

Original Article

Extrapair paternity, song, and genetic quality in song sparrows

Christopher E. Hill,^a Çağlar Akçay,^b S. Elizabeth Campbell,^b and Michael D. Beecher^{b,c}

^aDepartment of Biology, Coastal Carolina University, P.O. Box 261954, Conway 29528, SC, USA,

^bDepartment of Psychology, University of Washington, Box 351525, Seattle 98195, WA, USA, and

^cDepartment of Biology, University of Washington, Box 351800, Seattle 98195, WA, USA

Most songbirds are socially monogamous, yet molecular studies have found that in most species, some offspring in the nest are sired by males other than the social mate of the female. The functional significance of extrapair paternity (EPP) in social monogamy is poorly understood, despite numerous theoretical and empirical studies in the last decades. We have examined EPP in the song sparrow (*Melospiza melodia*) using microsatellites and tested whether females choose as extrapair mates males that 1) had larger song repertoires, 2) shared more songs with their neighbors, 3) were more heterozygous, or 4) were less related to the females than the social mate of these. We found that 24% of offspring were sired by extrapair males and that the extrapair sires were invariably neighbors. However, neither song repertoire size nor song sharing with neighbors predicted a male's EPP success. Furthermore, neither heterozygosity of a male nor his relatedness to the female predicted EPP success. At the same time, males that did not lose paternity in their own nest or gain paternity in other nests tended to be younger. These results indicate that females are not using song repertoire size or song sharing as a basis for extrapair mate choice and are not likely to accrue significant genetic benefits from EPP. Instead, the occurrence and level of EPP in this population might be primarily governed by behavioral trade-offs between mate guarding and pursuing extrapair copulations. We suggest that detailed behavioral studies are needed to understand extrapair mating in this species. *Key words*: bird song, compatible genes, extrapair paternity, good genes, song sparrow. [*Behav Ecol* 22:73–81 (2011)]

Most songbirds are socially monogamous. Molecular studies have revealed, however, that in most socially monogamous songbirds, some females mate multiply, resulting in extrapair paternity (EPP; for a review, see Griffith et al. 2002). Although early studies of multiple mating in songbirds focused on multiple mating as a male strategy (e.g., Beecher MD and Beecher IM 1979), most of the recent research has focused on the female side of the equation, specifically, on the hypothesis that EPP may be a result of female mate choice benefiting the female (Petrie and Kempenaers 1998; Jennions and Petrie 2000). The standard research approach has been to identify the correlates of female mate choice in male traits, with much of the research focusing on traits that are expected to be under sexual selection.

Male song and EPP

Male song is a classical example of a sexually selected trait in songbirds that has been proposed to be the basis of female mate choice (reviewed in Catchpole and Slater 2008). As a complex trait, male song is likely to reflect multiple selection pressures and constraints (Gil and Gahr 2002), which means that not all aspects of song would be expected to be correlated with female mate choice. One aspect of male song, however, has been proposed to be important for mate choice, namely song repertoire size or complexity (Searcy and Andersson 1986).

Several studies have directly examined the link between EPP and song repertoire size. In one highly cited study, Hasselquist et al. (1996) found that great reed warbler (*Acrocephalus arundinaceus*) females ($n = 10$) preferred extrapair males with larger song repertoires than their social mates and that the repertoire size of the father was positively correlated with recruitment of offspring.

Not all studies, however, have found a positive correlation between repertoire size and extrapair mate choice. For instance, in a recent study in the sedge warbler (*A. schoenobaenus*), a congener to great reed warbler, females ($n = 10$) were shown to prefer males as extrapair mates that have smaller repertoires than their social mates (Marshall et al. 2007) despite evidence that females prefer males with larger repertoires as their social mates (Catchpole 2000). Several reviews have also concluded that repertoire size is not correlated generally with EPP success of a male (Garamszegi and Møller 2004; Akçay and Roughgarden 2007b; Byers and Kroodsma 2009).

Genetic diversity and EPP

Early research on genetic benefits in EPP assumed that females accrued additive genetic benefits from their extrapair partners for their offspring in the form of better alleles that determine fitness. This hypothesis has been called the “good genes” hypothesis. More recently, research has also focused on non-additive genetic benefits that females might accrue. Under this “compatible genes” hypothesis, females improve their offspring's genetic quality not by providing their offspring better alleles from the extrapair father but by increasing the offspring's heterozygosity (Brown 1997; Neff and Pitcher 2005). This can be accomplished either by choosing more distantly related males as extrapair fathers or by mating with more heterozygous males. Thus, extrapair males should be less related to the female than the social male, and/or they should be more heterozygous. Furthermore, extrapair young (EPY) are

Address correspondence to Ç. Akçay. E-mail: caglar@u.washington.edu.

Received 12 December 2009; revised 20 September 2010; accepted 21 September 2010.

expected to be more heterozygous than their maternal half siblings that are fathered by the social male (within-pair young, WPY) as has been found in the blue tit (*Cyanistes caeruleus*; Foerster et al. 2003).

Genetic diversity and male song

Some recent research has supported the hypothesis that sexually selected traits such as male song may reflect individual genetic diversity (i.e., heterozygosity) rather than better alleles at fitness-determining loci. For instance, in the sedge warbler, repertoire size has been shown to be correlated with the heterozygosity of the male (Marshall et al. 2003). Similarly, in an island population of song sparrows (*Melospiza melodia*), Reid et al. (2005) found that repertoire size was negatively correlated with the inbreeding level of a male estimated from a detailed pedigree. Therefore, the effect of sexually selected traits on EPP could be due to a compatible genes effect rather than (or in addition to) a good genes effect.

In the current study of a western population of song sparrows, we tested 2 hypotheses 1) that females choose their extrapair partners on the basis of sexually selected song traits and 2) that females accrue genetic benefits from EPP. The song sparrow is a socially monogamous species in which males have song repertoires of 5–13 songs. A contemporaneous study of another western population of song sparrows found fairly high levels of EPP (about 28% of offspring; O'Connor et al. 2006). Although there is evidence from laboratory studies that females prefer larger repertoires (Searcy and Marler 1981; Searcy 1984) and mixed evidence that repertoire size is positively correlated with social pairing success in song sparrows (Reid et al. 2004; but see Searcy 1984), no study to date has examined the effect of repertoire size on EPP in this species. If females show a preference for larger repertoires sizes, then extrapair fathers should have larger repertoires than social mates. Similarly, males with smaller repertoires should lose more paternity than do those with larger repertoires.

Repertoire size might not be the only trait that is correlated with male attractiveness. Previous research in our Seattle population has shown that song sharing with neighboring males (but not repertoire size) is positively correlated with territory tenure (Beecher et al. 2000), a consequence primarily of male–male competition. In addition, an earlier laboratory study with females from this population indicated that females preferred locally shared songs (O'Loughlen and Beecher 1999). Therefore, we tested the above predictions with respect to song sharing as well.

In addition to looking at song features of a male, we tested the predictions of the compatible genes hypothesis using the genetic data. These predictions are that 1) extrapair fathers should be less related to the female than her social mate is to her, 2) extrapair fathers should be more heterozygous than social mates, and 3) EPY should be more heterozygous than WPY in the same nest. We also tested whether social pairings were formed nonrandomly with respect to relatedness. Finally, we tested whether repertoire size or degree of song sharing was correlated with heterozygosity of the male in this population of song sparrows.

MATERIALS AND METHODS

Study site

Discovery Park, in Seattle, Washington, USA, is a 3 km² undeveloped wooded park bordering Puget Sound, where a long-term study of song sparrows is ongoing since 1986

(Beecher et al. 1994). Song sparrow habitat, in the under story of deciduous and mixed woods, is nearly continuous at our study site, only occasionally interrupted by mowed fields. Song sparrows at the study site are year-round residents. Territories are small (0.2–0.4 ha) and dense (30–40 per 10 ha). Territory boundaries were delineated by observing the movements and singing patterns of banded birds as well as interactions between neighbors at boundaries. Boundaries were marked onto small-scale maps.

Subjects

Sparrows were captured in mist nets or in Potter traps and banded with a unique combination of metal (US Fish and Wildlife Service) leg band and 3 colored plastic bands. Song sparrows that were adults in this study were generally banded as juveniles in their first summer or fall or on territory establishment in spring. Males in our population establish their territories by early spring in the year following their natal summer. Due to the long-term banding and surveying effort in our study population, territory tenure and thus age were known for most males in the study. All males for which we had territory tenure data had disappeared from our study site by the end of 2002 (the study was conducted between 1996 and 1998). A small blood sample (20–150 µl) was taken from each bird at the time of capture. A total of 360 individuals (females, males, and offspring) were genotyped.

Finding nests

In 1995 and 1996, we closely monitored 5–10 song sparrow territories and found most nests that fledged young in those territories. In 1997 and 1998, we monitored an area that included 35–40 territories and found all the nests that hatched in those territories. We generally found nests when the female was nest building or incubating. We visited the nest when chicks were 7 days old to band chicks and take blood samples. Song sparrow chicks usually fledge at 10 days old and achieve independence at about 30 days. We did not confirm that all banded chicks fledged.

Recording songs and measuring song characteristics

Male song sparrows sing discrete song types and deliver them with what has been termed eventual variety, usually repeating each song 5–20 times before switching to a new type. The modal repertoire size in this population is 8 songs. Each male's song repertoire was recorded onto cassette tapes using Sennheiser ME-88 directional microphones and Sony TC-D5M tape recorders. A bird's repertoire was considered fully recorded if we recorded at least 15 consecutive switches between song types or at least 20 switches if recording was not continuous (Cassidy 1993). The average bird was recorded for >30 total switches. Tapes were analyzed on a Kay DSP 5500 Sonograph. We printed sonograms of all song types and all major variants of each song type (Stoddard et al. 1988; Podos et al. 1992). A song sparrow does not modify his repertoire size after his first year (Nordby et al. 2002).

Shared songs, defined as songs showing close similarity for 50–100% of their length, were identified from printed sonograms by consensus of 3 independent judges. We calculated a sharing index between each pair of neighboring birds equal to $2N_s/(R_1 + R_2)$, where N_s is the number of shared songs and R_1 and R_2 are the repertoire sizes of the 2 birds (McGregor and Krebs 1982). We also calculated for each male the average of his sharing indexes with all adjacent neighbors.

Genetic analysis

On collection, all blood samples were immediately stored in 750 μ l lysis buffer (0.1 M Tris-HCl, pH 8, 0.1 M ethylenediaminetetraacetic acid, 0.01 M NaCl, and 0.5% sodium dodecyl sulfate). We stored samples in buffer at room temperature or 4 °C for up to 3 years before DNA extraction, with no apparent degradation of the DNA. DNA was extracted from samples with a phenol-chloroform extraction, followed by ethanol precipitation. We amplified 7 polymorphic microsatellite DNA loci for each sparrow: Mme2, Mme7, Mme8, and Mme12 (Jeffery et al. 2001); Escu1 and GF05 (Hanotte et al. 1994; Petren 1998); and Psap335 (Temple M, Leonard M, Hamilton L, personal communication). One locus, Mme7, is sex linked in song sparrows carried on the Z chromosome. Amplifications (8 μ l total volume) were carried out in 1 \times polymerase chain reaction (PCR) reaction buffer (Promega, Madison, WI), with 0.5 mM dNTPs, 1.5–3.0 mM MgCl₂, with forward and reverse primers at 0.05–0.2 mM, and 0.3 units *Taq* DNA polymerase. Forward primers were fluorescently labeled. PCR products were separated using an automated DNA sequencer (ABI 373A; PE Biosystems, Inc., Foster City, CA). We analyzed the results using GeneScan and Genotyper software (PE Biosystems, Inc.). Alleles were recognized based on the differences in length. DNA from one bird was used as a positive control and amplified and electrophoresed on each gel to check for sizing accuracy.

Paternity exclusion power, *Q* (Weir 1996), varied across loci from 0.28 to 0.82, averaging 0.66. Combined exclusion power of all 7 loci was 0.9991. Using 7 loci, the power to exclude 4 random males (a typical number of immediate neighbors) was 0.9946 and to exclude 12 males (which would exclude all males within 2 territories) was 0.9894. Power to exclude 40 males (the maximum number of males at the study site in any season) was still 0.9650. Power to exclude neighboring males when only 6 loci were available for comparison remained high: 0.9873 for 4 males, 0.9748 for 12, and 0.9186 for 40.

Paternity assignment

We compared each chick's genotype to that of its mother. If a chick mismatched its mother at a locus, the data for that locus were discarded. Mother-chick mismatches occurred in 4% of comparisons, presumably due to a combination of genotyping errors and germ line mutations. If a chick matched its mother at a given locus, that information was used to identify the chick's paternal alleles at that locus. The paternal alleles of the chick were then compared with the genotypes of all males in the population at the time the chick was hatched. A preliminary analysis of cases where one male matched a chick perfectly at all loci and every other adult male in the population mismatched the chick at 2 or more loci (i.e., a tabulation of unequivocal paternity assignments) showed that of 20 extrapair (EPF) fathers in that sample, 19 held territories adjacent to the natal territory of the chicks they had fathered and the 20th was only one territory further away. We therefore disregarded males more than 2 territory widths distant in subsequent analyses if there was a local male that matched as well. A similar spatial pattern of EPP was found in an island population of song sparrows (O'Connor et al. 2006).

Social fathers were assigned paternity of a chick if they mismatched the chick at no more than one locus (usually none) and also matched the chick better than any other male within 2 territory widths of the nest. Extrapair males were assigned paternity if the social father had been rejected and the extrapair male mismatched the chick at no more than one locus (again, usually none) and no other nearby male matched the chick as well. If a social father and a neighbor matched

a chick equally well, paternity was considered equivocal and that chick was excluded from further analyses.

Heterozygosity and relatedness measures

One locus (Mme7) was not used in heterozygosity and relatedness analyses because it is a sex-linked locus. We also excluded all individuals that were typed at less than 4 loci (not counting Mme7) from the analyses involving heterozygosity and relatedness. One hundred and sixty-eight individuals were typed at all 6 autosomal loci, 122 at 5, and 40 at 4 loci. The rest of the individuals (including 9 chicks) were not considered for these analyses. We used the measure homozygosity by loci (HL; Aparicio et al. 2006) to estimate heterozygosity. This measure gives a measure of homozygosity that is weighted by the heterozygosity at each locus and has been shown to outperform both internal relatedness (Amos et al. 2001) and unweighted homozygosity (Aparicio et al. 2006). For the purpose of presentation, we used $1 - \text{HL}$, which we refer to as HL henceforth. HL scores were calculated using a Microsoft Excel Macro authored by Bill Amos (University of Cambridge, Cambridge, UK, <http://www.zoo.cam.ac.uk/zoostaff/amos/>).

Pairwise relatedness values were calculated using the software ML-Relate (Kalinowski et al. 2006). The program uses a simulation to calculate a maximum likelihood estimate of relatedness between 0 and 1 and is able to accommodate null alleles. As a further estimate of individual males' outbreeding level, we calculated the mean relatedness of each male to the genotyped females in the population ($n = 37$ females).

Statistical analyses

We carried out 2 types of comparisons. First, we compared males that lost at least one EPF with those that did not lose an EPF in their nest. Second, we carried out pairwise comparisons in which we compared the social father of an EPY with the extrapair father. Some males ($n = 13$) were present in multiple years. For age analyses, to avoid pseudoreplication, we took the first year the male was present in the study (for males that never lost an EPF), the first year the male had an EPF in his nest (for males that lost an EPF), or the first year a male gained an EPF (for males gaining at least an EPF). When a male lost EPF to multiple extrapair fathers, we took the average values weighted by the number of chicks these extrapair fathers sired in pairwise comparisons (so that, for instance, a male that sired 2 offspring was twice as heavily weighted as a male that sired a single offspring). Statistical analyses were carried out in SPSS 14.0 (SPSS Inc., Chicago, IL).

Consistent with recent discussions of the weaknesses of the null hypothesis testing approach, our statistical analyses focus primarily on effect sizes and their associated confidence intervals (CIs). We report unstandardized effect sizes and 95% CI in the text along with means and standard deviations (SDs). This approach has the advantage of giving the reader an appreciation of where the real effect size is likely to be in place of categorical statements that a particular effect is simply significant or not significant (Nakagawa and Cuthill 2007). We have also calculated standardized effect sizes (Hedges' *d*) and estimated 95% CI with the formulas given in Nakagawa and Cuthill (2007) for the major variables of interest. The standardized effect sizes allow us to compare our results with previous studies. These effect sizes and associated CIs are given in Table 1, supplemented with standard *t*-tests and their associated *P* values for the major analyses.

To test whether social pairs form nonrandomly with respect to relatedness, we performed a permutation test based on 38 random pairs (the number of social pairings in our study we were able to estimate genetic relatedness for) drawn from all

Table 1

Effect sizes (calculated from means and pooled SDs), associated 95% CIs, and the results of *t*-tests and their associated *p* values for the central variables examined

Trait	Hedges' <i>d</i>	95% CIs for effect sizes		Null hypothesis significance tests	
		Lower limit	Upper limit	<i>t</i> values (df)	<i>P</i> values
Unpaired samples					
Song repertoire size	0.103	-0.637	0.850	0.30 (26)	0.77
Song sharing	-0.010	-0.788	0.768	0.03 (24)	0.98
Age	-0.630	-1.287	0.001	-2.02 (37)	0.05
Relatedness to social mate	-0.100	-0.739	0.536	-0.31 (36)	0.76
Heterozygosity	0.135	-0.491	0.767	0.51 (37)	0.61
Paired samples					
Song repertoire size	-0.144	-0.972	0.684	0.34 (12)	0.74
Song sharing	0.001	-0.423	0.424	0.00 (12)	0.99
Age	-0.025	-0.606	0.557	0.10 (18)	0.92
Relatedness to social mate	0.216	-0.411	0.844	0.70 (17)	0.49
Heterozygosity	0.393	-0.315	1.102	1.13 (18)	0.27
Offspring heterozygosity	-0.507	-1.242	0.229	-1.46 (14)	0.17

Positive effects indicate that nonlosers (extrapair mates) have larger trait values than losers (social mates). In the case of relatedness, positive effect sizes indicate lesser relatedness of the nonloser (extrapair mates) to the female.

possible pairings of males and females and calculated the mean relatedness in these pairs. The random pairings were iterated 10 000 times, resulting in a distribution of mean relatedness among random pairs. The observed mean relatedness in social pairings was compared with this distribution. The permutation test was carried out using Matlab R2007 (Mathworks, Natick, MA).

RESULTS

Extrapair fertilizations

Forty-five of 191 chicks (24.0%) were the result of EPFs. For 9 additional chicks, we had insufficient data to determine paternity unequivocally; these chicks are not included in any of the subsequent analyses. Twenty-six of 72 nests (36.1%) had at least 1 extrapair chick, and 3 nests (5.5%) had more than 1 extrapair father. Thirty-nine social males successfully reared a clutch during the study (many of them multiple nests within or across breeding years resulting in 72 nests), and 19 (48.7%) of these males lost at least one extrapair fertilization in their nest. Of these 19 males, 9 (47%) gained paternity in at least one other nest. Only 5 of 20 (25%) males that did not lose paternity in their own nest gained paternity in another nest. This tendency to either both gain and lose EPFs or to neither gain nor lose EPFs, however, was not significant ($\chi^2 = 5.20$, degrees of freedom [df] = 3, $p = 0.16$). Of the 45 extrapair chicks, we were able to unambiguously assign paternity for 44. The one chick for which we were not able to assign paternity was apparently fathered by a male from which we had not obtained a blood sample as the chick mismatched the social father and all neighbors at 3 or more loci and mismatched all other males bled in that year at 2 or more loci. The proportion of chicks in the population that were due to EPFs did not vary significantly between years (range 20–30%, $\chi^2 = 4.88$, df = 3, $P = 0.18$), so data from all years were combined.

Song repertoire size and sharing

Among the males that raised at least one brood, the song repertoires of 28 males were recorded (15 lost EPF and 13 did not lose EPF). The males that lost an EPF in their nest had similar repertoire sizes to those that did not lose an EPF

(8.47 ± 1.46 vs. 8.61 ± 1.12 , Figure 1a, mean difference = 0.14, 95% CI: -1.17 to 0.87, df = 26). In pairwise comparisons, social and extrapair fathers also had very similar repertoire sizes, with the social fathers' repertoire size being slightly larger (8.54 ± 1.44 vs. 8.33 ± 1.38 , mean difference = -0.21, 95% CI: -1.56 to 1.14, df = 12). In 6 (of 13) cases, the extrapair father had a larger repertoire than the social father; in 3 cases, they were equal; and in 4 cases, the social father had a larger repertoire (Figure 2).

For 26 males (15 lost EPF and 11 did not lose EPF), song sharing information with all neighbors was available. Average sharing with neighbors did not differ between males that lost an EPF and those that did not ($18.3 \pm 17.9\%$ vs. $18.1 \pm 21.4\%$, respectively, mean difference = 0.2%, CI: -15.8 to 16.2%, df = 24). In pairwise comparisons, social and extrapair fathers did not differ in their sharing level with their neighbors ($20.5 \pm 18.3\%$ vs. $20.5 \pm 16.2\%$, respectively, mean difference = 0.0%, 95% CIs: -8.2 to 8.2%, df = 12).

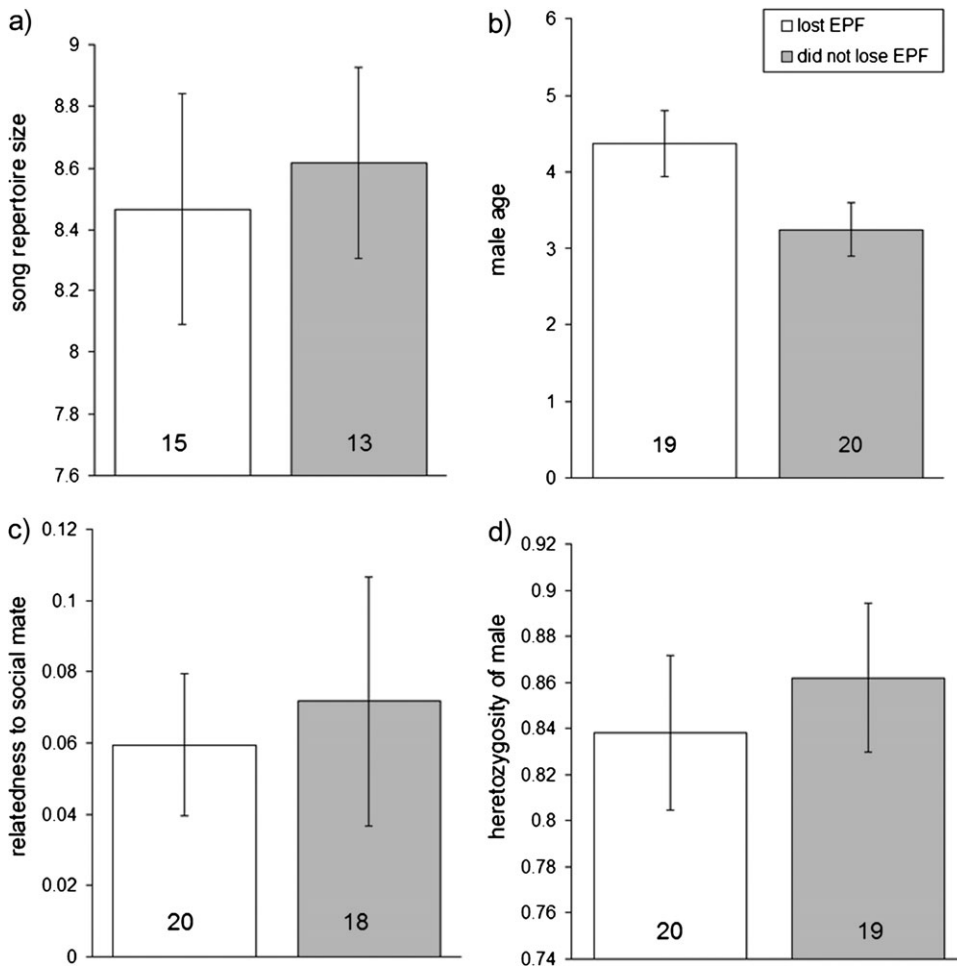
Age

Males that lost at least one EPF tended to be older than those that did not lose an EPF in their nests, the average difference being about 1 year (4.36 vs. 3.25 , mean difference = 1.11, 95% CI: -0.001 to 2.23, df = 37; Figure 1b). In pairwise comparisons, however, social and extrapair fathers had similar ages (4.36 vs. 4.32 , mean difference = 0.05, 95% CIs: -1.03 to 1.13, df = 18). Males that gained at least one EPF tended to be older than those that (as far as we know) did not gain an EPF (4.17 vs. 3.23 , mean difference = 0.93, 95% CIs: -0.03 to 1.89, df = 60).

Relatedness to social mate and female population

The observed mean relatedness in social pairings was 0.065, which was not significantly different ($P = 0.23$) from the mean of the distribution of average relatedness in random pairings (mean $r = 0.081$).

Males that lost at least one EPF in their nest were not on average more or less related to their social partners than those that did not lose an EPF (lost EPF, $n = 18$, mean $r = 0.059$ vs. did not lose EPF, $n = 20$, mean $r = 0.071$, mean difference = -0.012, 95% CI: -0.09 to 0.67, df = 36; Figure 1c). In pairwise comparisons, social and extrapair fathers did not differ in

**Figure 1**

Comparisons between males that lost at least one extrapair fertilization in their own nest versus males that did not lose in their (a) song repertoire size, (b) age, (c) relatedness to their social mate, and (d) heterozygosity. The numbers in each bar refer to the sample size. Error bars indicate ± 1 standard error of the mean.

their relatedness to the female (0.072 vs. 0.103, mean difference = -0.03 , 95% CI: -0.12 to 0.06 , $df = 17$).

We looked at correlations of average relatedness of a male to the female population with his song repertoire size ($n = 43$) and sharing ($n = 32$; for males present at multiple years, we took the first year in which sharing information with all neighbors was available). Neither repertoire size (Pearson's

$r = 0.00$, $P = 0.99$) nor song sharing (Pearson's $r = -0.26$, $P = 0.16$) was significantly correlated with average relatedness to the female population.

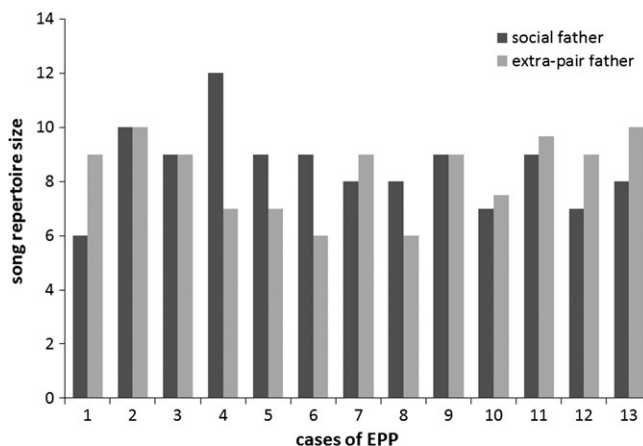
Heterozygosity of males

Males that lost at least one EPF in their nest were not less heterozygous than those that did not lose an EPF (mean $HL \pm SD$: 0.84 ± 0.15 vs. 0.86 ± 0.14 , mean difference = 0.02 , 95% CI: -0.07 to 0.11 , $df = 37$; Figure 1d). In pairwise comparisons, social males were less heterozygous than extrapair males (mean $HL \pm SD$: 0.86 ± 0.14 vs. 0.91 ± 0.11) but only slightly so (mean difference = -0.05 , 95% CI: -0.14 to 0.04 , $df = 18$). Heterozygosity was not correlated with either repertoire size (Pearson's $r = -0.02$, $P = 0.90$, $n = 43$) or sharing (Pearson's $r = 0.05$, $P = 0.78$, $n = 32$).

Heterozygosity of offspring

Of the 26 broods with an EPY, 15 also contained at least one WPY. For each of these broods, we calculated the mean heterozygosity of WPY and EPY. WPY in these broods were on average more heterozygous than their half siblings but not significantly so (WPY mean $HL \pm SD$: 0.78 ± 0.11 and EPY: 0.72 ± 0.12 , mean difference = 0.06 , 95% CIs: -0.15 to 0.03 , $df = 14$).

As a check, we tested whether offspring heterozygosity was correlated with relatedness of the parents. A total of 61 male–female pairings for which we had relatedness estimates produced at least one offspring during the study, including

**Figure 2**

Pairwise comparisons of song repertoire sizes of social fathers and genetic (extrapair) fathers of the EPY in each of the 13 cases of EPY where this information was available. The average difference is slightly negative (extrapair fathers have slightly smaller repertoires).

within-pair and extrapair fertilizations. Offspring heterozygosity was significantly and negatively correlated with relatedness of the parents (Pearson's $r = -0.28$, $P < 0.03$, $n = 61$).

Territory tenure, quality of males, and extrapair success

We tested whether a male's territory tenure (or survival on territory) was correlated with his repertoire size and heterozygosity. Territory tenure was not correlated with song repertoire size (Pearson's $r = -0.097$, $P = 0.61$, $n = 30$) or heterozygosity (Pearson's $r = 0.03$, $P = 0.88$, $n = 35$).

We also looked for a correlation between extrapair mating success of a male and territory tenure (range: 1–8 years). For this analysis, we combined all EPFs that a male gained and lost over the study. As a result, each male had a single score (positive if gained more EPFs than he lost, negative if vice versa, and zero if he did not gain or lose any or if he gained and lost equal numbers). This measure of EPF success also was not significantly correlated with survival on territory (Pearson's $r = 0.26$, $P = 0.13$, $n = 35$).

DISCUSSION

We tested multiple hypotheses suggesting that females choose extrapair mates based on song and/or accrue indirect (genetic) benefits as a result of extrapair fertilizations. We failed to find support for either class of hypotheses in any of the measures we have taken. Most of the effect sizes we found were quite small and at times inconsistent with the predicted effects. We also tested for correlations between song repertoire size and the genetic diversity of the males but again failed to find any evidence for strong relationship between these variables. We discuss these results below. First, however, we address the issue of power.

Power, bias, and confidence

The possibility of low power haunts all studies but particularly field studies with modest sample sizes (Jennions and Møller 2003). Traditionally, researchers assumed that when a small effect size (a nonsignificant effect in a null hypothesis significance test, such as t -tests with high P values) is found, low observed power is to blame. Under this view, nonsignificant findings would be more credible had we achieved high observed power and still failed to find effect. Consequently, authors are usually asked to carry out post hoc power analyses in order to determine the observed power of their tests. This argument, although commonly seen in literature, is flawed. The reason is that there is a one-to-one correspondence between a given P value and observed power in a given statistical test to reject a null hypothesis; the higher the P , the lower the power (Hoenig and Heisey 2001; Colegrave and Ruxton 2003). Therefore, as Hoenig and Heisey (2001, p. 20) put it, "computing the observed power after observing the P value should cause nothing to change about our interpretation of the P value.". Furthermore, decisions to publish or not based on achieved power run into the problem of creating a publication bias, which is further discussed below.

Therefore, in line with recent suggestions (Colegrave and Ruxton 2003; Nakagawa and Cuthill 2007), we have presented the observed effect sizes and 95% CIs (both in terms of mean difference and in standardized form in Table 1) for the effects of interest. The effect sizes and the CIs together allow a more sophisticated understanding of what the real effect might be and further facilitate comparisons with similar studies. We note that our CIs are rather large, which implies that we have a rather imprecise estimate of the real effect size. This caveat, however, is not unique to our study but applies to most field

studies, which tend to have modest sample sizes, even if they find large effects. Further studies with larger sample sizes would be the ideal solution. However, large sample sizes, particularly for behavioral traits like song repertoire size, are hard to achieve. An alternative is to compare and compile results from multiple studies in a meta-analysis, which is facilitated by reporting effect sizes and associated CIs. Although a full-fledged meta-analysis is beyond the scope of this paper, 2 relevant meta-analyses on extrapair mating in birds have been published in recent years (Garamszegi and Møller 2004; Akçay and Roughgarden 2007b), and in the following discussion, we draw heavily on these meta-analyses.

Male song repertoires and extrapair mate choice

We found no evidence that females choose extrapair partners based on their song repertoire size or song sharing with their neighbors. The effect sizes were small and not always consistent with the hypothesis of female mate choice for larger repertoires or more sharing (Table 1). Before we compare our results with other studies, we address a potential criticism: It could have been that in the cases where extrapair fertilizations occurred, the females were constrained in their choice as there were not any males with larger repertoires in the neighborhood. However, in 9 of 13 cases of EPP where we knew both the genetic and the social father's repertoire size, there was a neighboring male with a larger repertoire than the social male; in the remaining 4 cases, the social mate had the largest repertoire among the neighbors (according to the repertoire size hypothesis, we should not have seen any extrapair fertilizations in these cases). Had females chosen the neighboring male with the largest repertoire size in the 9 cases where potential extrapair sires with larger repertoires were available, we would have observed a marked difference in the song repertoire of social fathers (mean = 8.00) and ideal extrapair fathers (9.88). The mean difference of social and extrapair fathers in this "ideal-world scenario" (1.88 songs) contrasts sharply with the actual mean difference in our results (-0.21 songs). Thus, we can safely conclude that availability of larger repertoire males was not a constraint that kept females from showing a strong preference. We also note that our sample size, although modest, is still larger than that of most other comparable studies. In particular, to the best of our knowledge, our sample sizes (28 birds for the unpaired comparisons and 13 pairs for the paired comparisons) are larger than those of almost all studies to date that tested the hypothesis that song repertoire size is a basis of female extrapair mate choice (the only exceptions we know of are Bouwman et al. 2007; Byers 2007).

To compare our results with the results of previous studies, we plotted the standardized effect sizes for secondary sexual traits (including song traits) obtained from the meta-analysis of Akçay and Roughgarden (2007b) against the sample sizes of those studies (Figure 3). The effect sizes are from direct (paired) comparisons of extrapair mates and social mates; thus, we included only the paired tests for song repertoire size and song sharing from our study (see Akçay and Roughgarden 2007b, for details of the analysis). Positive effects refer to higher trait values in the extrapair mates. There have been only a few studies that compared song repertoire size between social and extrapair mates directly. As discussed in the INTRODUCTION, 2 of these studies, of the congeneric great reed warblers and sedge warblers, reported rather large effect sizes, positive and negative, respectively. The rest of the studies, including the current one, reported very small effect sizes on either side of zero. Therefore, there seems to be no general effect of song repertoire size in extrapair mate choice. Garamszegi and Møller (2004) came to the same conclusion

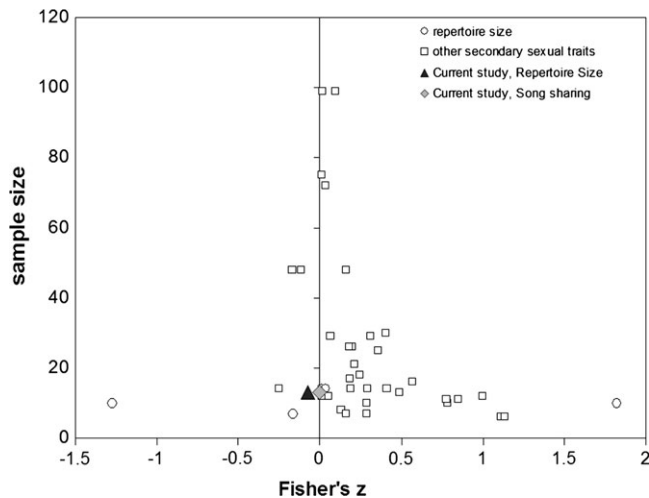


Figure 3

Plot of sample size versus effect size for secondary sexual traits obtained from the meta-analysis by Akçay and Roughgarden (2007b). The effect sizes are from paired comparisons between extrapair mates and social mates, with positive effect sizes indicating higher trait values for extrapair mates. Open circles are effect sizes for previous studies looking at song repertoire, open squares are for the rest of the secondary sexual traits, and the closed triangle and diamond are for the song repertoire size and song sharing comparisons from the current study (converted to Fisher's z using the same procedure as in Akçay and Roughgarden (2007b)).

in their meta-analysis of a larger set of studies (including comparisons between EPF losers and nonlosers).

As can be seen in Figure 3, most of the studies with large effects also have small samples. Furthermore, for small samples, there seems to be a preponderance of large effect sizes on the right side (higher trait values for extrapair mates compared with social mates), suggesting a publication bias against studies with small sample sizes that failed to find a large positive effect. Such a publication bias has the potential to distort the general patterns occurring in nature. Indeed, Akçay and Roughgarden (2007b) identified a statistically significant bias, and after correcting for it, the overall effect of male secondary sexual characteristics (which include song repertoire size) on EPF success became nonsignificant.

Also especially relevant to our results are earlier studies examining whether female song sparrows prefer males with larger song repertoires (Searcy 1984; Hiebert et al. 1989; Reid et al. 2004). Looking at social mate choice, these studies have produced only equivocal support for female choice based on repertoire size. Searcy (1984) found evidence that in the laboratory, females solicited more in response to playback of larger repertoires but failed to find any evidence in the field that males with larger repertoires had higher mating success. On the other hand, 2 studies in an island population of song sparrows in British Columbia (Hiebert et al. 1989; Reid et al. 2004) found that males with larger repertoires had higher probabilities of acquiring mates in their first year.

The pattern of equivocal support for female choice for large or complex repertoires is not unique to song sparrows. Recently, Byers and Kroodsma (2009) reviewed all studies testing for female choice for larger repertoires in songbirds (in both social and extrapair mate choice) and found that although laboratory studies tended to support female choice for larger repertoires, most field studies failed to find any effect. Therefore, there seems only equivocal evidence for female mate choice based on song repertoire size either in song sparrows or as a general pattern in songbirds.

Relatedness, heterozygosity, and extrapair mate choice

We also tested whether females accrue nonadditive genetic benefits by choosing extrapair mates that are either less related or more heterozygous than their social mate and again failed to find a strong and consistent effect of relatedness or heterozygosity of males in the extrapair mate choice of females. Although extrapair males tended to be slightly more heterozygous than social mates, this could be due to the fact that using the same set of markers to estimate heterozygosity and determine parentage creates a bias toward detected extrapair males being more heterozygous (Wetzel and Westneat 2009). Furthermore, in broods with both WPY and EPY, WPY were in fact slightly more heterozygous than EPY.

Another issue is that we used a limited number of loci in determining relatedness and heterozygosity (between 4 and 6, with most individuals typed at 5 and 6 loci). Microsatellite estimates of genome-wide heterozygosity and individual relatedness improve with the number of loci used, and our numbers are at the low end of studies in this vein so far published. However, we have carried out a "reality check" by looking at whether parental relatedness predicted offspring heterozygosity and were able to detect a significant negative correlation based on our estimates. Therefore, we believe that our estimates are informative. We also note that some previous studies (e.g., Foerster et al. 2003) found highly significant effects even with as few as 5 microsatellites, prompting Mays et al. (2008, p. 156) to suggest in a recent review that "researchers should not ignore testing genetic complementarity solely because they believe they have too few markers."

Interestingly, Reid (2007) found that song repertoire size correlated negatively with the average kinship of a male to the female population (determined from genealogical data) in the song sparrow population on Mandarte Island. Together with the finding that males with larger repertoires have higher (social) mating success in their first year (Reid et al. 2004), this suggests that female song sparrows on Mandarte Island could select as mates males less related to themselves if they used song repertoire size as a proxy (although an earlier study found no evidence for disassortative mating in the same population; Keller and Arcese 1998). We tested whether there was such a correlation in our population (with comparable sample sizes) but failed to find any correlation between song repertoire size and either heterozygosity of the male or the average relatedness of a male to the female population. This raises the interesting possibility that females in inbred populations might show stronger directional preferences for sexually selected traits, which become correlated with genetic diversity only at significant levels of inbreeding. Unfortunately, no extrapair mate choice data have been reported for the Mandarte population, which displays similar levels of EPFs to our population (O'Connor et al. 2006).

What is driving EPP in song sparrows?

Given that we could find no evidence that females choose on the basis of any of the male characteristics we have measured (song repertoire size, song sharing, genetic relatedness, and heterozygosity), how can we explain the occurrence of a significant level of EPP in song sparrows? At least 4 possibilities exist. The first possibility is that there is female choice for some male trait but that female choice is constrained by the ecological conditions or the female's own genetic quality such that only a fraction of females actually show choice (Gowaty 1996). This constraint may stem from the intense mate guarding by males or the need for the social mate's contribution to parental care. Because this hypothesis predicts that only some females will show preferences, experimental approaches are

probably needed to test this hypothesis with enough power (e.g., Hoi-Leitner et al. 1999).

The second possibility is that female choice is based not on male traits that reflect genetic quality but instead on direct benefits females obtain from males that they mate with (Akçay and Roughgarden 2007a). Such direct benefits have been shown in several species (Gray 1997; Tryjanowski and Hromada 2005; Rubenstein 2007). Song sparrows show no mate feeding or cooperative breeding. One likely form of direct benefit that females can accrue is foraging on an extrapair mate's territory.

As a third possibility, females may not accrue any benefit from extrapair mating but may simply be mating with whichever male they happen on. In most cases, this will be the social mate, assuming females and males spent most of their time on their own territory. This possibility is essentially a null hypothesis with respect to the function of EPP from the female perspective. Note that even if "any dude will do" for the females, males would still be expected to evolve strategies that will maximize their own paternity both in their own nests and extrapair matings. Interestingly, in our data, the only trend toward an effect was that males that lost at least one EPF in their own nest were on average older than those that did not lose an EPF. This, however, was not due to a preference for younger males. In fact, males that gained at least one EPF tended to be older than those that did not gain an EPF during our study. Together, these results suggest that most EPFs may be gained and lost between older males, which may point to different age-dependent strategies in the trade-off between pursuing EPCs versus mate guarding. This is an interesting possibility that warrants further investigation.

Finally, the last possibility is that extrapair mating is not a female strategy at all but instead extrapair males harass females into mating with them, presumably inflicting a net adaptive cost on the females (Arnqvist and Kirkpatrick 2005). Although theoretically possible, we suggest that this last possibility is very unlikely in song sparrows: In this species, copulation requires significant female cooperation (as it does in almost all songbirds), and females can further exert control over the situation by simply flying away and recruiting their own mate to avoid harassment.

Recently, Westneat and Stewart (2003) reviewed the variation in the pattern and occurrence of EPP in birds and forcefully argued that we need more studies on the behavioral events that lead to EPP. We fully agree with this suggestion. The above mentioned hypotheses in fact can only be distinguished with detailed behavioral observations of the behavior of both males and females during the breeding season. We have started doing just that by employing radiotelemetry to observe the movement patterns and behavior of the pairs during the breeding season in our population of song sparrows.

FUNDING

National Science Foundation (Grant number #9632193, research grants to M.D.B., a graduate fellowship to C.E.H., and a Doctoral Dissertation Improvement Grant to Ç.A.).

Animal use was approved the Institutional Animal Care and Use Committee (protocol #2207-03) at the University of Washington. We thank Erol Akçay for making the data from the meta-analysis available and Bill Searcy for his comments on the manuscript.

REFERENCES

- Akçay E, Roughgarden J. 2007a. Extra-pair parentage: a new theory based on transactions in a cooperative game. *Evol Ecol Res.* 9: 1223–1243.
- Akçay E, Roughgarden J. 2007b. Extra-pair paternity in birds: review of genetic benefits. *Evol Ecol Res.* 9:855–868.
- Amos W, Worthington Wilmer J, Fullard K, Burg TM, Croxall JP, Bloch D, Coulson T. 2001. The influence of paternal relatedness on reproductive success. *Proc R Soc Lond B Biol Sci.* 268:2021–2027.
- Aparicio JM, Ortego J, Cordero PJ. 2006. What should we weigh to estimate heterozygosity, alleles or loci? *Mol Ecol.* 15:4659–4665.
- Arnqvist G, Kirkpatrick M. 2005. The evolution of infidelity in socially monogamous passerines; the strength of direct and indirect selection on extra-pair copulation behavior in females. *Am Nat.* 165: S26–S37.
- Beecher MD, Beecher IM. 1979. Sociobiology of bank swallows: reproductive strategy of the male. *Science.* 205:1282–1285.
- Beecher MD, Campbell SE, Nordby JC. 2000. Territory tenure in song sparrows is related to song sharing with neighbors, but not to repertoire size. *Anim Behav.* 59:29–37.
- Beecher MD, Campbell SE, Stoddard PK. 1994. Correlation of song learning and territory establishment strategies in the song sparrow. *Proc Natl Acad Sci.* 91:1450–1454.
- Bouwman KM, van Dijk RE, Wilmenga JJ, Komdeur J. 2007. Older male reed buntings are more successful at gaining extra-pair fertilizations. *Anim Behav.* 73:15–27.
- Brown JL. 1997. A theory of mate choice based on heterozygosity. *Behav Ecol.* 8:60–65.
- Byers BE. 2007. Extra-pair paternity in chestnut-sided warblers is correlated with consistent vocal performance. *Behav Ecol.* 18:130–136.
- Byers BE, Kroodsmma DE. 2009. Female mate choice and songbird song repertoires. *Anim Behav.* 77:13–22.
- Cassidy ALEV. 1993. Song variation and learning in island populations of song sparrows [PhD dissertation]. Vancouver (Canada): University of British Columbia.
- Catchpole CK. 2000. Sexual selection and the evolution of song and brain structure in *Acrocephalus* warblers. *Adv Study Behav.* 29: 45–97.
- Catchpole CK, Slater PJB. 2008. Bird song: biological themes and variations. Cambridge(UK): Cambridge University Press.
- Colegrave N, Ruxton GD. 2003. Confidence intervals are a more useful complement to nonsignificant tests than are power calculations. *Behav Ecol.* 14:446–450.
- Foerster K, Delhey K, Johnsen A, Lifjeld JT, Kempenaers B. 2003. Females increase offspring heterozygosity and fitness through extra-pair matings. *Nature.* 425:714–717.
- Garamszegi LZ, Møller AP. 2004. Extra-pair paternity and the evolution of bird song. *Behav Ecol.* 15:508–519.
- Gil D, Gahr M. 2002. The honesty of bird song: multiple constraints for multiple traits. *Trends Ecol Evol.* 17:133–141.
- Gowaty PA. 1996. Battles of the sexes and origins of monogamy. In: Black JM, editor. *Partnership in birds: the study of monogamy.* Oxford: Oxford University Press. p. 21–52.
- Gray EM. 1997. Female red-winged blackbirds accrue material benefits from copulating with extra-pair males. *Anim Behav.* 53:625–639.
- Griffith SC, Owens IPF, Thuman KA. 2002. Extra pair paternity in birds: a review of interspecific variation and adaptive function. *Mol Ecol.* 11:2195–2212.
- Hanotte O, Zanon C, Pugh A, Greig C, Dixon A, Burke T. 1994. Isolation and characterization of micro-satellite loci in a Passerine bird—the reed bunting *Emberiza schoeniclus*. *Mol Ecol.* 3:529–530.
- Hasselquist D, Bensch S, von Schantz T. 1996. Correlation between male song repertoire, extra-pair paternity and offspring survival in the great reed warbler. *Nature.* 381:229–232.
- Hiebert SM, Stoddard PK, Arcese P. 1989. Repertoire size, territory acquisition and reproductive success in the song sparrow. *Anim Behav.* 37:266–273.
- Hoening JM, Heisey DM. 2001. The abuse of power: the pervasive fallacy of power calculations for data analysis. *Am Stat.* 55:19–24.
- Hoi-Leitner M, Hoi H, Romero-Pujante M, Valera F. 1999. Female extra-pair behavior and environmental quality in the serin (*Serinus serinus*): a test of the "constrained female hypothesis". *Proc R Soc Lond B Biol Sci.* 266:1021–1026.
- Jeffery KJ, Keller LF, Arcese P, Bruford MW. 2001. The development of microsatellite loci in the song sparrow, *Melospiza melodia* (Aves) and genotyping errors associated with good quality DNA. *Mol Ecol Notes.* 1:11–13.
- Jennions MD, Møller AP. 2003. A survey of the statistical power of research in behavioral ecology and animal behavior. *Behav Ecol.* 14:438–445.

- Jennions MD, Petrie M. 2000. Why do females mate multiply? a review of the genetic benefits. *Biol Rev.* 75:21–64.
- Kalinowski ST, Wagner AP, Taper ML. 2006. ML-Relate: a computer program for maximum likelihood estimation of relatedness and relationship. *Mol Ecol Notes.* 6:576–579.
- Keller LF, Arcese P. 1998. No evidence for inbreeding avoidance in a natural population of song sparrows (*Melospiza melodia*). *Am Nat.* 152:380–392.
- Marshall RC, Buchanan KL, Catchpole CK. 2003. Sexual selection and individual genetic diversity in a songbird. *Proc R Soc Lond B Biol Sci.* 270:S248–S250.
- Marshall RC, Buchanan KL, Catchpole CK. 2007. Song and female choice for extrapair copulations in the sedge warbler, *Acrocephalus schoenobaenus*. *Anim Behav.* 73:629–635.
- Mays HL, Albrecht T, Liu M, Hill GE. 2008. Female choice for genetic complementarity in birds: a review. *Genetica.* 134:147–158.
- McGregor PK, Krebs JR. 1982. Song types in a population of great tits (*Parus major*)—their distribution, abundance and acquisition by individuals. *Behaviour.* 79:126–152.
- Nakagawa S, Cuthill IC. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol Rev.* 82:591–605.
- Neff BD, Pitcher TE. 2005. Genetic quality and sexual selection: an integrated framework for good genes and compatible genes. *Mol Ecol.* 14:19–38.
- Nordby JC, Campbell SE, Beecher MD. 2002. Adult song sparrows do not alter their song repertoires. *Ethology.* 108:39–50.
- O'Connor KD, Marr AB, Arcese P, Keller LF, Jeffery KJ, Bruford MW. 2006. Extra-pair fertilization and effective population size in the song sparrow *Melospiza melodia*. *J Avian Biol.* 37:572–578.
- O'Loughlen AL, Beecher MD. 1999. Mate, neighbor and stranger songs: a female song sparrow perspective. *Anim Behav.* 58:13–20.
- Petren K. 1998. Microsatellite primers from *Geospiza fortis* and cross-species amplification in Darwin's finches. *Mol Ecol.* 7:1782–1784.
- Petrie M, Kempnaers B. 1998. Extra-pair paternity in birds: explaining variation between species and populations. *Trends Ecol Evol.* 13:52–58.
- Podos J, Peters S, Rudnicki T, Marler P, Nowicki S. 1992. The organization of song repertoires in song sparrows: themes and variations. *Ethology.* 90:89–106.
- Reid JM. 2007. Secondary sexual ornamentation and non-additive genetic benefits of female mate choice. *Proc R Soc Lond B.* 274:1395–1402.
- Reid JM, Arcese P, Cassidy ALEV, Hiebert SM, Smith JNM, Stoddard PK, Marr AB, Keller LF. 2004. Song repertoire size predicts initial mating success in male song sparrows, *Melospiza melodia*. *Anim Behav.* 68:1055–1063.
- Reid JM, Arcese P, Cassidy ALEV, Marr AB, Smith JNM, Keller LF. 2005. Hamilton and Zuk meet heterozygosity? Song repertoire size indicates inbreeding and immunity in song sparrows (*Melospiza melodia*). *Proc R Soc Lond B Biol Sci.* 272:481–487.
- Rubenstein DR. 2007. Female extra-pair mate choice in a cooperative breeder: trading sex for help and increasing offspring heterozygosity. *Proc R Soc Lond B Biol Sci.* 274:1895–1903.
- Searcy WA. 1984. Song repertoire size and female preferences in song sparrows. *Behav Ecol Sociobiol.* 14:281–286.
- Searcy WA, Andersson M. 1986. Sexual selection and the evolution of song. *Ann Rev Ecol Syst.* 17:507–533.
- Searcy WA, Marler P. 1981. A test for responsiveness to song structure and programming in female sparrows. *Science.* 213:926–928.
- Stoddard PK, Beecher MD, Willis MS. 1988. Response of territorial male song sparrows to song types and variations. *Behav Ecol Sociobiol.* 22:124–130.
- Tryjanowski P, Hromada M. 2005. Do males of the great grey shrike, *Lanius excubitor*, trade food for extra-pair copulations? *Anim Behav.* 69:529–533.
- Weir BS. 1996. Genetic data analysis II: methods for discrete population genetic data. Sunderland (MA): Sinauer.
- Westneat DF, Stewart IRK. 2003. Extra-pair paternity in birds: causes, correlates and conflict. *Ann Rev Ecol Evol Syst.* 34:365–396.
- Wetzel DP, Westneat DF. 2009. Heterozygosity and extra-pair paternity: biased tests result from the use of shared markers. *Mol Ecol.* 18:2010–2021.