate F-statistics. *J. Hered.* 86: 485–486.
- Guillot G. 2008: Inference of structure in subdivided populations at low levels of genetic differentiation: the correlated allele frequencies model revisited. *Bioinformatics* 24: 2222–2228.
- Guillot G., Estoup A., Mortier F. & Cosson J.F. 2005: A spatial statistical model for landscape genetics. *Genetics* 170: 1261–1280.
- Guillot G., Leblois R., Coulon A. & Frantz A.C. 2009: Statistical methods in spatial genetics. *Mol. Ecol.* 18: 4734–4756.
- Hanotte O., Zanon C., Pugh A., Grieg C., Dixon A. & Burke T. 1994: Isolation and characterization of microsatellite loci in a passerine bird: the reed bunting *Emberiza schoeniclus*. *Mol. Ecol.* 3: 529–530.
- Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G. & Jarvis A. 2005: Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25: 1965–1978.
- Irwin D.E., Bensch S., Irwin J.W. & Price T. 2005: Speciation by distance in a ring species. *Science* 307: 414–416.
- Irwin D.E. & Irwin J.H. 2002: Circular overlaps: rare demonstrations of speciation. *Auk* 119: 596–602.
- Irwin D.E., Irwin J.H. & Price T. 2001: Ring species as bridges between microevolution and speciation. *Genetica* 112/113: 223–243.
- Jakobsson M. & Rosenberg N.A. 2007: Clump: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23: 1801–1806.
- Jeffery K.J., Keller L.F., Arcese P. & Bruford M.W. 2001: The development of microsatellite loci in the song sparrow, *Melospiza melodia* (Aves) and genotyping errors associated with good quality DNA. *Mol. Ecol. Notes* 1: 1–2.
- Jordan D.S. 1905: The origin of species through isolation. *Science* 22: 545–562.
- Joseph L., Dolman G., Donnellan S., Saint K., Berg M. & Bennet A. 2008: Where and when does a ring start and end? Testing the ring-species hypothesis in a species complex of Australian parrots. *Proc. R. Soc. Lond. B* 275: 2431–2440.
- Kuchta S.R., Parks D.S., Mueller R.L. & Wake D.B. 2009: Closing the ring: historical biogeography of the salamander ring species *Ensatina eschscholtzii*. *J. Biogeogr.* 36: 982–995.
- Liebers D., Knijff P. & Helbig A. 2004: The herring gull complex is not a ring species. *Proc. R. Soc. Lond. B* 271: 893–901.
- Mayr E. 1942: Systematics and the origin of species. *Columbia University Press, New York*.
- Nosil P. 2012: Ecological speciation. *Oxford University Press, Oxford*.
- Packert M., Martens J., Eck S., Nazarenko A., Valchuk O., Petri B. & Veith M. 2005: The great tit (*Parus major*) – a misclassified ring species. *Biol. J. Linn. Soc.* 86: 153–174.
- Paetkau D., Slade R., Burden M. & Estoup A. 2004: Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Mol. Ecol.* 13: 55–65.



- Patten M.A. 2010: Evolution and historical biogeography of a song sparrow ring in western North America. In: Pontarotti P. (ed.), Evolutionary biology: concepts, molecular and morphological evolution. *Springer, Berlin*: 329–342.
- Patten M.A., Erickson R.A. & Unitt P. 2004a: Population changes and biogeographic affinities of the birds of the Salton Sink, California/Baja California. *Stud. Avian Biol.* 27: 24–32.
- Patten M.A. & Pruett C.L. 2009: The song sparrow, *Melospiza melodia*, as a ring species: patterns of geographic variation, a revision of subspecies, and implications for speciation. *Syst. Biodivers.* 7: 33–62.
- Patten M.A., Rotenberry J. & Zuk M. 2004b: Habitat selection, acoustic adaptation, and the evolution of reproductive isolation. *Evolution* 58: 2144–2255.
- Petren K. 1998: Microsatellite primers from *Geospiza fortis* and cross species amplification in Darwin's finches. *Mol. Ecol.* 7: 1782–1784.
- Piry S., Alapetite A., Cornuet J., Paetkau D., Baudouin L. & Estoup A. 2004: GeneClass 2: a software for genetic assignment and first-generation migrant detection. *J. Hered.* 95: 536–539.
- Postma E. & van Noordwijk A.J. 2005: Gene flow maintains a large genetic difference in clutch size at a small spatial scale. *Nature* 433: 65–68.
- Pruett C.L., Arcese P., Chan Y.L., Wilson A.G., Patten M.A., Keller L.F. & Winker K. 2008a: The effects of contemporary processes in maintaining the genetic structure of western song sparrows (*Melospiza melodia*). *Heredity* 101: 67–74.
- Pruett C.L., Arcese P., Chan Y.L., Wilson A.G., Patten M.A., Keller L.F. & Winker K. 2008b: Concordant and discordant signals between genetic data and described subspecies of Pacific coast song sparrows. *Condor* 110: 359–364.
- Pruett C.L., Topp C.M., Maley J.M., McCracken K.G., Rohwer S., Birks S., Sealy S.G. & Winker K. 2013: Evidence from the genetics of landbirds for a forested Pleistocene glacial refugium in the Haida Gwaii area. *Condor* 115: 725–737.
- Pruett C.L. & Winker K. 2005: Northwestern song sparrow populations show genetic effects of sequential colonization. *Mol. Ecol.* 14: 1421–1434.
- Radosavljevic A. & Anderson R.P. 2014: Making better MaxEnt models of species distributions: complexity, overfitting and evaluation. *J. Biogeogr.* 41: 629–643.
- Rosenberg N.A. 2004: Distruct: a program for the graphical display of population structure. *Mol. Ecol. Notes* 4: 137–138.
- Schluter D. 2009: Evidence for ecological speciation and its alternative. *Science* 323: 737–741.
- Swenson N.G. & Howard D.J. 2005: Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *Am. Nat.* 166: 581–591.
- Templeton A.R. 2009: Statistical hypothesis testing in intraspecific phylogeography: nested clade phylogeographical analysis vs. approximate Bayesian computation. *Mol. Ecol.* 18: 319–331.
- Topp C.M. & Winker K. 2008: Genetic patterns of differentiation among five species of landbirds on the Queen Charlotte Islands, British Columbia. *Auk* 125: 461–472.
- Van Devender T.R. & Spaulding W.G. 1979: Development of vegetation and climate in the southwestern United States. *Science* 204: 701–710.
- Van Oosterhout C., Hutchinson W.F., Wills D.P.M. & Shipley P. 2004: Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4: 535–538.
- Waltari E., Hijmans R.J., Peterson A.T., Nyári Á.S., Perkins S.L. et al. 2007: Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS ONE* 2: e563.
- Wilson A.G., Arcese P., Chan Y.L. & Patten M.A. 2011: Micro-spatial genetic structure in song sparrows (*Melospiza melodia*). *Conserv. Genet.* 12: 213–222.
- Zink R.M. & Dittmann D.L. 1993: Gene flow, refugia, and evolution of geographic variation in the song sparrow (*Melospiza melodia*). *Evolution* 47: 717–729.

**Appendix I.** Collection locations, subspecies, and accession numbers of song sparrows. All catalogue numbers are Florida Institute of Technology Bird Collection (FIT).

Locality	Subspecies	Catalogue numbers (FIT)
Nevada	<i>M. m. montana</i>	197, 204, 214, 217, 225, 226, 255, 257, 274, 275, 277, 288, 316-322, 325
Oregon: Malheur	<i>M. m. montana</i>	158, 169, 212, 213, 234, 235, 239, 240, 245, 248, 250, 259, 260, 265, 266, 276, 287, 331-333
California: Modoc	<i>M. m. montana</i>	124, 127, 135-137, 142, 148, 151, 153, 154, 179, 188-190, 192, 193, 195, 267, 312, 313
California: Los Banos	<i>M. m. heermanni</i>	269-271, 273, 278, 279, 283, 286, 296, 302, 304-311, 314, 329
Utah: Dixie	<i>M. m. fallax</i>	280-282, 284, 285, 297, 303, 326, 327, 330
Arizona	<i>M. m. fallax</i>	194, 196, 198, 199, 201, 202, 203, 232, 244, 246, 256, 258, 323, 324, 328

**Supplementary online materials**

**Appendix II.** Locations of birds used in ecological niche model. Blue points represent the subspecies *Melospiza melodia heermanni*, purple points are *M. m. montana*, and red points are *M. m. fallax* (URL: [http://www.ivb.cz/fofia/download/smyth\\_et\\_al.\\_appendix\\_II.jpg](http://www.ivb.cz/fofia/download/smyth_et_al._appendix_II.jpg)).