

Behavioural response to song and genetic divergence in two subspecies of white-crowned sparrows (*Zonotrichia leucophrys*)

SARA E. LIPSHUTZ,*  ISAAC A. OVERCAST,† MICHAEL J. HICKERSON,† ‡ §
ROBB T. BRUMFIELD¶ and ELIZABETH P. DERRYBERRY*¶

*Department of Ecology and Evolutionary Biology, Tulane University, 400 Lindy Boggs, New Orleans, LA 70118, USA, †Subprogram in Ecology, Evolution and Behavior, The Graduate Center, City University of New York, New York, NY 10016, USA, ‡Department of Biology, Marshak Science Building, City College of New York, Room 526, 160 Convent Avenue, New York, NY 10031, USA, §Division of Invertebrate Zoology, The American Museum of Natural History, Central Park West and 79th Street, New York, NY 10024, USA, ¶Museum of Natural Science and Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803, USA

Abstract

Divergence in sexual signals may drive reproductive isolation between lineages, but behavioural barriers can weaken in contact zones. Here, we investigate the role of song as a behavioural and genetic barrier in a contact zone between two subspecies of white-crowned sparrows (*Zonotrichia leucophrys*). We employed a reduced genomic data set to assess population structure and infer the history underlying divergence, gene flow and hybridization. We also measured divergence in song and tested behavioural responses to song using playback experiments within and outside the contact zone. We found that the subspecies form distinct genetic clusters, and demographic inference supported a model of secondary contact. Song phenotype, particularly length of the first note (a whistle), was a significant predictor of genetic subspecies identity and genetic distance along the hybrid zone, suggesting a close link between song and genetic divergence in this system. Individuals from both parental and admixed localities responded significantly more strongly to their own song than to the other subspecies song, supporting song as a behavioural barrier. Putative parental and admixed individuals were not significantly different in their strength of discrimination between own and other songs; however, individuals from admixed localities tended to discriminate less strongly, and this difference in discrimination strength was explained by song dissimilarity as well as genetic distance. Therefore, we find that song acts as a reproductive isolating mechanism that is potentially weakening in a contact zone between the subspecies. Our findings also support the hypothesis that intraspecific song variation can reduce gene flow between populations.

Keywords: birdsong, genotyping-by-sequencing, hybrid zone, reproductive isolation, subspecies, *Zonotrichia leucophrys*

Received 21 June 2016; revision received 6 December 2016; accepted 8 December 2016

Introduction

Discrimination between diverged sexual signals can contribute to reproductive isolation by reducing gene

flow between populations (Dobzhansky 1940; Mayr 1963; West-Eberhard 1983; Coyne & Orr 2004). Behavioural responses to a sexual signal provide a measure of the salience of that signal in mate choice (Searcy 1992) and in territorial interactions (Kroodsma 1986; Nowicki *et al.* 1998). Numerous empirical studies suggest that mating signals used to attract mates and repel

Correspondence: Elizabeth P. Derryberry, Fax: (504) 862 8706; E-mail: ederrybe@tulane.edu

competitors can also promote behavioural isolation among closely related populations (reviewed in West-Eberhard 1983; Andersson 1994; Price 1998; Panhuis *et al.* 2001). Concordance in geographic patterns of sexual signal and genetic variation occurs in a number of taxa (e.g. crickets, Shaw *et al.* 2007; mice, Campbell *et al.* 2010; gibbons, Thinh *et al.* 2011; and frogs, Warwick *et al.* 2015), supporting the hypothesis that divergence in mating signals between populations can act as a behavioural reproductive isolating mechanism.

Much emphasis has been placed on the importance of song, and in particular learned song, in facilitating speciation in birds (Marler & Tamura 1964; Nottebohm 1969; Baker & Cunningham 1985; Grant & Grant 1996; Martens 1996; Price 1998, 2008; Slabbekoorn & Smith 2002; Podos & Warren 2007). As a long-distance signal, song is often the first aspect of the phenotype that can be assessed by would-be mates or competitors (Catchpole & Slater 2008), and so has potential as a behavioural barrier to gene flow (Coyne & Orr 2004). Many birds produce distinct songs, and birds typically respond strongest to the song of their own species (reviewed in Andersson 1994; Martens 1996), supporting a role for song as an isolating mechanism. Within a species, there is also substantial geographic variation in song (reviewed in Podos & Warren 2007). Both male and female receivers often discriminate between songs of different cultural populations (Searcy *et al.* 1997; Derryberry 2007; Seddon & Tobias 2007), which should reduce gene flow between populations with diverged signals. However, there is little evidence for song acting as an intraspecific barrier to gene flow between cultural populations (reviewed in Slabbekoorn & Smith 2002). The majority of empirical studies that focus within a species do not find genetic substructuring based on song divergence for songbirds (oscines) (e.g. Payne & Westneat 1988; Loughheed & Handford 1992; Soha *et al.* 2004; Ruegg *et al.* 2006; Leader *et al.* 2008; Ortiz-Ramírez *et al.* 2016), nor for nonoscines in which vocal learning evolved independently (e.g. Wright & Wilkinson 2001; Saranathan *et al.* 2007; Gonzalez & Ornelas 2014). Thus, although there is abundant evidence that song acts as an isolating mechanism between bird species, when and how intraspecific song variation facilitates reproductive divergence is less clear (Slabbekoorn & Smith 2002; Lachlan & Servedio 2004).

Interactions between hybridizing lineages present an opportunity to investigate how intraspecific song divergence affects the process of mate selection and resource acquisition. Hybrid zones are natural laboratories for studying the speciation process because they facilitate the testing of behavioural and genetic barriers between differentiated lineages (Endler 1977; Barton & Hewitt 1985; Hewitt 1988; Harrison 1993). Although studies of

hybridizing lineages typically find song divergence in allopatry (e.g. Halfwerk *et al.* 2016), songs are often more convergent in sympatry, due to local acoustic adaptation or interspecific learning (Secondi *et al.* 2003; Haavie *et al.* 2004; Qvarnström *et al.* 2006; Kenyon *et al.* 2011), or hybridization itself (de Kort *et al.* 2002). Likewise, in many cases, individuals in hybrid zones discriminate between diverged songs (Patten *et al.* 2004; Turčoková *et al.* 2011; Greig & Webster 2013), while in other cases, individuals do not discriminate between nonlocal and local songs (Matessi *et al.* 2000; Gee 2005; den Hartog *et al.* 2008), or there is asymmetric song recognition (Kershner & Bollinger 1999; Dingle *et al.* 2010; Ruegg *et al.* 2012; McEntee 2014; Pegan *et al.* 2015). Concordance between (i) genetic and (ii) song divergence, along with (iii) behavioural discrimination between lineage-specific songs, suggests that songs have the potential to maintain, if not drive reproductive isolation. However, few studies test for an association among all three components, and fewer studies place these patterns in the context of evolutionary history. This context can help us understand how behavioural isolating barriers function in the transition from populations to species (Coyne & Orr 2004; Edwards *et al.* 2005).

Although recently diverged lineages offer a window into the speciation process (Hewitt 1988), their evolutionary history can be problematic to determine (Durrett *et al.* 2000; Pettengill & Moeller 2012). Historical demographic inference based on the coalescent can be used to estimate the relative time since divergence between lineages, which may indicate the role of historical, ecological or biogeographic processes (e.g. Hickerson *et al.* 2006), as well as to calculate the degree of historical gene flow between current lineages (e.g. Carling *et al.* 2010; Field *et al.* 2011). Model-based approaches employing coalescent-based analyses of multilocus sequence data can also test alternative hypotheses of evolutionary histories (Rosenberg & Nordborg 2002; Gutenkunst *et al.* 2009; Excoffier *et al.* 2013), which can provide a framework for interpreting reproductive barriers. Examples of evolutionary histories for recently diverged lineages include primary divergence with either strict isolation or ongoing symmetrical or asymmetrical migration, migration after a period of allopatric divergence (e.g. secondary contact) and panmixia. Differentiating between primary divergence and secondary contact can provide information on whether behavioural divergence could have occurred in allopatry. If evidence of asymmetric gene flow between taxa coincides with a pattern of asymmetric recognition of songs, then this pattern would provide support for song as an incomplete behavioural barrier (e.g. Halfwerk *et al.* 2016). In comparison,

concordance between symmetrical song discrimination and secondary contact with reduced gene flow would suggest song acting as a behavioural barrier. A supported model of panmixia, on the other hand, would indicate a limited role for reproductive isolation between lineages (e.g. Oomen *et al.* 2011).

Here, we investigate whether song is a reproductive isolating mechanism in an oscine species widely studied for song evolution: the white-crowned sparrow (*Zonotrichia leucophrys*). The question of whether song is a behavioural barrier to gene flow has been asked in the white-crowned sparrow for decades (Baker 1975; Baker *et al.* 1984; Baker & Cunningham 1985; MacDougall-Shackleton & MacDougall-Shackleton 2001; Soha *et al.* 2004) because of a strong pattern of discrete song types known as dialects across small geographic scales (Marler & Tamura 1964), with male (Nelson & Soha 2004) and female (Petrinovich & Patterson 1981) discrimination between dialects. Empirical data for genetic differentiation between cultural dialect populations within subspecies of white-crowned sparrows is mixed (Baker *et al.* 1982 as revisited by Soha *et al.* 2004; MacDougall-Shackleton & MacDougall-Shackleton 2001). However, behavioural studies suggest that examining this question in the context of hybridization between subspecies may lend insight into this question. Experiments with white-crowned sparrows demonstrate that males have a genetic predisposition to learn the song of their own subspecies, and females to pay closer attention to the song of their own subspecies (Whaling *et al.* 1997; Nelson 2000), suggesting an innate behavioural barrier to gene flow between subspecies despite learned song.

We focus on two subspecies distributed along the western coast of North America: the Puget Sound subspecies (hereafter *Z. l. pugetensis*), and the Nuttall's subspecies (hereafter *Z. l. nuttalli*). *Z. l. pugetensis* is migratory and breeds from northern California to southern British Columbia, whereas *Z. l. nuttalli* is a year-round resident that breeds in coastal central and northern California (Grinnell 1928; Blanchard 1941; Banks 1964). The subspecies are hypothesized to have diverged in glacial refugia during the Pleistocene (Banks 1964; Baker *et al.* 1984), and a putative contact zone exists in northern California (Banks 1964; Mewaldt *et al.* 1968; Corbin & Wilkie 1988). Previous studies on *Z. l. nuttalli* and *Z. l. pugetensis* documented cultural differences (Baker 1987) as well as behavioural discrimination between subspecific songs (Lampe & Baker 1994), but found limited genetic divergence based on allozymes (Corbin 1981; Corbin & Wilkie 1988) and mitochondrial haplotypes (Weckstein & Zink 2001). If the two subspecies are distinct and hybridizing, we expect to find (i) two genetic clusters with admixture

between them, as well as (ii) support for a historical demographic model of secondary contact. If song functions as an isolating mechanism between the subspecies, we expect to find (iii) song divergence between the subspecies; (iv) differential male response to subspecific songs; and (v) an association between song divergence, genetic divergence and the strength of discrimination between songs.

Materials and methods

Genetic sampling and sequencing

We sampled 190 individuals from 17 localities along a coastal transect spanning the ranges of *Z. l. nuttalli* and *Z. l. pugetensis* (Fig. 1, Table 1). We collected blood samples (20 µL) by brachial venipuncture from 132 mist-netted males in 2004 and 2005 and released birds after metal banding. We transferred blood to EDTA-saturated filter paper and stored in airtight containers on DrieRite at room temperature. Our sample also included tissues from 51 vouchered specimens collected in 2010. Voucher specimens are deposited in the Museum of Natural Science at Louisiana State University. We also collected vocal data for these 183 males (see Song recording and analysis). Seven additional samples, including four females, were provided as tissue loans from the Museum of Vertebrate Zoology and the Burke Museum. We extracted total genomic DNA using a DNeasy blood and tissue extraction kit following the manufacturer's recommended instructions (Qiagen, Valencia, CA, USA).

We sent DNA extracts to the Institute of Genomic Diversity at Cornell University in Ithaca, NY, USA, for genotyping-by-sequencing (GBS). GBS reduced-representation libraries were prepared and analysed according to the methods of Elshire *et al.* (2011) using the restriction enzyme PstI (CTGCAG) for digestion and creating a library with 95 unique barcodes, one for each individual for each plate. Samples were sequenced on two lanes of the Illumina HiSeq platform, generating 551 083 045 reads.

Single-nucleotide polymorphisms (SNPs) were called by processing the raw 100-bp single-end sequence reads using the UNEAK pipeline (Lu *et al.* 2013), an extension of the Java program of TASSEL 4.0 (Bradbury *et al.* 2007). Reverse complement tag-pairs were collapsed, and loci with greater than 20% missing data, as well as minor allele frequency less than 1%, were excluded. Samples were defined as failed if the number of sequences produced for that sample was less than 10% of the mean number of sequences for all samples sequenced in that flow cell lane—four failed samples were excluded from further analysis.

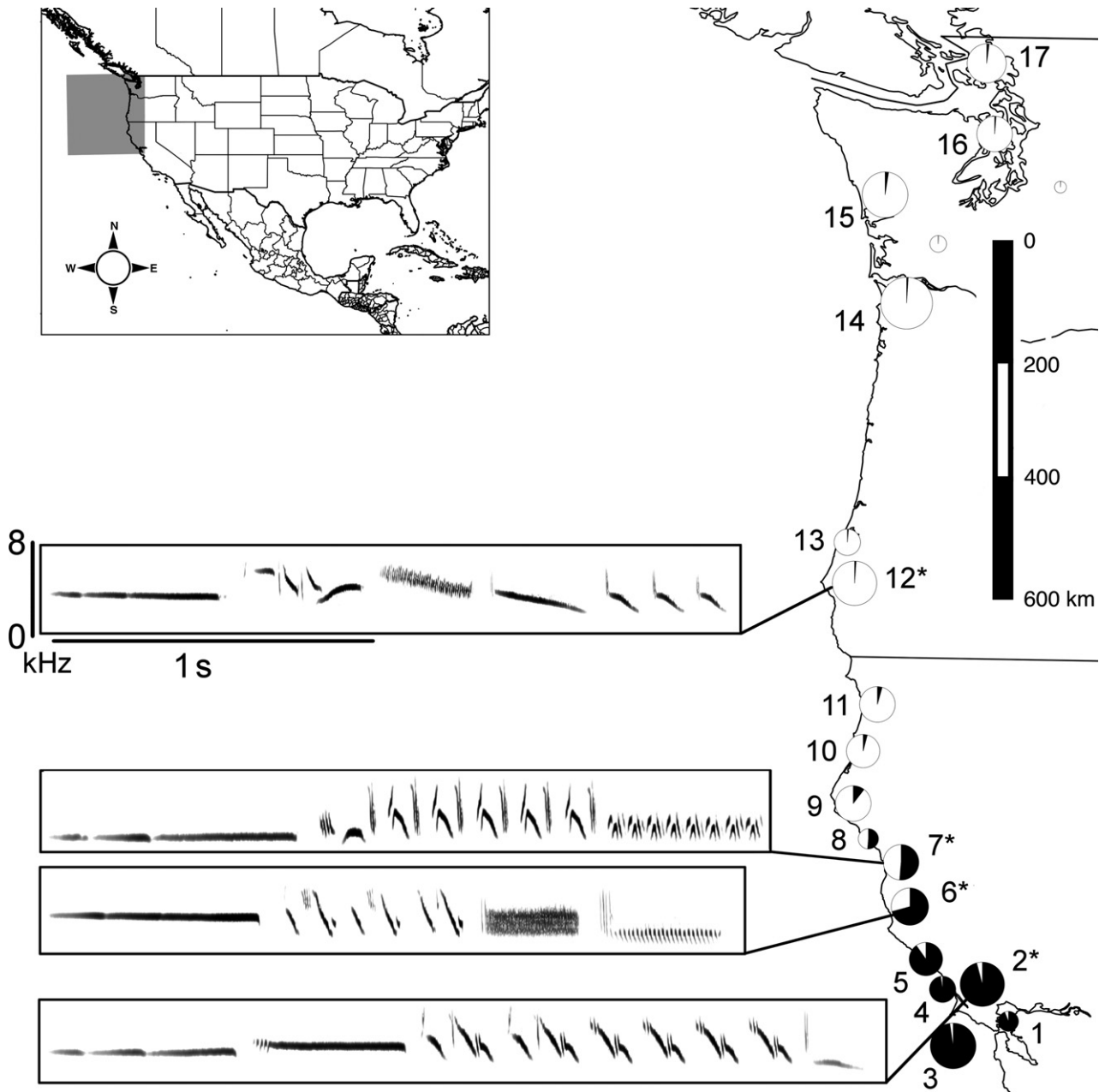


Fig. 1 Sampling map of numbered localities along a transect from the southern subspecies *Z. l. nuttalli* to the northern subspecies *Z. l. pugetensis*. Song spectrograms correspond to the sites of behavioural playback experiments. Pie charts represent admixture proportions from STRUCTURE for each locality, and circle size corresponds to the number of individuals selected for sequencing at each site. Grey pie charts indicate two samples collected outside of main localities. Asterisks indicate localities for territorial playback experiments.

The UNEAK pipeline identified a total of 79 130 biallelic SNPs. After filtering in the pipeline, the final data matrix for all 186 individuals contained 1583 SNPs. This reduction in loci was due in part to a handful of individuals with very few reads overall. To address this, a restricted data set was created by removing individual samples that produced fewer than 500 000 reads. The restricted data set resulted in 6419 SNPs for 169 individuals, with total missing data (number of missing genotypes

per locus per sample) equal to 16.9%. All subsequent analyses were conducted with this restricted data set.

Population structure analysis

To characterize patterns in genetic structure and assign individuals to populations, we used the program STRUCTURE v. 2.3.4 (Pritchard *et al.* 2000). Structure analyses were performed using a burn-in length of 200 000,

Table 1 Sampling information for each locality. *Q* values from STRUCTURE refer to mean admixture proportions for individuals in each locality

Site	Locality	County	State	Latitude	Longitude	Year	<i>Q</i> value	Song <i>N</i>	Genetic <i>N</i>	Subspecies
1	San Francisco	San Francisco	CA	37.803	-122.478	2005	0.061	12	3	nuttalli
2	Bolinas	Marin	CA	37.908	-122.722	2005	0.024	34	15	nuttalli
3	Schooner Bay	Marin	CA	38.083	-122.914	2004	0.041	9	14	nuttalli
4	Abbotts Lagoon	Marin	CA	38.122	-122.953	2010	0.023	9	5	nuttalli
5	Sonoma	Sonoma	CA	38.419	-123.105	2010	0.098	11	8	nuttalli
6	Manchester	Mendocino	CA	38.981	-123.702	2010	0.292	5	10	admixed
7	MacKerricher	Mendocino	CA	39.489	-123.8	2005	0.487	15	9	admixed
8	Sinkyone	Mendocino	CA	39.832	-123.85	2010	0.487	0	3	admixed
9	Ferndale	Humboldt	CA	40.544	-124.358	2010	0.896	10	9	pugetensis
10	Eureka	Humboldt	CA	40.742	-124.239	2005	0.96	12	8	pugetensis
11	Trinidad	Humboldt	CA	41.258	-124.099	2010	0.957	10	9	pugetensis
12	Bandon	Coos	OR	43.071	-124.435	2005	0.989	12	14	pugetensis
13	Bullards Beach	Coos	OR	43.127	-124.416	2005	0.986	6	5	pugetensis
14	Nehalem	Tillamook	OR	45.684	-123.938	2005	0.99	16	19	pugetensis
	Frances	Pacific	WA	46.557	-123.4		0.995	0	2	pugetensis
15	Ocean Shores	Grays Harbor	WA	46.928	-124.17	2005	0.975	7	15	pugetensis
	Enumclaw	King	WA	47.248	-122.013		0.996	0	1	pugetensis
16	Dosewallips	Jefferson	WA	47.692	-122.895	2005	0.987	6	9	pugetensis
17	San Juan Island	San Juan	WA	48.461	-123.014	2004	0.98	34	11	pugetensis

and 500 000 MCMC repetitions to test clusters ranging from $K = 1$ to $K = 20$ to allow for substructuring within sampling sites, with 10 independent replicates, after which all parameters converged. Parameter settings also included an admixture model of ancestry and correlated allele frequencies. An individual was classified as parental *Z. l. nuttalli* if $Q \leq 0.1$, parental *Z. l. pugetensis* if $Q \geq 0.9$ and admixed if $0.1 < Q < 0.9$. The optimal number of clusters (K value) was calculated using ΔK likelihood evaluations (Evanno *et al.* 2005) in Structure Harvester (Earl & vonHoldt 2012). We used CLUMPP (Jakobsson & Rosenberg 2007) to identify potential multimodality and account for label switching among replicates and DISTRICT (Rosenberg 2004) to visual admixture proportions. We recognize that STRUCTURE is merely an exploratory starting point for downstream population-based analyses, and it can be hazardous to read its results as actual inference (Falush *et al.* 2016). Therefore, we also assessed population structure with a model-free method based on multidimensional statistics: principal component analysis (PCA) implemented in the R (R Core Team 2015) package ADEGENET v.2.0 (Jombart & Ahmed 2011). We used the function SCALEGEN to scale allele frequencies and replace missing genotype data with the mean allele frequencies. We performed the PCA with the function dudi.pca. For visualization, each individual was labelled according to sampling location, allowing us to examine relationships among individuals without a priori assumption about subspecies assignment. Alternative methods of analysis (e.g. FASTSTRUCTURE (Raj *et al.* 2014) and DAPC (Jombart *et al.* 2010))

showed qualitatively similar findings and are not presented here.

Inference of historical demographic parameters

To distinguish the demographic history of these populations and to estimate key parameters of interest, such as divergence time and effective population size, we used the composite-likelihood simulation-based approach of fastsimcoal2 (Excoffier *et al.* 2013). We generated the observed joint, folded site frequency spectrum (SFS) using custom python scripts and *daði* (Gutenkunst *et al.* 2009). *daði* provides a facility for projecting an SFS from a larger sample size to a smaller sample size. For RAD-like data sets, which often contain significant missing data, projecting down can increase the number of usable SNPs by averaging over resamplings of the larger data set. We explored multiple possible values for projection and selected two values for full downstream analysis: one larger (75×75) and one smaller (20×20), as measured in number of individuals per population. The larger projection was selected to maximize the number of segregating sites. The smaller projection was selected to minimize total size of the SFS without sacrificing demographically important signal in the data, in an effort to explore the trade-off between computation time and accuracy of inference. We tested five different historical demographic models for goodness of fit to the observed data: (i) primary divergence with strict isolation, (ii) divergence with ongoing symmetrical and (iii) asymmetrical migration, (iv) migration

after a period of allopatric divergence (e.g. secondary contact) and (v) panmixia. For all models, we explored two partitions of the data: (i) admixed individuals assigned to subspecies based on admixture proportions from the aggregated *STRUCTURE* runs for $K = 2$ and (ii) only parental individuals by excluding those with admixture proportions between 0.1 and 0.9. We selected wide, uninformative, uniform search ranges for all estimated parameters. We performed 50 independent runs per model, per data partition to obtain likelihood values for the observed data under each given model. For each run, we performed 100 000 simulations and 40 expectation–conditional maximization (ECM) cycles. We evaluated model fit with both information theoretic (AIC) and likelihood (LRT)-based methods. We generated 95% confidence intervals for demographic parameters of interest using 100 parametric bootstrap replicates. For each bootstrap replicate, we simulated a new SFS with the same number of SNPs as our observed data using the maximum-likelihood parameters and the best fitting model from the model selection step. We re-estimated parameters across replicates for the simulated SFS and aggregated maximum-likelihood parameters. We generated bootstrap confidence intervals using the python package *SCIKITS-BOOTSTRAP* (<https://github.com/geevans/scikits-bootstrap>).

Song recording and analysis

Within these two subspecies, males produce one stereotyped song type (e.g. dialect). Most males in each location produce the same dialect, and males in different locations produce different dialects. We recorded 208 males defending territories during the breeding season in 16 different localities (we did not have songs for site 8), each with their own unique song dialect, in 2004, 2005 and 2010, with an average of 12 individuals (range = 5–34) per site (Table 1). Song dialects in these localities have been stable over 30 years (Derryberry 2009; Luther & Derryberry 2012), so have not likely changed across this six-year spread in sampling. Recordings were made using a Sony TCM-5000EV cassette recorder, a PRO-302 Unidirectional Dynamic microphone, and a Sony PBR330 parabolic reflector. All songs were digitized with 16-bit precision at a 25 kHz sampling rate using *SYRINX 2.2b* (Burt 2001) and an Echo Digital Audio sound card. All songs were high pass filtered to eliminate noise below 1500 Hz. We measured eight acoustic parameters shared by all song types: song maximum and minimum frequencies (Hz); the dominant (peak) frequency of the whistle; the duration of the whistle, the introduction and the average syllable duration; and the rate of trill note delivery and frequency bandwidth of the trill (Table S1, Supporting

information). All measurements were taken using *SIGNAL* version 3.1 or 5 (Beeman 1999). We took minimum and maximum frequency measurements at -36 dB relative to the peak amplitude frequency in the song from digital spectrograms (256 pt transform, frequency resolution = 97.7 Hz). We calculated frequency bandwidth as the difference between the maximum and minimum frequencies. We measured dominant frequency as the frequency at which the most sound energy was transmitted during production of the relevant song section from a smoothed power spectrum (smoothing resolution 100 points). Temporal variables were measured from oscillograms (time waveforms). Trill rate was calculated as the number of notes produced per second. Following Podos (2001), we calculated a ninth acoustic parameter, vocal performance, as the orthogonal distance between each song and an upper-bound regression for the plot of trill frequency bandwidth as a function of trill rate for 375 white-crowned sparrow songs from 15 different dialects that has been shown to be robust to different methods for estimating the performance trade-off between bandwidth and trill rate (Derryberry 2009; Wilson *et al.* 2014). Songs closer to the limit are higher performance. All raw song data were transformed to a scale with a mean of 0 and a SD of 1 (a z-score) to allow for scale free comparisons.

We assessed if song acoustic parameters could distinguish the two subspecies using two approaches. First, we ran a discriminant function analysis (DFA) of the individual song variables using *JMP v.12* (Sall 2015). Songs from genetically admixed individuals were classified to subspecies based on their admixture proportions from *STRUCTURE*. Next, we summarized seven of the acoustic parameters using a PCA in *JMP*. We excluded trill rate and frequency bandwidth from the PCA, as these were used to calculate vocal performance. This yielded four independent factors with an eigenvalue greater than 1, explaining a total of 75% of the original song variation (Table 2). To assess whether songs have diverged between subspecies, we used a linear mixed model approach with locality ($n = 16$) as a random effect and subspecies as the predictive factor.

Territorial playback experiment

To measure subspecies-specific discrimination among songs, we tested free-living, territorial adult males using territorial playbacks, a standard experimental design that quantifies male response to simulated intrusion on their territories (McGregor *et al.* 1992). We conducted playback experiments in two parental localities (sites 2 and 12) and two admixed localities (sites 6 and 7), hereafter referred to as playback localities. Playbacks were conducted between May and June in 2010 (sites 2, 7, 12)

Table 2 Factor loadings for the first five principal components (eigenvalue >1) derived from song variables

Song parameter	PC1	PC2	PC3	PC4
Eigenvalue	1.27	1.21	1.07	1.03
Per cent variation	23	21	16	15
Whistle length (ms)	0.59			
Avg. note length (ms)	-0.59	0.24	0.15	
Song maximum frequency (Hz)	0.18	0.66	-0.17	-0.16
Song minimum frequency (Hz)	0.2	-0.45	-0.52	-0.39
Whistle dominant frequency (Hz)	-0.41		-0.21	-0.69
Introduction length (ms)	-0.26	-0.28	-0.46	0.56
Vocal performance		-0.46	0.65	-0.17

Bold values indicate significant loadings.

and in 2013 (site 6), when males were actively defending breeding territories. Focal males were not genotyped.

Stimuli were presented in a paired, balanced design. In each parental locality, we assessed male response to songs from their own location ('own') and to songs from a parental locality of the other subspecies ('other'), and in each admixed location, we assessed male response to 'own' local song vs. 'other' for each of the two subspecies. We used 10 exemplars for each song category and tested 10 males for each comparison. Each male heard a different pair of exemplars to avoid pseudoreplication (Kroodsma 1990; Kroodsma *et al.* 2001). Order of presentation and selection of exemplars were randomized across males. Stimuli from site 2 were selected from recordings made in 2004, sites 7 and 12 from 2005, and site 6 from 2010. Adult males in the wild live an average of 16 months (Cortopassi & Mewaldt 1965), so it is unlikely that males tested were familiar with the males that produced the stimulus songs.

We separated treatments by 48 hours to minimize habituation and did not test neighbours on the same day. Songs were amplitude normalized and broadcast at a constant level (80–82 dB SPL 1 m) and at a natural rate of six songs per minute from a speaker (Altec Lansing IMT320 inMotion) near the centre of the focal male's territory as determined by behavioural observations. We measured four response variables: mean distance from the speaker during the (i) three-min playback period and the (ii) three-min postplayback period, (iii) number of flights over the speaker during the playback period and (iv) song rate (songs/min) during the playback period. To facilitate accurate measures of distance, we placed markers at 4 and 8 m on either

side of the speaker before the trial began. Distances and observed behaviours were recorded at 10-second intervals. Males were considered to have a stronger response to the stimulus when they approached the speaker more closely (Searcy *et al.* 2006), flew over the speaker more often, and produced songs at a higher rate. One observer narrated observations while another recorded the observations onto datasheets. Experiments were not blind to stimulus type.

For each experiment, we reduced the four behavioural response measures using PCA and used the PC scores in statistical testing in R (R Core Team 2015). The original behavioural variables were not statistically independent (Rice 1989), and separate tests would not reflect the multivariate nature of the males' responses (McGregor *et al.* 1992). Playback and postplayback distance were highly correlated, so we calculated an average distance to include in PCAs. We retained PCs with an eigenvalue greater than 1 for analyses, or the minimum number of PCs required to explain 50% or more of the variance in the original variables (Table 3). To test whether males discriminated between 'own' and 'other' within each locality, we analysed paired comparisons separately using Wilcoxon signed-rank tests on the respective PCs. Repeated-measures MANOVAS were then used to test for differences in male response due to (i) locality and (ii) genetic status (putative parental vs. putative admixed). Joint analyses standardized the postplayback data to the first 3 min. For combined analyses using MANOVAS, male response was normally distributed (all Shapiro–Wilk $P > 0.25$) and equal in variance (all Leven's $P > 0.18$). Effect size (Cohen's d) and power to reject a false null hypothesis were also determined (Cohen 1988). We calculated the power for our given N , effect sizes (estimated from means and SDs) and alpha level of 0.05 using the G^* POWER 3.1 (Faul *et al.* 2009) for Wilcoxon signed-rank tests (matched pairs).

Song dissimilarity, behavioural response and F_{ST}

We asked whether variation in male response to 'own' vs. 'other' song was explained by acoustic dissimilarity, as well as pairwise F_{ST} between the parental (2 and 12) and admixed (6 and 7) playback locations. We calculated a dissimilarity score between all songs used for playback analyses using the dynamic time warping function in LUSCINIA v.2.02.10.15 (Lachlan 2007). This function searches for the optimal alignment of two signals and then calculates a dissimilarity score based on temporal and spectral characters. For each playback experiment, we averaged song dissimilarity scores and the response difference to 'own' vs. 'other'. We calculated genetic distance among the playback localities

Table 3 Factor loadings for the principal components (eigenvalue >1) derived from behavioural responses to individual playback experiments and all experiments standardized and combined

Experiments (own vs. other)	Response variables	PC1	PC2
Bolinas (2) vs. Bandon (12)	Average distance	-0.76	
	Song rate	0.87	
	Fly overs	0.65	
	Eigenvalues	1.75	
	Cum. per cent variance	58.5	
Manchester (6) vs. Bolinas (12)	Average distance	0.02	0.99
	Song rate	-0.76	0.05
	Fly overs	0.76	0.01
	Eigenvalues	1.17	1
	Cum. per cent variance	39	72.4
Manchester (6) vs. Bandon (12)	Average distance	-0.78	0.47
	Song rate	0.89	0.1
	Fly overs	0.3	0.92
	Eigenvalues	1.5	1.08
	Cum. per cent variance	50	86
MacKerricher (7) vs. Bolinas (2)	Average distance	-0.83	
	Song rate	0.63	
	Fly overs	0.73	
	Eigenvalues	1.62	
	Cum. per cent variance	54	
MacKerricher (7) vs. Bandon (12)	Average distance	-0.83	
	Song rate	0.8	
	Fly overs	0.71	
	Eigenvalues	1.8	
	Cum. per cent variance	61	
Combined own vs. other	Average distance	-0.8	
	Song rate	0.79	
	Fly overs	0.63	
	Eigenvalues	1.7	
	Cum. per cent variance	55	

using a pairwise F_{ST} matrix in ARLEQUIN v3.5 (Excoffier & Lischer 2010) based on the 6419 SNP data set (Table S2, Supporting information). We then used linear regression to compare male response to song dissimilarity and male response to genetic distance in R (R Core Team 2015).

Multiple matrix regression with randomization

To quantify the independent contributions of song and geographic distance on genetic differentiation, we used a multiple matrix regression with randomization (MMRR) (Wang 2013). Unlike a partial Mantel, MMRR

uses a randomized permutation procedure to account for nonindependence between variables (Wang 2013). We quantified isolation by song (IBS) and isolation by distance (IBD) with the 'MMRR' function in R for 10 000 permutations, using pairwise genetic distance as the response variable and geographic and song distances as the explanatory variables. We calculated genetic distance among locations using a pairwise F_{ST} matrix in ARLEQUIN v3.5, based on the 6419 SNP data set (Excoffier & Lischer 2010). We calculated the geographic distance matrix from GPS coordinates using the 'earthdist' function in the R package fossil (Vavrek 2011). We calculated the song distance matrices for each of the nine song parameters as well as song PC1 using the 'write.matrix' function in the R package MASS (Venables & Ripley 2002). We excluded Sinkyone (site 8; Table 1) from analyses because we did not have song data.

Results

Population structure

The optimal number of populations in the Bayesian assignment probability analysis using STRUCTURE for all replicates was $K = 2$, based on the Evanno method, with $\Delta K = 922.63$ (Fig. 2). Increasing K did not provide a better explanation of the data as measured by log likelihood. Based on admixture proportions (Q values), individuals from localities 1–5 were confidently assigned to the southern subspecies (*Z. l. nuttalli*), localities 6–8 were admixed, and localities 9–17 were assigned to the northern subspecies (*Z. l. pugetensis*). For the PCA, we found two separate clusters corresponding to the two subspecies (Fig. S1, Supporting information). The admixed MacKerricher (site 7) and Sinkyone (site 8) samples were distributed between the two clusters, whereas Manchester (site 6) samples did not cluster with the rest of the *Z. l. nuttalli* samples.

Historical demographic inference of secondary contact

Including admixed individuals in the data set consistently increased the fit of all models; therefore, we report only the results including admixed individuals in the data matrix. Likewise, the reduced projection data set provided a poor fit for all the models, so we report results only of the larger projection. All model comparison criteria decisively supported the secondary contact model over models of continuous symmetrical or asymmetric migration, indicating that some period of isolation was important in establishing the divergence between these subspecies (Fig. S2, see Table S3, Supporting information, for a comparison of all models). However, we caution here that we could not possibly

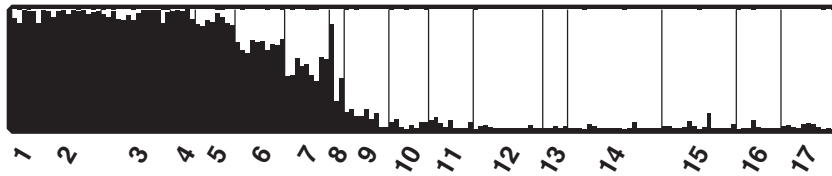


Fig. 2 Probability of assignment to *Z. l. nuttalli* (black) and *Z. l. pugetensis* (white) as determined from a STRUCTURE analysis using 6419 SNPs for $K = 2$ across 17 localities.

evaluate all historical scenarios and therefore cannot fully reject a model of primary differentiation. Our parameter estimates indicate incomplete isolation beginning during the last glacial cycle (~45 kya), followed by a short period of complete isolation after the Last Glacial Maximum (LGM) (~9 kya), and secondary contact only relatively recently (~2 kya). Recent migration rates (~10 individuals per generation) were on the order of 5× higher than the migration rate between the time of initial divergence and isolation (~2 individuals per generation), indicating a period of drastically reduced connectivity (Table 4).

Song variation between subspecies

Songs have diverged between subspecies, although not in all parameters [see measurements of the nine acoustic parameters in Table S1 (Supporting information) and factor loadings for the first four PCs derived from song parameters in Table 2]. In general, *Z. l. nuttalli* produce songs with longer whistles and shorter other notes. Using all acoustic parameters, all but two songs (1%) were correctly classified to subspecies by a DFA. A forward, stepwise DFA revealed that the best variable to distinguish between the subspecies was whistle length. Whistle length had a jackknifed classification accuracy of 80%; adding song minimum frequency and note length increased accuracy to 90%. Using DFA, songs from admixed individuals in Manchester (site 6) and MacKerricher (site 7) were classified as *Z. l. nuttalli*. A linear mixed model revealed that genetic subspecies identity was a significant predictor of song structure for

PC1 ($F = 22$, d.f. = 1, $P < 2.29 \times 10^{-6}$) but not for PC2-4 (PC2: $F = 0.1$, d.f. = 1, $P = 0.8$; PC3: $F = 0.6$, d.f. = 1, $P = 0.44$; PC4: $F = 0.4$, d.f. = 1, $P = 0.5$). Plotting song PC1 against geographic distance illustrates that the subspecies have diverged in song, and the songs of admixed individuals group with *Z. l. nuttalli* (Fig. 3).

Behavioural response to playbacks

Coding of playback populations as parental in Bolinas (site 2) and Bandon (site 12), and as admixed in Manchester (site 6) and MacKerricher (site 7), was corroborated by genetic clustering analyses (see Results Population structure, Table 1). Note that admixture proportions in Manchester (site 6) are 71% *Z. l. nuttalli* and in MacKerricher (site 7) are 51%. Within the two parental playback localities (sites 2 and 12), males responded more strongly to their local song than to the song of the other subspecies (site 2: PC1: $S = -25.5$, $P < 0.0059$, effect size Cohen's $d = 1.5$; site 12: PC1: $S = -23.5$, $P < 0.0137$, $d = 0.92$; Tables 5 and S4, Supporting information). In the two admixed playback localities (sites 6 and 7), males gave equal responses to local and nonlocal songs of their more genetically similar subspecies, *Z. l. nuttalli* (site 6: PC1: $S = 3.5$, $P < 0.78$, $d = 0.27$, PC2: $S = 1.5$, $P < 0.92$, $d = 0.22$; site 7: PC1: $S = -12.5$, $P < 0.23$, $d = 0.22$; Tables 5 and S4, Supporting information), but responded more strongly to local song than to *Z. l. pugetensis* song (site 6: PC1: $S = -21.5$, $P < 0.0273$, $d = 0.86$, PC2: $S = 1.5$, $P < 0.92$, $d = 0.09$; site 7: PC1: $S = -21.5$, $P < 0.0273$, $d = 0.99$; Tables 5 and S4, Supporting information).

Table 4 Maximum-likelihood parameter point estimates and 95% bootstrap confidence intervals for the secondary contact model

Parameter	Max likelihood	Bootstrap values		
	Point estimate	Median	Lower 95%	Upper 95 %
Ancestral N_e	74 422	54 872	49 553	67 125
<i>Pugetensis</i> N_e	7478	35 096	15 114	63 980
<i>Nuttalli</i> N_e	8802	17 785	7779	31 860
Ancient migration rate	0.00013	0.00006	0.00003	0.00013
Recent migration rate	0.00078	0.00374	0.00289	0.00506
Time of initial isolation with migration	45 372	45 324	40 585	63 789
Time of full isolation	9142	6225	8447	37 283
Time of secondary contact	2281	6139	8431	36 695

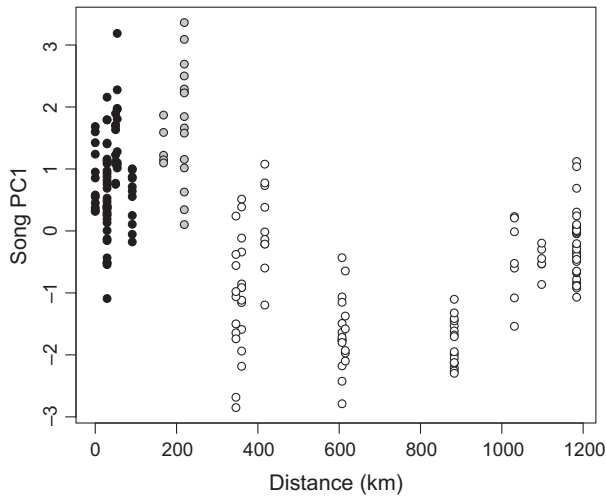


Fig. 3 Plot of song PC1 for *Z. l. nuttalli* (black), admixed individuals (grey) and *Z. l. pugetensis* (white) across hybrid zone transect.

In testing whether male response to own vs. other song varied among playback localities, the locality main effect was significant ($F_{3,56} = 4.16, P = 0.0099$), indicating that males from the four localities (sites 2, 6, 7 and 12) varied in their overall level of response to simulated intruders. The song main effect was also significant ($F_{1,56} = 40.1, P = 0.0001$), indicating that males responded more strongly to simulated territorial intruders with their own song than intruders with songs from other locations. The interaction term for this test was not significant ($F_{1,56} = 0.42, P = 0.42$), indicating that males in all locations gave a stronger response to ‘own’ than to ‘other’ songs.

In testing whether male response to ‘own’ vs. ‘other’ varied according to whether the subject males are from

Table 5 Results of statistical comparisons of response to ‘own’ vs. ‘other’ for each pairwise comparison using Wilcoxon sign-ranked tests. Significant *P* values indicated by a (*), and *d* is the post hoc calculated effect size

Experiments (Own vs. Other)	Response variable	<i>N</i>	<i>S</i>	<i>P</i>	<i>d</i>
Bolinas (2) vs. Bandon (12)	PC1	10	-25.5	0.0059*	1.5
Manchester (6) vs. Bolinas (12)	PC1	10	3.5	0.78	0.27
Manchester (6) vs. Bandon (12)	PC2	10	1.5	0.92	0.22
Manchester (6) vs. Bandon (12)	PC1	10	-21.5	0.0273*	0.86
MacKerricher (7) vs. Bolinas (2)	PC2	10	1.5	0.92	0.09
MacKerricher (7) vs. Bolinas (2)	PC1	10	-12.5	0.23	0.22
MacKerricher (7) vs. Bandon (12)	PC1	10	-21.5	0.0273*	0.99
Bandon (2) vs. Bolinas (12)	PC1	10	-23.5	0.0137*	0.92

parental or admixed populations, the genetic status main effect was significant ($F_{1,58} = 10.35, P = 0.0021$), indicating that putative admixed males responded more strongly to territorial intruders than did putative parental males (Fig. 4). Consistent with the previous model, the song main effect was also significant ($F_{1,58} = 41.3, P = 0.0001$), indicating that overall males responded more strongly to their ‘own’ than to the ‘other’ song phenotype. The interaction term for this test was not significant ($F_{1,58} = 2.69, P = 0.11$), indicating that both putative parental and putative admixed males showed similar levels of discrimination between ‘own’ and ‘other’ songs, although there was a trend towards weaker discrimination among putative admixed males (Fig. 4).

Song dissimilarity and F_{ST} predict strength of discrimination

Dissimilarity between own and other stimuli predicted the strength of behavioural discrimination. Individuals were more different in response to own vs. other song when there was greater dissimilarity between own and other song, such that they responded significantly less to songs more dissimilar to their own ($R^2 = 0.77, F_{1,4} = 13.41, P < 0.02$; Fig. S3, Supporting information). Likewise, pairwise genetic distance predicted the strength of behavioural discrimination. Individuals were more different in response to own vs. other song

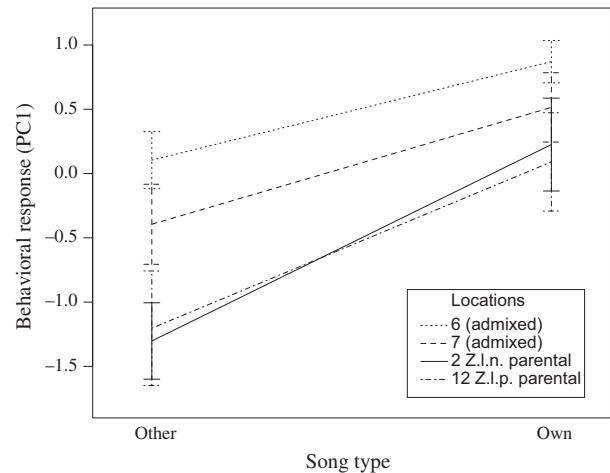


Fig. 4 Results of territorial playback experiments comparing male responses to their ‘own’ songs and the songs of the ‘other’ subspecies in four localities along a hybrid zone transect. Localities (from south to north): 2. Bolinas (solid line), 6. Manchester (dotted lines), 7. MacKerricher (dashed lines) and 12. Bandon (dash dot lines). Larger values of PC1 indicate a stronger behavioural response to simulated territorial intrusion. A steeper slope indicates a stronger discrimination between local and foreign stimuli. Bars represent standard error.

when there was a larger F_{ST} between the focal playback locality and stimulus locality, such that individuals responded less to songs from populations more genetically distant from their own ($R^2 = 0.88$, $F_{1,4} = 29.23$, $P < 0.006$; Fig. S4, Supporting information).

Isolation by song

Whistle length and song PC1, which included whistle length and average note length as significant loadings, were both stronger predictors of genetic distance than geography across localities. The regression coefficient for song PC1 distance ($\beta_S = 0.51$, $3 P = 0.0001$) was over twice as large as the regression coefficient for geographic distance ($\beta_D = 0.2$, $P = 0.12$), and the regression coefficient for whistle length distance ($\beta_S = 0.53$, $P = 0.0002$) was over seven times as large as the regression coefficient for geographic distance ($\beta_D = 0.075$, $P = 0.63$), suggesting that isolation by song explained genetic distance more strongly than isolation by distance for these parameters (Fig. S5, Supporting information).

Discussion

Overall, we found acoustic, behavioural and genetic evidence that *Z. l. nuttalli* and *Z. l. pugetensis* are distinct evolutionary units and that song is acting as a barrier to gene flow between them. Historical demographic inference suggests that the subspecies diverged relatively recently, and subspecific differences in song have been maintained in the face of hybridization. Putative parental individuals discriminate between the two subspecies based on song, as do putative admixed individuals. We found strong evidence that as songs become more dissimilar, males respond less to these songs in an important functional context: territory defence. There was some evidence that song is a weaker barrier in the hybrid zone, as putative admixed individuals tended to discriminate less strongly between songs of the two subspecies than did putative parental individuals, although the effect was not statistically significant. On further examination of song divergence, we found that whistle length—an important species recognition cue in song learning in this species—may also function in subspecies recognition.

Our genetic clustering analyses found that *Z. l. nuttalli* and *Z. l. pugetensis* are two distinct genetic populations. This contrasts with previous attempts that could not distinguish the subspecies using allozyme or mitochondrial loci (Corbin 1981; Weckstein & Zink 2001). In both our study and others, pairwise genetic distances among localities both within and between the subspecies were low (Corbin 1981; Zink & Barrowclough

1984). While these two subspecies are genetically distinct in allopatry, population assignment tests revealed genetically admixed individuals at three localities in the hybrid zone. Our simulation-based demographic analyses suggest that this admixture could be explained by secondary contact, with glacial refugia during the last glacial maximum a plausible mechanism for allopatric differentiation. However, further testing of more complex models with selection is warranted in future studies (Roux *et al.* 2016). While many studies of individuals with intermediate admixture proportions assume that taxa are exchanging genes in secondary contact, few explicitly test this model against other evolutionary scenarios (Payseur & Rieseberg 2016; but see Nadachowska-Brzyska *et al.* 2013). Coalescent simulations are a powerful tool for investigating the history of populations, but there are several drawbacks. For one, it can be computationally demanding for genome-scale data, especially as sample size and model complexity increase. Additionally, the stochastic nature of the coalescence process introduces some uncertainty into the estimated demographic parameters (Terhorst & Song 2015). Finally, we interpret these analyses with the caveat that although our results supported secondary contact, it is impossible to evaluate all possible historical scenarios of differentiation.

Songs were divergent between *Z. l. nuttalli* and *Z. l. pugetensis*. The best parameters to distinguish between the subspecies were whistle length and average note duration, and *Z. l. nuttalli* had longer whistles and shorter durations of other notes. Putative admixed individuals from Manchester (6) and MacKerricher (7) had songs more similar to *Z. l. nuttalli*, and individuals from Manchester (6) were also more genetically similar to this subspecies. The pattern of song divergence may provide insight into the features of song that individuals use to discriminate between subspecies. We found that whistle length was the best song parameter to distinguish between the subspecies. The whistle is a likely candidate as a conspecific marker for recognition, as it is universally present across song dialects for both subspecies and is the introductory component of their song. Although song is culturally inherited, song learning is directed by a genetic template (Nelson *et al.* 1995; Soha & Marler 2001a,b). Several song learning experiments with white-crowned sparrow nestlings indicate that the whistle is innate rather than learned (Whaling *et al.* 1997), important for acoustic imprinting (Margoliash 1983), and may reflect an innate template in conspecific song memorization, production and recognition (Whaling *et al.* 1997; Soha & Marler 2000). Thus, the whistle's importance in song learning and recognition for nestlings may also influence adult recognition of potential competitors and mates. Innate recognition of

conspecifics may be especially important given that in the nonbreeding season, southern localities contain a mixture of overwintering *Z. l. pugetensis* and resident *Z. l. nuttalli* (Blanchard 1941). A rich area of future study lies in exploring regions of the genome responsible for divergence in song between the subspecies, potentially related to innate aspects of song learning.

Males responded less strongly to heterotypic than homotypic song in parental populations. Signals that elicit a stronger territorial response from males are typically interpreted as signals more effective at maintaining a territory and competing for mates (Searcy & Nowicki 2005) (but see Baker & Mewaldt 1978; Baker *et al.* 1981 for the argument that a stronger response to heterotypic song can facilitate reproductive isolation). Therefore, we interpret a lower response to heterotypic song as evidence of a behavioural barrier between *Z. l. nuttalli* and *Z. l. pugetensis*, such that individuals of one subspecies would not be as effective in territory defence and mate acquisition in a population of the other subspecies. These results are consistent with previous studies that found males in parental populations are more responsive to their own subspecies song, both in white-crowned sparrows (Lampe & Baker 1994) and in other taxa (Turčoková *et al.* 2011; Greig & Webster 2013). Discrimination was not explained solely by familiarity with the local song type, as putative admixed individuals did not discriminate between their own songs and nonlocal songs *Z. l. nuttalli*, which more closely matched their genotype.

Male territorial responses to playbacks of bird songs are less logistically challenging than measuring female response and are therefore more typically used to test signal discrimination as an indirect measure of reproductive isolation (e.g. Irwin *et al.* 2001; Dingle *et al.* 2010; Podos 2010; Derryberry 2011; Turčoková *et al.* 2011), although theory also suggests that male–male competition can contribute directly to isolation (Ellers & Slabbekoorn 2003). A useful follow-up experiment could involve testing admixed and parental female preferences for admixed and parental songs, as testing females would provide more direct evidence of whether song is acting as a behavioural barrier between the subspecies. A study in the *Z. l. oriantha* subspecies indicated that females prefer their natal-dialect song over a foreign-dialect or heterospecific song (MacDougall-Shackleton *et al.* 2001), so we predict that females will discriminate between subspecies songs in the parental populations.

We not only found that individuals discriminated between homotypic and heterotypic song, but also that the strength of discrimination between songs was predicted by similarity of the stimulus song to the receiver's song. Finding this association supports the hypothesis that divergence in the signal itself is driving

behavioural discrimination between the signals. Covariation between signal and response has been found in some playback studies (e.g. Sosa-López *et al.* 2016), but not others (e.g. den Hartog *et al.* 2008). However, these studies focus on the strength of behavioural response to a stimulus (e.g. height of behavioural response PC1) rather than the strength of discrimination between stimuli (e.g. slope of behavioural response PC1). Our pairwise design enabled us to relate the difference in song stimuli directly to the difference in behavioural responses to those stimuli. The strength of discrimination between stimulus songs was also positively correlated with pairwise genetic distances for those playback localities. These associations among genetic distance, song divergence and behavioural response support the role of song as a behavioural isolating mechanism in the hybrid zone.

We also attempted to infer what processes are driving patterns of genetic, acoustic and behavioural divergence between subspecies. Multiple forces of selection may act on song, causing divergence among populations. Acoustic signals may diverge and converge via sexual and social selection (Fisher 1930; West-Eberhard 1983), acoustic adaptation to environmental conditions affecting sound transmission (Morton 1975; Wiley & Richards 1982; Derryberry 2009), morphological divergence in shape and size that constrains signal production (Podos 1996; Podos *et al.* 2004), genetic or cultural drift (Lemon 1975; Lynch 1996; Irwin *et al.* 2008) or a combination of these social, ecological and stochastic factors (Mundinger 1982; Price 1998; Wilkins *et al.* 2013). Divergent migratory behaviour and/or allopatric temporal isolation between the subspecies could also play a role in genetic differentiation (e.g. Ruegg *et al.* 2012; Delmore & Irwin 2014), given that *Z. l. pugetensis* is migratory and *Z. l. nuttalli* is a year-round resident. However, migratory behaviour as an isolating mechanism is not mutually exclusive with our hypothesis that song is a behavioural barrier between the subspecies. After controlling for geographic distance, we still found a significant association between whistle length and genetic differentiation, which suggests that this song feature could be driving genetic divergence between the subspecies. Many studies have looked for an association between song distance and genetic variation in white-crowned sparrows (Baker 1975; Baker *et al.* 1982; MacDougall-Shackleton & MacDougall-Shackleton 2001; Soha *et al.* 2004) and other taxa (Wright & Wilkinson 2001; Nicholls *et al.* 2006; Alstrom *et al.* 2007; Irwin *et al.* 2008; Kenyon *et al.* 2011; Sosa-López *et al.* 2016), but few other than our study have found this association independent of geographic variation (but see MacDougall-Shackleton & MacDougall-Shackleton 2001; Rendell *et al.* 2012).

Conclusions

Our aim in this study was to evaluate when and how song acts as an isolating mechanism to gain insight into the evolution and maintenance of behavioural barriers. Our findings provide strong support to the hypothesis that song is a behavioural barrier to gene flow between *Z. l. nuttalli* and *Z. l. pugetensis*, although clearly other barriers to mating are incomplete as introgression is ongoing. We investigated both the signal—bird song—and receiver response to that signal and found that both songs and behavioural responses to song are divergent between the subspecies, although these differences may be weakening in the contact zone and facilitating hybridization. Our finding that song and genetic distance predict the strength of behavioural discrimination between songs gives insight into the processes driving the evolution of this behavioural barrier.

Acknowledgements

We thank B. Wayne, four anonymous reviewers, and members of the Derryberry Lab writing workshop for helpful comments on earlier drafts of this manuscript. This research was conducted under Duke University Institutional Animal Care and Use Committee (IACUC) Protocol A099-06-03 and LSU IACUC protocol 10-009. The fieldwork in 2004–2005 was conducted under Federal Fish and Wildlife Banding Permit 22712-G and Collecting Permit MB-813248, California Scientific Collecting Permit 801208-05, Oregon Scientific Taking Permit 050-04 and Washington State Scientific Collection Permit 04-110. The fieldwork in 2010 was conducted under Federal Fish and Wildlife Collecting Permit MB-679782-2 and California Scientific Collecting Permit 802010-02. We thank J. Phillips for assistance in the field and S. Hunter for assistance in the laboratory. G. Derryberry, A. Xue, I. Prates and G. Seeholzer provided feedback on bioinformatics analyses. This material is based upon work supported by Louisiana State Board of Regents grants NSF(2010)-Pfund-167 and LEQSF(2012-25)-RD-A-21 to RTB and EPD, a National Science Foundation (NSF) Graduate Research Fellowship under Grant No. 1154145 to SEL, and the NASA Dimensions of Biodiversity Program and NSF DOB 1342578 and DEB 1253710 provided support to IO and MJH. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors(s) and do not necessarily reflect the views of the Louisiana Board of Regents, nor the National Science Foundation.

References

Alstrom P, Olsson U, Rasmussen PC *et al.* (2007) Morphological, vocal and genetic divergence in the *Cettia acanthizoides* complex (Aves: Cettiidae). *Zoological Journal of the Linnean Society*, **149**, 437–452.

Andersson M (1994) *Sexual Selection*. Princeton University Press, Princeton, New Jersey.

Baker MC (1975) Song dialects and genetic differences in white-crowned sparrows (*Zonotrichia leucophrys*). *Evolution*, **29**, 226–241.

Baker MC (1987) Integration of song between two subspecies of White-crowned Sparrows on the west coast of North America. *Ornis Scandinavica*, **18**, 265–268.

Baker MC, Cunningham MA (1985) The biology of bird-song dialects. *Behavioral and Brain Sciences*, **8**, 85–100.

Baker MC, Mewaldt LR (1978) Song dialects as barriers to dispersal in White-crowned Sparrows, *Zonotrichia leucophrys nuttalli*. *Evolution*, **32**, 712–722.

Baker MC, Thompson DB, Sherman GL, Cunningham MA (1981) The role of male vs male interactions in maintaining population dialect structure. *Behavioral Ecology and Sociobiology*, **8**, 65–69.

Baker MC, Thompson DB, Sherman GL, Cunningham MA, Tomback DF (1982) Allozyme frequencies in a linear series of song dialect populations. *Evolution*, **36**, 1020–1029.

Baker MC, Cunningham MA, Thompson AD (1984) Cultural and genetic differentiation of two subspecies of White-crowned Sparrow. *Condor*, **86**, 359–367.

Banks RC (1964) Geographic variation in the White-crowned Sparrow, *Zonotrichia leucophrys*. *University of California Publications in Zoology*, **70**, 1–123.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113–148.

Beeman K (1999) *SIGNAL*. Engineering Design, Belmont, Massachusetts.

Blanchard BD (1941) The White-crowned Sparrows (*Zonotrichia leucophrys*) of the Pacific seaboard: environment and annual cycle. *University of California Publications in Zoology*, **46**, 1–178.

Bradbury PJ, Zhang Z, Kroon DE *et al.* (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, **23**, 2633–2635.

Burt J (2001) *Syrinx*. Univ. of Washington, Seattle.

Campbell P, Pasch B, Pino JL *et al.* (2010) Geographic variation in the songs of neotropical singing mice: testing the relative importance of drift and local adaptation. *Evolution*, **64**, 1955–1972.

Carling MD, Lovette IJ, Brumfield RT (2010) Historical divergence and gene flow: coalescent analyses of mitochondrial, autosomal and sex-linked loci in passerina buntings. *Evolution*, **64**, 1762–1772.

Catchpole CK, Slater PJB (2008) *Bird Song: Biological Themes and Variations*. Cambridge University Press, Cambridge, Massachusetts.

Cohen J (1988) *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Corbin KW (1981) Genic heterozygosity in the White-crowned Sparrow: a potential index to boundaries between subspecies. *Auk*, **98**, 669–680.

Corbin KW, Wilkie PJ (1988) Genetic similarities between subspecies of the White-crowned Sparrow. *Condor*, **90**, 637–647.

Cortopassi AJ, Mewaldt LR (1965) The circumannual distribution of White-crowned Sparrows. *Bird-Banding*, **36**, 141–169.

Coyne JA, Orr AH (2004) *Speciation*. Sinauer Associates, Sunderland, Massachusetts.

Delmore KE, Irwin DE (2014) Hybrid songbirds employ intermediate routes in a migratory divide. *Ecology Letters*, **17**, 1211–1218.

Derryberry EP (2007) Evolution of bird song affects signal efficacy: an experimental test using historical and current signals. *Evolution*, **61**, 1938–1945.

- Derryberry EP (2009) Ecology shapes birdsong evolution: variation in morphology and habitat explains variation in White-crowned Sparrow song. *The American Naturalist*, **174**, 24–33.
- Derryberry EP (2011) Male response to historical and geographical variation in bird song. *Biology Letters*, **7**, 57–59.
- Dingle C, Poelstra JW, Halfwerk W, Brinkhuizen DM, Slabbekoorn H (2010) Asymmetric response patterns to subspecies-specific song differences in allopatry and parapatry in the gray-breasted wood-wren. *Evolution*, **64**, 3537–3548.
- Dobzhansky T (1940) Speciation as a stage in evolutionary divergence. *The American Naturalist*, **74**, 312–321.
- Durrett R, Buttel L, Harrison R (2000) Spatial models for hybrid zones. *Heredity*, **84**, 9–19.
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.
- Edwards SV, Kingan SB, Calkins JD *et al.* (2005) Speciation in birds: genes, geography, and sexual selection. *Proceedings of the National Academy of Sciences*, **102**, 6550–6557.
- Ellers J, Slabbekoorn H (2003) Song divergence and male dispersal among bird populations: a spatially explicit model testing the role of vocal learning. *Animal Behaviour*, **65**, 671–681.
- Elshire RJ, Glaubitz JC, Sun Q *et al.* (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, **6**, e19379.
- Ender JA (1977) *Geographic Variation, Speciation and Clines*. Princeton University Press, Princeton, New Jersey.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.
- Excoffier L, Dupanloup I, Huerta-Sánchez E, Sousa VC, Foll M (2013) Robust demographic inference from genomic and SNP data. *PLoS Genetics*, **9**, e1003905.
- Falush D, van Dorp L, Lawson D (2016) A tutorial on how (not) to over-interpret STRUCTURE/ADMIXTURE bar plots. *bioRxiv*, **66431**, 066431.
- Faul F, Erdfelder E, Buchner A, Lang A-G (2009) Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, **41**, 1149–1160.
- Field DL, Ayre DJ, Whelan RJ, Young AG (2011) Patterns of hybridization and asymmetrical gene flow in hybrid zones of the rare *Eucalyptus aggregata* and common *E. rubida*. *Heredity*, **106**, 841–853.
- Fisher R (1930) The genetical theory of natural selection. *Genetics*, **154**, 272.
- Gee JM (2005) No species barrier by call in an avian hybrid zone between California and Gambel's quail (*Callipepla californica* and *C. gambelii*). *Biological Journal of the Linnean Society*, **86**, 253–264.
- Gonzalez C, Ornelas JF (2014) Acoustic divergence with gene flow in a lekking hummingbird with complex songs. *PLoS ONE*, **9**, e109241.
- Grant BR, Grant PR (1996) Cultural inheritance of song and its role in the evolution of Darwin's finches. *Evolution*, **50**, 2471–2487.
- Greig EI, Webster MS (2013) Spatial decoupling of song and plumage generates novel phenotypes between 2 avian subspecies. *Behavioral Ecology*, **24**, 1004–1013.
- Grinnell J (1928) Notes on the systematics of west American birds. III. *Condor*, **30**, 185–189.
- Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009) Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics*, **5**, e1000695.
- Haavie J, Borge T, Bures S *et al.* (2004) Flycatcher song in allopatry and sympatry – convergence, divergence and reinforcement. *Journal of Evolutionary Biology*, **17**, 227–237.
- Halfwerk W, Dingle C, Brinkhuizen DM *et al.* (2016) Sharp acoustic boundaries across an altitudinal avian hybrid zone despite asymmetric introgression. *Journal of Evolutionary Biology*, **29**, 1356–1367.
- Harrison RG (1993) Hybrids and hybrid zones: historical perspective. In: *Hybrid Zones and the Evolutionary Process* (ed Harrison RG), pp. 3–12. University Press, Oxford.
- den Hartog PM, Slabbekoorn H, Ten Cate C (2008) Male territorial vocalizations and responses are decoupled in an avian hybrid zone. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**, 2879–2889.
- Hewitt GM (1988) Hybrid zones-natural laboratories for evolutionary studies. *Trends in Ecology & Evolution*, **3**, 158–167.
- Hickerson MJ, Stahl EA, Lessios HA (2006) Test for simultaneous divergence using approximate Bayesian computation. *Evolution*, **60**, 2435–2453.
- Irwin DE, Bensch S, Price TD (2001) Speciation in a ring. *Nature*, **409**, 333–337.
- Irwin DE, Thimman MP, Irwin JH (2008) Call divergence is correlated with geographic and genetic distance in greenish warblers (*Phylloscopus trochiloides*): a strong role for stochasticity in signal evolution? *Journal of Evolutionary Biology*, **21**, 435–448.
- Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.
- Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, **27**, 3070–3071.
- Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, **11**, 94.
- Kenyon HL, Toews DPL, Irwin DE (2011) Can song discriminate between MacGillivray's and Mourning Warblers in a narrow hybrid zone? *Condor*, **113**, 655–663.
- Kershner EL, Bollinger EK (1999) Aggressive response of chickadees towards Black-capped and Carolina chickadee calls in central Illinois. *Wilson Bulletin*, **111**, 363–367.
- de Kort SR, den Hartog PM, ten Cate C (2002) Diverge or merge? The effect of sympatric occurrence on the territorial vocalizations of the vinaceous dove *Streptopelia vinacea* and the ring-necked dove *S. capicola*. *Journal of Avian Biology*, **33**, 150–158.
- Kroodsma DE (1986) Design of song playback experiments. *Auk*, **103**, 640–642.
- Kroodsma DE (1990) Using appropriate experimental designs for intended hypotheses in 'song' playbacks, with examples

- for testing effects of song repertoire sizes. *Animal Behaviour*, **40**, 1138–1150.
- Kroodsma DE, Byers BE, Goodale E, Johnson S, Liu W-C (2001) Pseudoreplication in playback experiments, revisited a decade later. *Animal Behaviour*, **61**, 1029–1033.
- Lachlan RF (2007) Luscinia: a bioacoustics analysis computer program. Available at <http://rflachlan.github.io/Luscinia/>.
- Lachlan RF, Servedio MR (2004) Song learning accelerates allopatric speciation. *Evolution*, **58**, 2049–2063.
- Lampe HM, Baker MC (1994) Behavioral response to song playback by male and female white-crowned sparrows of two subspecies. *Bioacoustics*, **5**, 171–185.
- Leader N, Geffen E, Mokady O, Yom-Tov Y (2008) Song dialects do not restrict gene flow in an urban population of the orange-tufted sunbird, *Nectarinia osea*. *Behavioral Ecology and Sociobiology*, **62**, 1299–1305.
- Lemon RE (1975) How birds develop song dialects. *Condor*, **77**, 385–406.
- Louheed SC, Handford P (1992) Vocal dialects and the structure of geographic variation in morphological and allozymic characters in the Rufous-collared Sparrow, *Zonotrichia capensis*. *Evolution*, **46**, 1443–1456.
- Lu F, Lipka AE, Glaubitz J *et al.* (2013) Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. *PLoS Genetics*, **9**, e1003215.
- Luther DA, Derryberry EP (2012) Birdsongs keep pace with city life: changes in song over time in an urban songbird affects communication. *Animal Behaviour*, **83**, 1059–1066.
- Lynch A (1996) The population memetics of birdsong. In: *Ecology and Evolution of Acoustic Communication in Birds* (eds Kroodsma DE, Miller EH), pp. 181–197. Cornell University Press, Ithaca, New York.
- MacDougall-Shackleton EA, MacDougall-Shackleton SA (2001) Cultural and genetic evolution in mountain white-crowned sparrows: song dialects are associated with population structure. *Evolution*, **55**, 2568–2575.
- MacDougall-Shackleton SA, MacDougall-Shackleton EA, Hahn TP (2001) Physiological and behavioural responses of female mountain white-crowned sparrows to natal- and foreign-dialect songs. *Canadian Journal of Zoology*, **79**, 325–333.
- Margoliash D (1983) Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *The Journal of Neuroscience*, **3**, 1039–1057.
- Marler P, Tamura M (1964) Culturally transmitted patterns of vocal behavior in sparrows. *Science*, **146**, 1483–1486.
- Martens J (1996) Vocalizations and speciation of palearctic birds. In: *Ecology and Evolution of Acoustic Communication in Birds* (eds Kroodsma DE, Miller EH), pp. 221–240. Cornell University Press, Ithaca, New York.
- Matessi G, Dabelsteen T, Pilastro A (2000) Responses to playback of different subspecies songs in the Reed Bunting *Emberiza schoeniclus*. *Journal of Avian Biology*, **31**, 96–101.
- Mayr E (1963) *Animal Species and Evolution*. Harvard University, Cambridge, Massachusetts.
- McEntee JP (2014) Reciprocal territorial responses of parapatric African sunbirds: species-level asymmetry and intraspecific geographic variation. *Behavioral Ecology*, **25**, 1380–1394.
- McGregor PK, Catchpole CK, Dabelsteen T *et al.* (1992) Design of playback experiments: the Thornbridge Hall NATO ARW Consensus. *Playback and Studies of Animal Communication*, pp. 1–9. Plenum Press, New York.
- Mewaldt LR, Kibby SS, Morton ML (1968) Comparative biology of Pacific coastal white-crowned sparrows. *Condor*, **70**, 14–30.
- Morton ES (1975) Ecological selection on avian sounds. *The American Naturalist*, **109**, 12–34.
- Munding PC (1982) Microgeographic and macrogeographic variation in the acquired vocalizations of birds. In: *Acoustic Communication in Birds* (eds Kroodsma DE, Miller EH), pp. 147–208. Academic Press, New York.
- Nadachowska-Brzyska K, Burri R, Olason PI *et al.* (2013) Demographic divergence history of pied flycatcher and collared flycatcher inferred from whole-genome re-sequencing data. *PLoS Genetics*, **9**, e1003942.
- Nelson DA (2000) A preference for own-subspecies' song guides vocal learning in a song bird. *Proceedings of the National Academy of Sciences*, **97**, 13348–13353.
- Nelson DA, Soha JA (2004) Perception of geographical variation in song by male Puget Sound white-crowned sparrows, *Zonotrichia leucophrys pugetensis*. *Animal Behaviour*, **68**, 395–405.
- Nelson D, Marler P, Palleroni A (1995) A comparative approach to vocal learning: intraspecific variation in the learning process. *Animal Behaviour*, **50**, 83–97.
- Nicholls J, Austin J, Moritz C, Goldizen A (2006) Genetic population structure and call variation in a passerine bird, the satin bowerbird, *Ptilonorhynchus violaceus*. *Evolution*, **60**, 1279–1290.
- Nottebohm F (1969) The song of the chingolo, *Zonotrichia capensis*, in Argentina: description and evaluation of a system of dialects. *Condor*, **71**, 299–315.
- Nowicki S, Searcy WA, Hughes M (1998) The territory defense function of song in song sparrows: a test with the speaker occupation design. *Behaviour*, **135**, 615–628.
- Oomen RA, Reudink MW, Nocera JJ *et al.* (2011) Mitochondrial evidence for panmixia despite perceived barriers to gene flow in a widely distributed waterbird. *Journal of Heredity*, **102**, 584–592.
- Ortiz-Ramírez MF, Andersen MJ, Zaldívar-Riverón A, Ornelas JF, Navarro-Sigüenza AG (2016) Geographic isolation drives divergence of uncorrelated genetic and song variation in the Ruddy-capped Nightingale-Thrush (*Catharus frantzii*; Aves: Turdidae). *Molecular Phylogenetics and Evolution*, **94**, 74–86.
- Panhuis TM, Butlin R, Zuk M, Tregenza T (2001) Sexual selection and speciation. *Trends in Ecology & Evolution*, **16**, 364–371.
- Patten MA, Rotenberry J, Zuk M (2004) Habitat selection, acoustic adaptation, and the evolution of reproductive isolation. *Evolution*, **58**, 2144–2155.
- Payne RB, Westneat DF (1988) A genetic and behavioral analysis of mate choice and song neighborhoods in indigo buntings. *Evolution*, **42**, 935–947.
- Payseur BA, Rieseberg LH (2016) A genomic perspective on hybridization and speciation. *Molecular Ecology*, **25**, 2337–2360.
- Pegan TM, Rumelt RB, Dzielski SA *et al.* (2015) Asymmetric response of Costa Rican White-breasted Wood-Wrens (*Henicorhina leucosticta*) to vocalizations from allopatric populations. *PLoS ONE*, **10**, e0144949.

- Petrinovich L, Patterson TL (1981) The responses of white-crowned sparrows to songs of different dialects and subspecies. *Zeitschrift für Tierpsychologie*, **57**, 1–14.
- Pettengill JB, Moeller DA (2012) Phylogeography of speciation: allopatric divergence and secondary contact between outcrossing and selfing *Clarkia*. *Molecular Ecology*, **21**, 4578–4592.
- Podos J (1996) Motor constraints on vocal development in a songbird. *Animal Behaviour*, **51**, 1061–1070.
- Podos J (2001) Correlated evolution of morphology and vocal signal structure in Darwin's finches. *Nature*, **409**, 185–188.
- Podos J (2010) Acoustic discrimination of sympatric morphs in Darwin's finches: a behavioural mechanism for assortative mating? *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 1031–1039.
- Podos J, Warren PS (2007) The evolution of geographic variation in birdsong. *Advances in the Study of Behavior*, **37**, 403–458.
- Podos J, Huber SK, Taft B (2004) Bird song: the interface of evolution and mechanism. *Annual Review of Ecology and Systematics*, **35**, 55–87.
- Price T (1998) Sexual selection and natural selection in bird speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **353**, 251–260.
- Price T (2008) *Speciation in Birds*. Roberts and Company Publishers, Greenwood Village, Colorado.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Qvarnström A, Haavie J, Saether S, Eriksson D, Pärt T (2006) Song similarity predicts hybridization in flycatchers. *Journal of Evolutionary Biology*, **19**, 1202–1209.
- R Core Team (2015) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Raj A, Stephens M, Pritchard JK (2014) FastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*, **197**, 573–589.
- Rendell L, Mesnick SL, Dalebout ML, Burtenshaw J, Whitehead H (2012) Can genetic differences explain vocal dialect variation in sperm whales, *Physeter macrocephalus*? *Behavior Genetics*, **42**, 332–343.
- Rice WER (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.
- Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.
- Rosenberg NA, Nordborg M (2002) Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nature Reviews Genetics*, **3**, 380–390.
- Roux C, Fraise C, Romiguier J *et al.* (2016) Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biology*, **14**, e2000234.
- Ruegg K, Slabbekoorn H, Clegg S, Smith TB (2006) Divergence in mating signals correlates with ecological variation in the migratory songbird, Swainson's thrush (*Catharus ustulatus*). *Molecular Ecology*, **15**, 3147–3156.
- Ruegg K, Anderson EC, Slabbekoorn H (2012) Differences in timing of migration and response to sexual signalling drive asymmetric hybridization across a migratory divide. *Journal of Evolutionary Biology*, **25**, 1741–1750.
- Sall J (2015) JMP® Version 12. SAS Institute Inc., Cary, North Carolina, 1989–2007.
- Saranathan V, Hamilton D, Powell GVN, Kroodsmas DE, Prum RO (2007) Genetic evidence supports song learning in the three-wattled bellbird *Procnias tricarunculata* (Cotingidae). *Molecular Ecology*, **16**, 3689–3702.
- Searcy WA (1992) Measuring responses of female birds to male song. In: *Playback and Studies of Animal Communication* (ed McGregor PK), pp. 175–189. Plenum, New York.
- Searcy WA, Nowicki S (2005) *The Evolution of Animal Communication: Reliability and Deception in Signaling Systems*. Princeton University Press, Princeton, New Jersey.
- Searcy WA, Nowicki S, Hughes M (1997) The response of male and female song sparrows to geographic variation in song. *Condor*, **99**, 651.
- Searcy WA, Anderson RC, Nowicki S (2006) Bird song as a signal of aggressive intent. *Behavioral Ecology and Sociobiology*, **60**, 234–241.
- Secondi J, Bretagnolle V, Compagnon C, Faivre B (2003) Species-specific song convergence in a moving hybrid zone between two passerines. *Biological Journal of the Linnean Society*, **80**, 507–517.
- Seddon N, Tobias JA (2007) Song divergence at the edge of Amazonia: an empirical test of the peripatric speciation model. *Biological Journal of the Linnean Society*, **90**, 173–188.
- Shaw KL, Parsons YM, Lesnick SC (2007) QTL analysis of a rapidly evolving speciation phenotype in the Hawaiian cricket *Laupala*. *Molecular Ecology*, **16**, 2879–2892.
- Slabbekoorn H, Smith TB (2002) Bird song, ecology and speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **357**, 493–503.
- Soha JA, Marler P (2000) A species-specific acoustic cue for selective song learning in the white-crowned sparrow. *Animal Behaviour*, **60**, 297–306.
- Soha JA, Marler P (2001a) Vocal syntax development in the white-crowned sparrow (*Zonotrichia leucophrys*). *Journal of Comparative Psychology*, **115**, 172–180.
- Soha JA, Marler P (2001b) Cues for early discrimination of conspecific song in the white-crowned sparrow (*Zonotrichia leucophrys*). *Ethology*, **826**, 813–826.
- Soha JA, Nelson DA, Parker PG (2004) Genetic analysis of song dialect populations in Puget Sound white-crowned sparrows. *Behavioral Ecology*, **15**, 636–646.
- Sosa-López JR, Martínez Gómez JE, Mennill DJ (2016) Divergence in mating signals correlates with genetic distance and behavioural responses to playback. *Journal of Evolutionary Biology*, **29**, 306–318.
- Terhorst J, Song YS (2015) Fundamental limits on the accuracy of demographic inference based on the sample frequency spectrum. *Proceedings of the National Academy of Sciences*, **112**, 7677–7682.
- Thinh VN, Hallam C, Roos C, Hammerschmidt K (2011) Concordance between vocal and genetic diversity in crested gibbons. *BMC Evolutionary Biology*, **11**, 11–36.
- Turčoková L, Pavel V, Chutný B, Petrussek A, Petrusková T (2011) Differential response of males of a subarctic population of Bluethroat *Luscinia svecica svecica* to playbacks of their own and foreign subspecies. *Journal of Ornithology*, **152**, 975–982.
- Vavrek MJ (2011) Fossil: palaeoecological and palaeogeographical analysis tools. *Palaeontologia Electronica*, **14**, 1T.

- Venables WN, Ripley BD (2002) *Modern Applied Statistics With S*. Springer, New York.
- Wang IJ (2013) Examining the full effects of landscape heterogeneity on spatial genetic variation: a multiple matrix regression approach for quantifying geographic and ecological isolation. *Evolution*, **67**, 3403–3411.
- Warwick AR, Travis J, Lemmon EM (2015) Geographic variation in the Pine Barrens Treefrog (*Hyla andersonii*): concordance of genetic, morphometric and acoustic signal data. *Molecular Ecology*, **24**, 3281–3298.
- Weckstein J, Zink RM (2001) Anomalous variation in mitochondrial genomes of White-crowned (*Zonotrichia leucophrys*) and Golden-crowned (*Z. atricapilla*) sparrows: pseudogenes, hybridization, or incomplete lineage sorting? *The Auk*, **118**, 231–236.
- West-Eberhard MJ (1983) Sexual selection, social competition, and speciation. *The Quarterly Review of Biology*, **58**, 155–183.
- Whaling CS, Solis MM, Doupe A, Soha JA, Marler P (1997) Acoustic and neural bases for innate recognition of song. *Proceedings of the National Academy of Sciences*, **94**, 12694–12698.
- Wiley RH, Richards DG (1982) Adaptations for acoustic communication in birds: sound transmission and signal detection. In: *Acoustic Communication in Birds* (eds Kroodsma DE, Miller EH), pp. 131–181. Academic Press, New York.
- Wilkins MR, Seddon N, Safran RJ (2013) Evolutionary divergence in acoustic signals: causes and consequences. *Trends in Ecology & Evolution*, **28**, 156–166.
- Wilson DR, Bitton P, Podos J, Mennill DJ (2014) Uneven sampling and the analysis of vocal performance constraints. *The American Naturalist*, **183**, 214–228.
- Wright TF, Wilkinson GS (2001) Population genetic structure and vocal dialects in an Amazon parrot. *Proceedings of the Royal Society B: Biological Sciences*, **268**, 609–616.
- Zink RM, Barrowclough GF (1984) Allozymes and song dialects: a reassessment. *Evolution*, **38**, 444–448.

E.P.D. and R.T.B. conceived of the study. E.P.D. conducted the behavioural experiments and song analyses. I.O. conducted historical demographic modelling in fastsimcoal. S.E.L. conducted all other analyses and wrote the manuscript. All authors provided edits and comments to the manuscript.

Data accessibility

Sampling information and SNP calls for each sequenced individual is available at the Dryad Digital Repository (<https://doi.org/10.5061/dryad.5jc0t>). Scripts for the historical demographic inference are available on Github (<https://github.com/isaacovercast/WCS>). Songs for each recorded individual were archived at the Borror Laboratory of Bioacoustics.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Inter-class principal component analysis of 6419 loci dataset for 17 sampling localities.

Fig. S2 Best supported demographic model of secondary contact.

Fig. S3 Individuals discriminated more strongly between songs when songs were more dissimilar.

Fig. S4 Individuals discriminated more strongly between their local song and the song of the other subspecies when the pairwise genetic distance was greater.

Fig. S5 Multiple matrix regression with randomization (MMRR) analysis performed on genetic and song distance.

Table S1 Nine acoustic variables analyzed to compare songs between subspecies *Z. l. nuttalli* and *Z. l. pugetensis*.

Table S2 Pairwise F_{ST} matrix based on the 6419 SNP dataset, not including locality 8.

Table S3 AIC values for historical demography models.

Table S4 Mean \pm SD for behavioral responses to individual playback experiments and all experiments standardized and combined.