JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.13183

Characterizing selection in black-throated blue warblers using a sexual network approach

E. R. A. CRAMER*†, S. A. KAISER‡, M. S. WEBSTER†, T. S. SILLETT* & T. B. RYDER*

*Migratory Bird Center, Smithsonian Conservation Biology Institute, Washington, DC, USA
†Cornell Lab of Ornithology, Ithaca, NY, USA
‡Center for Conservation Genomics, Smithsonian Conservation Biology Institute, Washington, DC, USA

Keywords:

body size;
competition;
selection gradient;
sexual selection;
social network;
social selection.

Abstract

Our understanding of trait evolution is built upon studies that examine the correlation between traits and fitness, most of which implicitly assume all individuals experience similar selective environments. However, accounting for differences in selective pressures, such as variation in the social environment, can advance our understanding of how selection shapes individual traits and subsequent fitness. In this study, we test whether variation in the social environment affects selection on individual phenotype. We apply a new sexual network framework to quantify each male's social environment as the mean body size of his primary competitors. We test for direct and social selection on male body size using a 10-year data set on black-throated blue warblers (*Setophaga caerulescens*), a territorial species for which body size is hypothesized to mediate competition for mates. We found that direct selection on body size was weak and nonsignificant, as was social selection via the body size of the males' competitors. Analysing both types of selection simultaneously allows us to firmly reject a role for body size in competitive interactions between males and subsequent male fitness in this population. We evaluate the application of the sexual network approach to empirical data and suggest that other phenotypic traits such as song characteristics and plumage may be more relevant than body size for male–male competition in this small passerine bird.

Introduction

Phenotypic selection analysis, whereby fitness is regressed on variation in a trait (Lande & Arnold, 1983), is a widely applied tool that has provided important insights into evolutionary processes (Siepielski *et al.*, 2011; Kingsolver *et al.*, 2012; Morrissey & Hadfield, 2012). However, a weakness of standard selection analyses is that all individuals are treated equally, even though they may experience different selective environments (Heisler & Damuth, 1987). This weakness has been mostly recognized in terms of the social environment (Moore *et al.*, 1997; Wolf *et al.*, 1999; McGlothlin *et al.*, 2010), which we define here as the phenotype of the individual's direct competitors (Aspi *et al.*, 2003; Formica *et al.*, 2011; Goodnight, 2015). The social environment can affect how an individual's phenotype relates to its reproductive success. For example, behavioural studies show that both contest outcomes (Arnott & Elwood, 2009) and mate choice (Wagner, 1998; Callander *et al.*, 2011, 2012) depend more on an individual's phenotype relative to its competitors than on its absolute phenotype. These studies provide strong rationale to control for variation in the social environment in selection analysis, if contests with conspecifics are an important determinant of fitness through indirect impacts on mating success. Controlling for the social environment could therefore improve our ability to detect environmentally induced variation in selection pressure, particularly when an individual interacts with a nonrandom subset of the population (Heisler & Damuth, 1987; McDonald *et al.*, 2013). Despite the

Correspondence: Emily R. A. Cramer, Migratory Bird Center, Smithsonian Conservation Biology Institute, MRC 5503 Box 37012, Washington, DC 20013, USA.
Tel.: +1 202-633-4035; fax: +1 202-633-1237;
e-mail: erc25@cornell.edu

importance of the social environment, few studies of selection have been designed to account for this variation, so the extent to which the phenotype of an individual's competitors affects selection on the individual's traits is not well understood. In this study, we assess the impact of an individual's own body size on fitness and the impact of the social environment – the body size of the competitors – on phenotypic selection. Hereafter, we refer to selection on an individual due to its social environment as 'social selection' (*sensu* Wolf *et al.*, 1999; see also Moore *et al.* 1997; McGlothlin *et al.*, 2010).

Correctly identifying an individual's relevant competitors is an important step in evaluating social selection. Previous work on social selection typically assumes that individuals located within a certain distance of each other, or within the same habitat patch, constitute the relevant competitors (Aspi *et al.*, 2003; Formica *et al.*, 2011). While this assumption is often upheld, because social interactions are often spatially constrained, descriptions of the social context can be further improved using social or sexual networks (McDonald *et al.*, 2013). For example, a sexual network could be built based on aggressive interactions between males for access to females and on subsequent copulations with females. Each male's relevant competitors are then identified based on network connections, and researchers can assess selection on male phenotypes while controlling for the mean phenotype of his competitors. Note that we use competition broadly to include both intrasexual contests and intersexual choice. Simulations show that this method substantially improves the ability to detect selective forces when males primarily compete against other males with a similar phenotype (McDonald *et al.*, 2013), but the approach has not yet been tested with empirical data.

A particularly salient feature of competitors' phenotypes may be body size. Intrasexual selection favours large body size in diverse taxa, because larger individuals tend to win physical contests against smaller individuals (Arnott & Elwood, 2009). However, other factors, such as differences in motivation and resource value, can also affect the outcome of contests (Renison *et al.*, 2000; Lindström & Pampoulie, 2005). Winning physical contests may, in turn, lead to larger males having greater access to females and/or controlling higher-quality resources that can improve reproductive success directly, for example by increasing fecundity or offspring survival (Kelly, 2008). Females may also prefer to mate with larger males either because they control higher-quality resources (Kelly, 2008) or because of direct preferences and/or pre-existing biases (Basolo, 2004; Callander *et al.*, 2011). Finally, larger males may be more able to overwhelm female resistance to courtship and copulations, in systems where forced copulations occur. Likely due to such advantages, body size is positively selected in a diversity of taxa (Kingsolver *et al.*, 2012).

Here, we test the hypothesis that selection favours larger-bodied males in a territorial, socially monogamous songbird, the black-throated blue warbler (*Setophaga caerulescens*). The opportunity for sexual selection is substantial in this species (Webster *et al.*, 2001), and competition with conspecifics likely plays an important role in male reproductive success. Males defend territories, and territory quality has a strong effect on reproductive success (Rodenhouse *et al.*, 2003; Kaiser *et al.*, 2015). Territory size is constrained by the density of conspecifics (Sillett *et al.*, 2004), suggesting an important role for competition in territory establishment. More competitive males may be able to settle on higher-quality territories (i.e. more food resources) and/or to increase the total amount of resources they defend by maintaining larger territories. Competition over access to females also appears to be important, as approximately 35% of offspring are sired by extra-pair males (Kaiser *et al.*, 2015). Males guard their social mates during the fertile stage, and males that mate guard effectively are less likely to be cuckolded, suggesting that extra-pair paternity is driven by both female choice and male–male competition (Chuang-Dobbs *et al.*, 2001). Previous work found no difference in body size between extra-pair males and the males they cuckolded (Webster *et al.*, 2001), but the role of body size in mediating competition for high-quality territories and for mating opportunities has not been thoroughly investigated.

We use a sexual network (McDonald *et al.*, 2013) to test predictions about how the variable social environment (i.e. the body size of competitors) affects both the fitness of the focal male and phenotypic selection on the focal male's body size. We use territory quality (estimated via an index of on-territory food abundance) and genetic reproductive success (i.e. the number of offspring a male sired, accounting for extra-pair paternity) as measures of fitness. Territory quality is relevant to fitness both via potential survival benefits for the individual and because it directly affects reproductive success (Rodenhouse *et al.*, 2003; Kaiser *et al.*, 2015). Assuming that large body size confers a competitive advantage, we predicted that larger males would defend higher-quality territories and produce more offspring ('direct' selection independent of the social environment) and that males with larger competitors would defend lower-quality territories and produce fewer offspring (social selection). If males interact primarily with similarly sized individuals, direct and social selection could counteract each other, such that statistically accounting for the bias introduced by social selection would enable us to detect direct selection. In contrast, if males interact with a random subset of other males, accounting for variation in the social environment should simply improve our power to detect direct selection by addressing some of the variation in how an individual's body size relates to his reproductive success. By identifying the relevant competitors with an unprecedented level of precision, in a system where extra-pair mating behaviours and territory quality effects are well described with long-term data, this study advances our understanding of the dynamics of selection in a complex social environment.

Materials and methods

Study species and field methods

We used data from 10 breeding seasons (2006–2015) from a population of black-throated blue warblers at the Hubbard Brook Experimental Forest, Woodstock, New Hampshire, USA (43°56′N, 71°45′W; for details on the long-term study, see Holmes, 2011). We included data from three study plots (low elevation, 250–350 m, 85 ha; mid elevation, 450–600 m, 65 ha; and high elevation, 750–850 m, 35 ha) that differ in a number of characteristics, including the density and age structure of black-throated blue warblers, habitat quality and weather conditions (Table S1; Rodenhouse *et al.*, 2003; Cline *et al.*, 2013; Kaiser *et al.*, 2014, 2015). Male black-throated blue warblers return to their breeding grounds approximately one week before females and defend territories against other males, with females then choosing a male (and territory) on which to breed (Holmes *et al.*, 2005). Nonterritorial floaters are not present at appreciable levels in this study population (Marra & Holmes, 1997; Sillett *et al.*, 2004). Moreover, nearly all offspring produced in the population can be assigned to sampled territorial males (Webster *et al.*, 2001; Kaiser *et al.*, 2017a); those young not assigned to sampled males are typically from nests near the edge of our study site and are likely to have been sired by males defending territories off-site. Males guard their social females during the female's peak of fertility by staying within a short distance of the female and following her as she moves through the forest (Chuang-Dobbs *et al.*, 2001). Copulations are observed only rarely in this species (Marra, 1993), and forced copulations have not been observed. Females build nests and incubate the eggs; both sexes feed nestlings and fledglings, with lepidopteran larvae representing the main food source (Rodenhouse & Holmes, 1992). Females are slightly smaller than males (Holmes *et al.*, 2005; Table S2). High-quality habitat is characterized as having a high density of understory leaves, which provides nesting and foraging substrates, and a high abundance of lepidopteran larvae (Rodenhouse *et al.*, 2003).

Reproductive success data were collected in the field by monitoring nests of individually marked warblers. Adult birds were captured using mist nets and banded with a unique combination of one USGS numbered aluminium band and two or three coloured plastic leg bands. We located nests primarily by observing females carrying nesting material, with some nests found at later stages by observing parental behaviours. Nest searching was conducted daily on all plots, and nests were checked every 2–3 days until fledging or failure. Therefore, nearly all nests that survived to the nestling feeding stage were detected and monitored. Nestlings were banded with a single aluminium band on day six after hatching. We took a blood sample from all individuals by puncture of the brachial vein and stored blood in lysis buffer (White & Densmore, 1992) until genetic analysis.

We assessed territory quality by estimating an index of food availability within mapped male territory boundaries, following Kaiser *et al.* (2015). Briefly, this index is calculated by multiplying an index of the abundance of lepidopteran larvae per leaf, which is derived from bi-weekly insect counts along transects, by an estimate of leaf abundance on the territory, which is obtained by extracting estimated leaf abundance within the territory boundaries from a kriged surface of leaf abundance across the entire study plot (see Kaiser *et al.* (2015) and supplementary materials for more details). We use this index of territory quality as a measure of male success.

Adult morphological measurements

We collected morphological data from each adult at capture. We measured unflattened wing chord with a standard wing rule to the nearest 0.1 mm, tarsus length to the nearest 0.1 mm with digital callipers and tail length (i.e. the length of the longest rectrix) to the nearest 0.5 mm with a standard ruler. We determined age (first breeding year vs. second or later breeding year) and sex based on plumage traits (Holmes *et al.*, 2005). We did not analyse body mass because mass can change dramatically over a short time and is likely related to environmental conditions.

Most adults were only captured and measured once, but some individuals were captured multiple times within a season or in different years. For recaptured individuals, body size measurements were repeatable across years (*sensu* Nakagawa & Schielzeth, 2010; see Table S2), so we calculated a single estimate of body size for each individual. This estimate of body size accounted for differences among banders and age-related changes in feather-based measurements of size, as follows. For each morphological measure, we fit a general linear mixed model (GLMM) with age, sex and their interaction as fixed effects, and bander and bird identities as random effects, and assuming normal residuals. We then extracted the best linear unbiased predictor (BLUP) for each bird as our measure of body size (Table S2; see e.g. Bolund *et al.* (2011) for the use of BLUPs to control for methodological variation in phenotypic measures). These models included all measurement occasions for each individual bird, although we excluded measurements greater than 3 SD from the mean for each size and age class, as those measurements were likely to be errors. We included female measurements in these models to improve the estimation of variation attributable to different banders, as well as to assess the degree of sexual size dimorphism in this population (Table S2).

Reproductive success

We assessed parentage of all sampled chicks using a panel of six microsatellite markers that give robust genotypes for parentage (Kaiser *et al.*, 2017b) and a total-evidence approach to paternity assignment (following Smith *et al.*, 2005; Kaiser *et al.*, 2015; see details in supplementary materials, including characterization of the molecular markers (Table S3), a breakdown of the number of mismatches between chicks and assigned parents for two representative years (Table S4) and a summary of the number of chicks assigned to parents in each year (Table S5)). We estimated annual genetic reproductive success as the total number of genetic offspring a male sired (i.e. day 6 chicks). We could not assess offspring production immediately following fertilization, because the high level of extra-pair paternity in our study species would make such an estimate inaccurate. Our analyses may therefore somewhat confound selection on the parents with viability selection on the young offspring (Lande & Arnold, 1983). Our parentage assignments relied on CERVUS (Kalinowski *et al.*, 2007) to evaluate genetic information. To ensure accuracy and reliability, we ran a pilot analysis using MasterBayes (Walling *et al.*, 2010), a program that allowed us to simultaneously account for genetic information and spatial information concerning the distance between putative fathers and offspring. This analysis returned highly similar results (see Supporting information and Table S6).

Sexual networks

To identify the competitive group for each male, we constructed a weighted network based on inferred copulation patterns (as inferred from paternity and social pairing data) and the spatial arrangement of male territories (Fig. 1). We used network weights, which represent the importance of competitive interactions, to calculate the weighted average body size of the competitors. To identify the male network based on copulations, we began with a network for which males were connected to females that they sired genetic offspring, with edge weight equalling the number of nests in which they produced genetic offspring together (regardless of whether they were socially paired or not). Under the assumption that social pairs have a stronger social connection than extra-pair mates, and more specifically, that social mates copulate at least once even if they did not produce genetic offspring, we added one to the edge weight of social mates (or assigned social mates an edge with weight one, if they sired no genetic offspring). These weights therefore reflect the minimum expected number of copulations between each male and female. We based edge weights on nests, rather than offspring produced together, because females might store sperm from a single copulation to fertilize multiple offspring within a brood. However, sperm storage between successive nests is unlikely, as females' sperm storage organs rapidly regress upon clutch completion (Briskie, 1996; Birkhead *et al.*, 1997). We determined the weight of male–male connections by projecting across the females in the network (Opsahl, 2009) to create edges between males that had edges to a common female. The weight for the edge connecting these males was the sum of the edges that had linked the two males to the female (Fig. 1a), and females were then removed from the projected network. To this projected network, we added new edges, or increased edge weight of existing edges, based on the spatial proximity of territories. Most extra-pair paternity happens among territories with centroids within 210 m of each other, with closer territorial neighbours likely representing more important competitors and being more likely to sire extra-pair offspring (Webster *et al.*, 2001; Kaiser *et al.*, 2017a). We therefore assigned an edge weight of 0.5 when territory centres were separated by less than 140 m (an approximation of immediate neighbours) and a weight of 0.25 if the centres were between 140 and 210 m apart (Fig. 1b; an approximation of second-degree neighbours). The final edge weight connecting males was the sum of the edge weight based on spatial proximity and the edge weight based on inferred copulations and is intended to represent the importance of interactions between each pair of males (Fig. 1c).

We chose this network weighting scheme *a priori* so that spatial proximity would have less influence than copulations on the final weight, because inferred copulations represent more certain interactions. The resulting mean number of competitors (network degree) was 5.32 ± 2.35 (mean ± SD), with a mean total weight (network strength) of 5.21 ± 4.48. Spatial edges in the network accounted for approximately 31% of total weight (1.58 ± 0.78), whereas sexual edges accounted for approximately 69% (3.62 ± 4.40) of total weight. We investigated an analysis using only the copulation-based edges to assess whether focusing only on social interactions produced different results but found no evidence of different results beyond what would be expected due to sampling variation (see supplementary materials). We did not explore other weighting options, to avoid inflating the likelihood of making type I errors.

Based on this network that accounts for both spatial and known social interactions, the competitive group for each male was defined as the subset of other males to whom he was directly connected, that is his one-step ego network (McDonald *et al.*, 2013). We calculated the

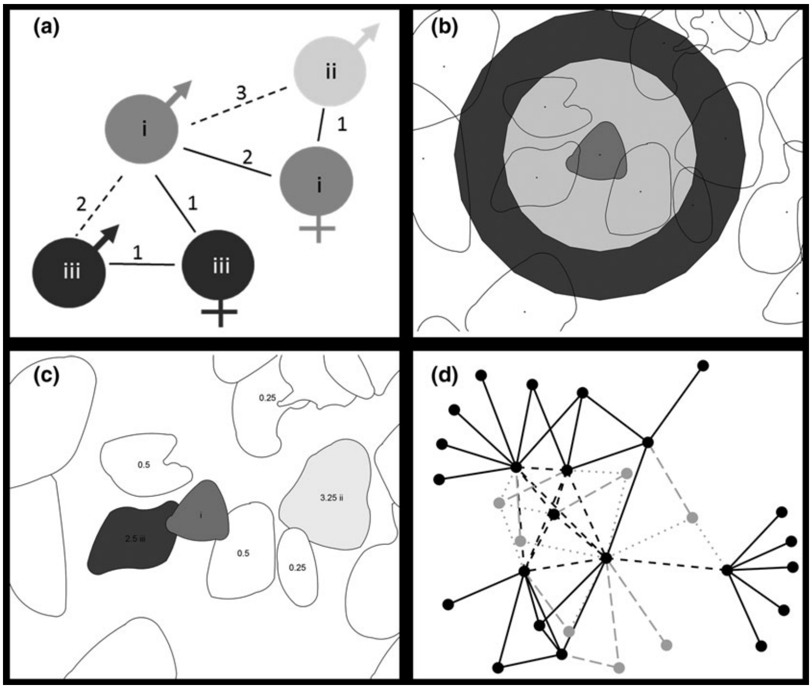

Fig. 1 Schematic of how networks were constructed (a–c) and sample network subset (d). Social pairing and paternity data (a) and spatial relationships (b) were used to obtain final edge weights (c), shown here for a hypothetical focal male in medium grey. (a) Males (♂) and females (♀) were connected by network edges (solid lines) whose weights were equal to the number of nests in which they produced genetic offspring together, plus one if the individuals were social mates. We projected across females to create a network that included only males (dotted lines), with edge weights equal to the sum of the edge weights of each male to the common female. Example edge weights are indicated as numbers above lines. In this example, social mates i (medium grey) produced offspring together in one nest, where male ii (light grey) also sired some offspring. Social mates iii (dark grey) had one nest together, but all offspring in it were sired by male i. (b) Males with a territory centre (small +) falling within 140 m of the focal male's centre (light grey circle) were assigned a weight of 0.5, and males with a centre between 140 and 210 m (dark grey circle) were assigned a weight of 0.25. (c) Total weighting was the sum of the two types of edge weights, here with labels and colours indicating male identity from panels a and b. (d) A partial network (truncated for ease of presentation) of black-throated blue warblers. Males and females are black and grey circles, respectively. Solid black lines indicate direct spatial connections between males. Grey lines indicate male–female connections (dotted, inferred extra-pair copulations; dashed, within-pair copulations/social connections), and dashed black lines indicate projected male connections via copulation with the same female. Note that many extra-pair copulations occur on a local spatial scale, not indicated directly in this figure panel.

weighted mean body size of the males in the competitive group using the edge weight as the weighting factor (McDonald *et al.*, 2013). We used weighted estimates in the presented models; using a flat weighting scheme (treating all edges as equal) did not qualitatively alter results (not shown), nor did using weightings based only on inferred copulations and excluding spatial information (see supplementary materials). Note that weights were used only to calculate weighted average phenotypes of competitors and were not directly included in the statistical models. We further explored a randomization approach to deal with nonindependence of network statistics (Croft *et al.*, 2008; Farine & Whitehead, 2015) and with the possible covariance between reproductive success and network weight; this approach gave similar results to the main analysis (see supplementary materials).

This protocol modified the procedures of McDonald *et al.* (2013), who constructed two separate sexual networks, one based on female visitation of males, and a second based on copulation patterns. However, as is true of most field studies, we were unable to directly observe female visitation of males or copulations with any degree of detail. We therefore modified the methods of McDonald *et al.* (2013) to account for the social environment using territorial social interactions and patterns of within-pair and extra-pair paternity (Webster *et al.*, 2001; Kaiser *et al.*, 2015, 2017a). Territorial neighbours that did not sire extra-pair offspring likely represent the pool of males that a female chose not to copulate with or that females copulated with but whose nest failed before sampling.

Because the sexual network approach was proposed to describe situations where males are more similar to the other members of their competitive groups than expected by chance, we measured the weighted assortativity in each network, following the procedures of Farine (2014). This analysis was conducted on the sexual network for each plot and year, not on the one-step ego networks. The weighted assortativity coefficient is derived from the Pearson correlation coefficient, with scores ranging from −1 to 1. Weighted assortativity combines information on the edge weight connecting each pair of males in the population with information on how similar those two males are in phenotype. Negative scores indicate that each male tends to be strongly, directly connected to males that differ from him in body size, while positive scores indicate that each male tends to be strongly, directly connected to individuals that are similar to him in body size. A value of zero indicates that network connections are random with respect to male phenotype. A single weighted assortativity score was calculated for each sexual network.

Selection analysis

We calculated linear selection gradients (Lande & Arnold, 1983) using annual measures of genetic reproductive success, territory quality and the social environment. Performing selection analysis on annual data, rather than summing across all years a male bred, appears likely to be more powerful, because the social environment and territory quality differ among years. Moreover, this approach improved our ability to account for age effects on reproductive success for those birds that were observed both in their first breeding season and in later breeding seasons. We used the same value for morphological measurements for all years that the bird was observed breeding (see Adult morphological measurements). To make our results comparable with previous studies and to facilitate the relation of selection gradients to evolutionary theory (Lande & Arnold, 1983), we standardized variables before analysis. We standardized morphological measurements to have a mean of 0 and a variance of 1 for each year and plot separately. The weighted mean body size of the competitors was also standardized in this way (after weighted means were calculated using unstandardized BLUP data). We standardized annual genetic reproductive success and territory quality by dividing by their respective mean values. Standardizing within plots is consistent with other studies (e.g. Husby *et al.*, 2011) and improves our ability to compare patterns across plots. Moreover, sexual networks were within-year and within-plot due to the limited geographic scale of extra-pair paternity behaviour (Webster *et al.*, 2001; Kaiser *et al.*, 2017a). We conservatively only included individuals that were not part of experimental manipulations in the selection analysis (see supplementary methods for details). Experimentally manipulated males were included in estimating the body size of competitors, because the experimental treatments likely reduced their social connections (or our ability to detect them), such that removing manipulated males from networks would be anticonservative.

For each dependent variable (territory quality and reproductive success) and each plot, we constructed a separate GLMM (Bates *et al.*, 2014) for each morphological measurement. These models included year and male identity as random effects, as well as fixed effects of age, the male's own morphological measurement, and the weighted mean morphological measurement of the male's competitors. We allowed for nonlinear selection by testing a quadratic term for the birds' own measurement. To test the prediction that the strength of direct and social selection would depend on the relative size of the individual and his competitors, we tested the interaction between the bird's own size and the size of his competitors in predicting fitness. We assessed significance using lmerTest (Kuznetsova *et al.*, 2014) and corrected for multiple testing using false discovery rate correction (Verhoeven *et al.*, 2005). We further calculated 95% confidence intervals on selection gradient parameters by bootstrapping, using 1000 iterations for each model. We applied a square root transformation to standardized territory quality to improve normality of model residuals and to ensure that statistical significance estimates were meaningful. However, to provide comparability across studies, we report the estimated strength of selection from the standardized, untransformed measure (Lande & Arnold, 1983). Model residuals describing reproductive success approached normality and were not improved by transformations. All statistics were performed in R v. 3.3.0 (R Development Core Team, Vienna, Austria).

Results

In total, 700 individual males were included in the sexual networks, with repeated observations bringing the total number of observations to 945 male-years. These males sired 2250 offspring, with paternity confirmed by genetic analysis. In total, 10 nonexperimental males (in one year only for all 10 males) had no identified competitors because they had no territorial neighbours within 210 m, were not cuckolded by identified extra-pair sires and did not sire extra-pair chicks that we detected. These isolated males had smaller wings (difference ± SE in standardized wing, −0.35 ± 0.16, $t_{253} = 2.23$, $P = 0.03$) and tended to have shorter tarsi (−0.32 ± 0.17, $t_{297} = 1.87$, $P = 0.06$) than connected males, though tail length did not differ (0.04 ± 0.18, $t_{292} = 0.25$, $P = 0.81$). After removing unconnected and experimentally manipulated males, selection analysis included 719–735 observations of 561–573 males across all three plots (with sample sizes differing slightly among morphological measurements; Table S7). Assortativity scores for male phenotypes generally were weak, ranging from −0.45 to 0.38 (mean ± *SD*:

−0.05 ± 0.13), and we found no significant variation in assortativity depending on plot, year or specific morphological measurement (all F-statistics < 1.00, P > 0.40). As assortativity became more extreme, population-wide variance in the mean phenotype of males' competitors increased (quadratic term: 2.94 ± 0.69, $F_{1,87}$ = 18.27, P < 0.001; linear term: 0.86 ± 0.18, $F_{1,87}$ = 11.36, P = 0.001). Covariance between the male's own phenotype and the phenotype of his competitors was low (wing: −0.09; tarsus: −0.09; tail: −0.13).

Selection gradients

Selection gradients on body size were weak, with confidence intervals typically encompassing 0 (Tables 1, 2). While some relationships between body size and territory quality (Table S8) or between body size and genetic reproductive success (Table S9) approached significance, none of these results were significant after correcting for multiple testing. Similarly, nonlinear selection on birds' own body size and interactions between the birds' own body size and the phenotype of the competitors were not significant after correction for multiple testing (parameters shown in Tables S8 and S9). Selection gradient estimates varied substantially among plots, but differences among plots were not consistent across morphological measurements (Tables 1, 2). Counter to our prediction, we did not find that social selection acted in opposition to direct selection, as the selection gradients on the birds' own body size and the mean body size of the competitive group were not consistently in opposite directions (Tables 1, 2). The strength of direct selection on body size across both measures of fitness and all plots was 0.05 ± 0.07 (mean ± SD). Older males defended higher-quality territories on one of the study plots (Table S8), consistent

Table 1 Standardized selection gradients relating morphological variables to territory quality in three study plots. We report standardized linear gradients (β) on the bird's own morphological measurements and on the morphological measurements of birds' competitors. Values presented here are from models including only linear effects of the birds' own and the competitors' trait measurement, and the covariate of age. See Table S8 for test statistics, quadratic selection gradients (γ), interactions between the birds' own and competitors' measurements and the effects of the covariate age. No gradients were statistically significant after correction for multiple testing.

		Selection gradient ± SE* (95% Confidence interval)		
Trait	Gradient type	Low elevation	Mid elevation	High elevation
Tail	Linear, own	0.12 ± 0.06 (0.00, 0.24)	0.04 ± 0.03 (−0.02, 0.10)	0.02 ± 0.07 (−0.11, 0.16)
	Linear, competitors'	−0.03 ± 0.06 (−0.15, 0.08)	0.00 ± 0.03 (−0.05, 0.06)	−0.01 ± 0.05 (−0.12, 0.10)
Tarsus	Linear, own	0.00 ± 0.07 (−0.14, 0.13)	−0.02 ± 0.03 (−0.09, 0.04)	−0.08 ± 0.06 (−0.21, 0.05)
	Linear, competitors'	0.03 ± 0.07 (−0.11, 0.17)	−0.00 ± 0.03 (−0.06, 0.05)	−0.09 ± 0.05 (−0.19, 0.02)
Wing	Linear, own	0.18 ± 0.06 (0.05, 0.31)	0.04 ± 0.03 (−0.02, 0.10)	0.13 ± 0.06 (0.01, 0.25)
	Linear, competitors'	0.02 ± 0.06 (−0.11, 0.14)	−0.04 ± 0.03 (−0.10, 0.01)	−0.08 ± 0.05 (−0.18, 0.03)

*Sample sizes varied slightly among tests because of different missing data for the different morphological variables. Full details are in Table S3. The lowest sample sizes were as follows: low elevation plot, 136 observations of 120 males; mid elevation plot, 382 observations of 292 males; and 197 observations of 145 males.

Table 2 Standardized selection gradients relating three morphological variables to number of offspring sired in three study plots. We report standardized linear gradients (β) on the bird's own morphological measurements and on the morphological measurements of birds' competitors. Values presented here are from models including only linear effects of the birds' own and the competitors' trait measurement, and the covariate age. See Table S9 for test statistics, quadratic selection gradients (γ), interactions between the birds' own and competitors' measurements, and the effect of the covariates age. No gradients were statistically significant after correction for multiple testing.

		Selection Gradient ± SE*		
Trait	Gradient type	Low elevation	Mid elevation	High elevation
Tail	Linear, own	0.18 ± 0.10, (−0.02, 0.38)	0.06 ± 0.05, (−0.03, 0.15)	0.15 ± 0.09, (−0.01, 0.33)
	Linear, competitors'	0.22 ± 0.10, (−0.01, 0.42)	−0.04 ± 0.05, (−0.14, 0.04)	0.05 ± 0.07, (−0.09, 0.19)
Tarsus	Linear, own	−0.05 ± 0.12, (−0.27, 0.16)	0.07 ± 0.05, (−0.03, 0.16)	0.03 ± 0.08, (−0.13, 0.18)
	Linear, competitors'	0.10 ± 0.12, (−0.15, 0.33)	0.01 ± 0.05, (−0.08, 0.09)	−0.10 ± 0.07, (−0.24, 0.06)
Wing	Linear, own	0.07 ± 0.11, (−0.15, 0.28)	−0.01 ± 0.05, (−0.10, 0.09)	0.16 ± 0.08, (0.01, 0.32)
	Linear, competitors'	0.14 ± 0.11, (−0.07, 0.36)	−0.07 ± 0.04, (−0.16, 0.02)	−0.07 ± 0.07, (−0.22, 0.07)

*Sample sizes varied slightly among tests because of different missing data for the different morphological variables. Full details are in Table S3. The lowest sample sizes were as follows: low elevation plot, 136 observations of 120 males; mid elevation plot, 382 observations of 292 males; and 197 observations of 145 males.

with previous studies. Older males typically produced more offspring (Table S9).

Discussion

Body size is a determinant of competitive ability in a diversity of taxa (Arnott & Elwood, 2009) and is a trait preferred by the choosy sex in several species (Basolo, 2004; Callander *et al.*, 2011). We therefore predicted that a bird's own body size should influence fitness via the effect of body size on competitive ability and that the strength of selection on body size would be mediated by the body size of competitors. However, we found little evidence that either male body size or the social environment affected individual fitness, despite using a network-based measurement of the social environment and a data set that encompassed 10 years and three study plots with diverse social and abiotic conditions. Because we simultaneously tested direct and social selection, we can be confident that body size has little effect on fitness during the breeding season. This study provides new insights into applying a network-based approach to evaluate social selection, and it provides new information about selection dynamics in a well-studied passerine species.

The impact of the social environment

We used a social network approach to test the effect of competitors' phenotypes on the reproductive success of focal males (McDonald *et al.*, 2013). The approach would ideally model male–female encounters, copulations and fertilizations as successive networks, with data on male success at each step. As is typical of field studies, we lacked such behavioural data. We instead used proxies of the relevant behaviours, based on the known spatial pattern of territories and the occurrence of extra-pair parentage in birds' nests. This should be a reasonable approximation, as territorial neighbours are certain to interact with each other, neighbouring males are the most common extra-pair sires and non-territorial 'floaters' are absent from the population (Marra & Holmes, 1997; Webster *et al.*, 2001; Sillett *et al.*, 2004; Kaiser *et al.*, 2017a). Given this knowledge of the species' biology, our designation of the competitive pool most likely represents the males that the focal individual's social mate would have evaluated as copulation partners. The lack of an association between the mean body size of a male's competitors and his own reproductive success therefore should not reflect problems with identifying the primary competitors. However, two additional issues we encountered with the social network approach merit discussion.

First, some individuals had to be excluded from analysis because they were not connected to other individuals in the network and therefore lacked data on the phenotype of their competitors. Isolated males were a nonrandom subset with respect to reproductive success because they either were not cuckolded or were not cuckolded by known males (though they may have been cuckolded by unidentified, off-plot, neighbours), and they did not gain detected extra-pair offspring. When isolated males also represent a nonrandom subset with respect to the phenotype of interest, excluding them from selection analysis may bias results. In our data set, only a small number of males were isolated, so they probably had a minimal effect on our analysis. This issue could be a more substantial problem for studies with a larger proportion of isolated individuals.

Second, statistical power to detect social selection may be limited because the variable describing the phenotype of the competitors is an average across several individuals and therefore has low variance compared to individual phenotypes. Note that we refer not to variance in the phenotype among each male's competitors, but rather to variation in the mean competitor phenotypes across all males in the population. Population-wide variance in the mean phenotype of males' competitors was highest in the years and plots with strongly assortative or disassortative sexual networks. Due to the nature of assortativity calculations, similar patterns are likely to be generally true for other data sets. That is, population-wide variance in the mean phenotype of the competitors should be highest when each individual interacts with a phenotypically homogenous subset of the population. When individuals interact with phenotypically homogenous subsets that are similar to themselves, the network will be highly assortative. In contrast, a network will be disassortative when individuals interact with subsets of the population that differ substantially from themselves. Maximum disassortativity will occur when those competitors are similar in phenotype to each other, but do not interact directly with each other. Statistical power to detect social selection may therefore be highest with strongly assortative or disassortative networks.

Strongly assortative networks may also represent the situations in which control for social selection is most statistically important: in these cases, social selection might be most likely to counteract direct selection. For example, a male gains fitness by being large himself, but loses fitness by interacting with other large males that are effective competitors, resulting in little apparent relationship between size and fitness in an analysis that fails to separate social and direct selection. However, the relatively small scale of differences between competitors in such a highly assortative network may reduce the importance of differences in phenotype in determining the outcome of competition. In intrasexual encounters, when the difference in inherent competitive ability is small, the duration and presumed costs of physical contests can increase substantially (Enquist & Leimar, 1983; Arnott & Elwood, 2009). Factors such as asymmetries in the value of the resource being contested can affect the outcome and may override the effects of different inherent competitive abilities (Renison *et al.*, 2000; Lindström & Pampoulie, 2005; Gherardi, 2006). Likewise, in intersexual selection, the strength of female preferences for one male over another may correlate with the relative difference between male phenotypes (Gerhardt *et al.*, 2000; Basolo, 2004), such that females might not exert strong choice, or perhaps not perceive differences, between two males with very similar phenotypes.

Strongly disassortative networks may represent the contexts where phenotypic differences among interacting individuals are the most biologically relevant to social selection: differences in phenotype between competitors would be substantial and statistically tractable because of high statistical power, due to the substantial variation in the mean phenotype of the competitors. Supporting this idea, Formica *et al.* (2011) found stronger directional selection in their study populations where focal individual body size was negatively correlated with the body size of the competitors (similar to a disassortative network). In such disassortative cases, direct and social selection may reinforce each other's effects, resulting in particularly strong total selection. This type of concern may affect most studies that control for the social environment, regardless of whether it is defined using network or other approaches.

Direct selection on body size

Direct observations of competitive or dominance interactions between individuals across a diversity of species, including some passerine birds, suggest that larger individuals are more likely to win such encounters (Alatalo & Moreno, 1987; Mönkkönen, 1990; Sandell & Smith, 1991; French & Smith, 2005; Jonart *et al.*, 2007; Arnott & Elwood, 2009). However, we found no significant effect of body size measures on reproductive success or territory quality in black-throated blue warblers. Selection on body size in small-bodied passerines may be relatively weak compared to selection on body size in other taxa: although some field studies on small- to medium-sized passerines find selection for larger body size (McGlothlin *et al.*, 2005; Dolan *et al.*, 2007; Jensen *et al.*, 2008; Husby *et al.*, 2011; Ryder *et al.*, 2012; Lessard *et al.*, 2014), many find no relationship between body size and reproductive success (Yezerinac & Weatherhead, 1997; Johnsen *et al.*, 2001; Westneat, 2006; Balenger *et al.*, 2009; Husby *et al.*, 2011; Cramer, 2013; Sousa & Westneat, 2013), and some find selection for smaller body size (Woolfenden *et al.*, 2002). Moreover, a recent meta-analysis in passerine birds shows that body size generally does not differ between social males and the extra-pair males that cuckold them (Hsu *et al.*, 2015). Extra-pair paternity, one of several potential agents of sexual selection in passerine birds, may not therefore generally promote larger body size in passerines, even though it does select for larger size in some species (e.g. Ryder *et al.*, 2012). Black-throated blue warblers are mildly sexually size-dimorphic, suggesting that either the weak selection we observed is sufficient to maintain sexual size dimorphism or that selection on male body size may occur during other episodes of selection, such as survival during migration or on nonbreeding territories.

The mean strength of direct selection on individual body size in our study was only 0.05, compared to a median strength of linear selection on body size of 0.08 (Kingsolver *et al.*, 2012). Our sample sizes were large and comparable to other studies in wild vertebrates that did find significant selection (e.g. McGlothlin *et al.*, 2005). However, for studies with large sample sizes, the estimated strength of selection typically falls from about 0.16 to below 0.1 (Kingsolver *et al.*, 2001), perhaps partly because those studies also encompass greater environmental heterogeneity, which can reduce the apparent strength of selection (Aspi *et al.*, 2003). Our analyses accounted for environmental heterogeneity, but selection still appeared to be quite weak across all plots combined.

The opportunity for selection on black-throated blue warblers appears to be high (Webster *et al.*, 2001), but realized selection may not reach this maximum potential, perhaps in part because spatial constraints on extra-pair paternity limit the scope for sexual selection to act (Kaiser *et al.*, 2017a). Song characteristics, plumage coloration or manoeuvreability during combat may be more important than body size in competitive interactions in this species, though these ideas remain to be tested. Selection analysis that focuses on females could reveal different patterns.

We recommend that future research examine the impact of competitors' phenotypes on selective pressures experienced by individuals. When social and direct selection act in opposition and are of similar magnitude, studies that fail to test both types of selection may wrongly conclude that the trait does not significantly affect selection dynamics. Our results suggest that social selection may be most detectable when each individual competes against a relatively homogenous subset of the population and where competitors differ substantially in phenotype from the focal individual. Given the likely importance of competition with conspecifics in many taxa, future studies that account for variation in the social environment are needed to advance our understanding of selection dynamics.

Acknowledgments

Funding was provided by the Smithsonian Migratory Bird Center; the Cornell Lab of Ornithology; US National Science Foundation (NSF) grants to Cornell (0640470), the Smithsonian Institution (0640732, 1353085) and Wellesley College (064082300); and the

Environmental Protection Agency to SAK (91723201). Dick Holmes, Nick Rodenhouse, Mike Hallworth, Nicole Kraus, the Fuller Evolutionary Biology Lab and many teams of field technicians provided invaluable assistance in the field and with analysis. Field work was conducted at the Hubbard Brook Ecosystem study, which is part of the Long-Term Ecological Research Network supported by the NSF.

References

Alatalo, R.V. & Moreno, J. 1987. Body size, interspecific interactions, and use of foraging sites in tits (Paridae). *Ecology* **68**: 1773–1777.

Arnott, G. & Elwood, R.W. 2009. Assessment of fighting ability in animal contests. *Anim. Behav.* **77**: 991–1004. Elsevier Ltd.

Aspi, J., Jäkäläniemi, A., Tuomi, J. & Siikamäki, P. 2003. Multilevel phenotypic selection on morphological characters in a metapopulation of *Silene tatarica*. *Evolution* **57**: 509–517.

Balenger, S.L., Johnson, L.S. & Masters, B.S. 2009. Sexual selection in a socially monogamous bird: male color predicts paternity success in the mountain bluebird, *Sialia currucoides*. *Behav. Ecol. Sociobiol.* **63**: 403–411.

Basolo, A.L. 2004. Variation between and within the sexes in body size preferences. *Anim. Behav.* **68**: 75–82.

Bates, D., Maechler, M., Bolker, B. & Walker, S. 2014. Lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-5.

Birkhead, T.R., Briskie, J.V., Lifjeld, J.T. & Slagsvold, T. 1997. Breeding-cycle patterns of sperm storage in the pied flycatcher (*Ficedula hypoleuca*). *Auk* **114**: 792–796.

Bolund, E., Schielzeth, H. & Forstmeier, W. 2011. Correlates of male fitness in captive zebra finches – a comparison of methods to disentangle genetic and environmental effects. *BMC Evol. Biol.* **11**: 327. BioMed Central Ltd.

Briskie, J.V. 1996. Spatiotemporal of sperm storage and last-male patterns in birds sperm precedence. *Funct. Ecol.* **10**: 375–383.

Callander, S., Jennions, M.D. & Backwell, P.R.Y. 2011. Female choice over short and long distances: neighbour effects. *Behav. Ecol. Sociobiol.* **65**: 2071–2078.

Callander, S., Hayes, C.L., Jennions, M.D. & Backwell, P.R.Y. 2012. Experimental evidence that immediate neighbors affect male attractiveness. *Behav. Ecol.* **24**: 730–733.

Chuang-Dobbs, H.C., Webster, M.S. & Holmes, R.T. 2001. The effectiveness of mate guarding by male black-throated blue warblers. *Behav. Ecol.* **12**: 541–546.

Cline, M.H., Strong, A.M., Sillett, T.S., Rodenhouse, N.L. & Holmes, R.T. 2013. Correlates and consequences of breeding dispersal in a migratory songbird. *Auk* **130**: 742–752.

Cramer, E.R.A. 2013. Physically challenging song traits, male quality, and reproductive success in house wrens. *PLoS ONE* **8**: e59208.

Croft, D.P., James, R. & Krause, J. 2008. *Exploring Animal Social Networks*. Princeton University Press, Princeton, NJ.

Dolan, A.C., Murphy, M.T., Redmond, L.J., Sexton, K. & Duffield, D. 2007. Extrapair paternity and the opportunity for sexual selection in a socially monogamous passerine. *Behav. Ecol.* **18**: 985–993.

Enquist, M. & Leimar, O. 1983. Evolution of fighting behavior: decision rules and assessment of relative strength. *J. Theor. Biol.* **102**: 387–410.

Farine, D.R. 2014. Measuring phenotypic assortment in animal social networks: weighted associations are more robust than binary edges. *Anim. Behav.* **89**: 141–153. Elsevier Ltd.

Farine, D.R. & Whitehead, H. 2015. Constructing, conducting and interpreting animal social network analysis. *J. Anim. Ecol.* **84**: 1144–1163.

Formica, V.A., McGlothlin, J.W., Wood, C.W., Augat, M.E., Butterfield, R.E., Barnard, M.E. *et al.* 2011. Phenotypic assortment mediates the effect of social selection in a wild beetle population. *Evolution* **65**: 2771–2781.

French, A.R. & Smith, T.B. 2005. Importance of body size in determining dominance hierarchies among diverse tropical frugivores. *Biotropica* **37**: 96–101.

Gerhardt, H.C., Tanner, S.D., Corrigan, C.M. & Walton, H.C. 2000. Female preference functions based on call duration in the gray tree frog (*Hyla versicolor*). *Behav. Ecol.* **11**: 663–669.

Gherardi, F. 2006. Fighting behavior in hermit crabs: the combined effect of resource-holding potential and resource value in *Pagurus longicarpus*. *Behav. Ecol. Sociobiol.* **59**: 500–510.

Goodnight, C.J. 2015. Multilevel selection theory and evidence: a critique of Gardner, 2015. *J. Evol. Biol.* **28**: 1734–1746.

Heisler, I.L. & Damuth, J. 1987. A method for analyzing selection in hierarchically structured populations. *Am. Nat.* **130**: 582–602.

Holmes, R.T. 2011. Avian population and community processes in forest ecosystems: long-term research in the Hubbard Brook Experimental Forest. *For. Ecol. Manage.* **262**: 20–32.

Holmes, R.T., Rodenhouse, N.L. & Sillett, T.S. 2005. Black-throated Blue Warbler (*Dendroica caerulescens*). In: *The Birds of North America Online* (A. Poole, ed), Cornell Lab of Ornithology, Ithaca, New York.

Hsu, Y.-H., Schroeder, J., Winney, I., Burke, T. & Nakagawa, S. 2015. Are extra-pair males different from cuckolded males? A case study and a meta-analytic examination. *Mol. Ecol.* **24**: 1558–1571.

Husby, A., Hille, S.M. & Visser, M.E. 2011. Testing mechanisms of Bergmann's rule: phenotypic decline but no genetic change in body size in three passerine bird populations. *Am. Nat.* **178**: 202–213.

Jensen, H., Steinsland, I., Ringsby, T.H. & Sæther, B.E. 2008. Evolutionary dynamics of a sexual ornament in the house sparrow (*Passer domesticus*): the role of indirect selection within and between sexes. *Evolution* **62**: 1275–1293.

Johnsen, A., Lifjeld, J., Andersson, S., Örnborg, J. & Amundsen, T. 2001. Male characteristics and fertilisation success in bluethroats. *Behaviour* **138**: 1371–1390.

Jonart, L.M., Hill, G.E. & Badyaev, A.V. 2007. Fighting ability and motivation: determinants of dominance and contest strategies in females of a passerine bird. *Anim. Behav.* **74**: 1675–1681.

Kaiser, S.A., Sillett, T.S. & Webster, M.S. 2014. Phenotypic plasticity in hormonal and behavioural responses to changes in resource conditions in a migratory songbird. *Anim. Behav.* **96**: 19–29.

Kaiser, S.A., Sillett, T.S., Risk, B.B. & Webster, M.S. 2015. Experimental food supplementation reveals habitat-dependent male reproductive investment in a migratory bird. *Proc. R. Soc. B* **282**: 20142523–20142523.

Kaiser, S.A., Risk, B.B., Sillett, T.S. & Webster, M.S. 2017a. Ecological and social factors constrain spatial and temporal opportunities for mating in a migratory songbird. *Am. Nat.* **189**: 283–296.

Kaiser, S.A., Taylor, S.A., Chen, N. & Sillett, T.S. 2017b. A comparative assessment of SNP and microsatellite markers for assigning parentage in a socially monogamous bird. *Mol. Ecol. Resour.* **17**: 183–193.

Kalinowski, S.T., Taper, M.L. & Marshall, T.C. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* **16**: 1099–1106.

Kelly, C.D. 2008. The interrelationships between resource-holding potential, resource-value and reproductive success in territorial males: how much variation can we explain? *Behav. Ecol. Sociobiol.* **62**: 855–871.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E. *et al.* 2001. The strength of phenotypic selection in natural populations. *Am. Nat.* **157**: 245–261.

Kingsolver, J.G., Diamond, S.E., Siepielski, A.M. & Carlson, S.M. 2012. Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. *Evol. Ecol.* **26**: 1101–1118.

Kuznetsova, A., Brockhoff, P.B. & Christensen, R.H.B. 2014. lmerTest: Tests for random and fixed effects for linear mixed effects models (lmer objects of lme4 package). R package version 2.0-11.

Lande, R. & Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution* **37**: 1210–1226.

Lessard, A., Bourret, A., Bélisle, M., Pelletier, F. & Garant, D. 2014. Individual and environmental determinants of reproductive success in male tree swallow (*Tachycineta bicolor*). *Behav. Ecol. Sociobiol.* **68**: 733–742.

Lindström, K. & Pampoulie, C. 2005. Effects of resource holding potential and resource value on tenure at nest sites in sand gobies. *Behav. Ecol.* **16**: 70–74.

Marra 1993. Reverse mounting in the black-throated blue warbler. *Wilson Bull.* **105**: 359–361.

Marra, P.P. & Holmes, R.T. 1997. Avian removal experiments: do they test for habitat saturation or female availability? *Ecology* **78**: 947–952.

McDonald, G.C., James, R., Krause, J. & Pizzari, T. 2013. Sexual networks: measuring sexual selection in structured, polyandrous populations. *Philos. Trans. R. Soc. London* **368**: 20120356.

McGlothlin, J.W., Parker, P.G., Nolan, V. & Ketterson, E.D. 2005. Correlational selection leads to genetic integration of body size and an attractive plumage trait in dark-eyed juncos. *Evolution* **59**: 658–671.

McGlothlin, J.W., Moore, A.J., Wolf, J.B. & Brodie, E.D. 2010. Interacting phenotypes and the evolutionary process. III. Social evolution. *Evolution* **64**: 2558–2574.

Mönkkönen, M. 1990. Removal of territory holders causes influx of small-sized intruders in passerine bird communities in northern Finland. *Oikos* **57**: 281–288.

Moore, A.J., Brodie, E.D.I. & Wolf, J.B. 1997. Interacting phenotypes and the evolutionary process: I. Direct and indirect genetic effects of social interactions. *Evolution* **51**: 1352–1362.

Morrissey, M.B. & Hadfield, J.D. 2012. Directional selection in temporally replicated studies is remarkably consistent. *Evolution* **66**: 435–442.

Nakagawa, S. & Schielzeth, H. 2010. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biol. Rev. Camb. Philos. Soc.* **85**: 935–956.

Opsahl, T. 2009. Structure and evolution of weighted networks. University of London (Queen Mary College), London.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Renison, D., Boersma, D. & Martella, M.B. 2000. Winning and losing: causes for variability in outcome of fights in male Magellanic penguins (*Spheniscus magellanicus*). *Behav. Ecol.* **13**: 462–466.

Rodenhouse, N.L. & Holmes, R.T. 1992. Results of experimental and natural food reductions for breeding black-throated blue warblers. *Ecology* **73**: 357–372.

Rodenhouse, N.L., Sillett, T.S., Doran, P.J. & Holmes, R.T. 2003. Multiple density-dependence mechanisms regulate a migratory bird population during the breeding season. *Proc. R. Soc. B* **270**: 2105–2110.

Ryder, T.B., Fleischer, R.C., Shriver, W.G. & Marra, P.P. 2012. The ecological-evolutionary interplay: density-dependent sexual selection in a migratory songbird. *Ecol. Evol.* **2**: 976–987.

Sandell, M. & Smith, H.G. 1991. Dominance, prior occupancy, and winter residency in the great tit (*Parus major*). *Behav. Ecol. Sociobiol.* **29**: 147–152.

Siepielski, A.M., DiBattista, J.D., Evans, J.A. & Carlson, S.M. 2011. Differences in the temporal dynamics of phenotypic selection among fitness components in the wild. *Proc. R. Soc. B* **278**: 1572–1580.

Sillett, T.S., Rodenhouse, N.L. & Holmes, R.T. 2004. Experimentally reducing neighbor density affects reproduction and behavior of a migratory songbird. *Ecology* **85**: 2467–2477.

Smith, S.B., Webster, M.S. & Holmes, R.T. 2005. The heterozygosity theory of extra-pair mate choice in birds: a test and a cautionary note. *J. Avian Biol.* **36**: 146–154.

Sousa, B.F. & Westneat, D.F. 2013. Variance in mating success does not produce strong sexual selection in a polygynous songbird. *Behav. Ecol.* **24**: 1381–1389.

Verhoeven, K.J.F., Simonsen, K.L. & McIntyre, L.M. 2005. Implementing false discovery rate control: increasing your power. *Oikos* **108**: 643–647.

Wagner, W.E. 1998. Measuring female mating preferences. *Anim. Behav.* **55**: 1029–1042.

Walling, C.A., Pemberton, J.M., Hadfield, J.D. & Kruuk, L.E.B. 2010. Comparing parentage inference software: reanalysis of a red deer pedigree. *Mol. Ecol.* **19**: 1914–1928.

Webster, M.S., Chuang-Dobbs, H.C. & Holmes, R.T. 2001. Microsatellite identification of extrapair sires in a socially monogamous warbler. *Behav. Ecol.* **12**: 439–446.

Westneat, D.F. 2006. No evidence of current sexual selection on sexually dimorphic traits in a bird with high variance in mating success. *Am. Nat.* **167**: E171–E189.

White, L. & Densmore, L. 1992. Mitochondrial DNA isolation. In: *Molecular Genetic Analysis of Populations: A Practical Approach* (A. Hoelzel, ed), pp. 29–58. IRL Press, New York.

Wolf, J.B., Brodie, E.D. III & Moore, A.J. 1999. Interacting phenotypes and the evolutionary process. II. Selection resulting from social interactions. *Am. Nat.* **153**: 254–266.

Woolfenden, B.E., Lisle Gibbs, H. & Sealy, S.G. 2002. High opportunity for sexual selection in both sexes of an obligate brood parasitic bird, the brown-headed cowbird (*Molothrus ater*). *Behav. Ecol. Sociobiol.* **52**: 417–425.

Yezerinac, S.M. & Weatherhead, P.J. 1997. Extra-pair mating, male plumage coloration and sexual selection in yellow warblers (*Dendroica petechia*). *Proc. R. Soc. B* **264**: 527–532.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Table S1 Summary of differences among the three study plots, using data from all non-experimental territories, 2006-2015.

Table S2 Estimated mean ± SE measurements for each age and sex category of black-throated blue warbler, from GLMMs including age, sex, and their interaction as fixed effects and bird and bander identities as random effects.

Table S3 Assessment of microsatellite loci used in parentage analysis, based on all 243 adults present in the population in 2015.

Table S4 Information on mismatches between offspring and their assigned genetic parents in 2014 and 2015 field seasons, a subset of the 10 years used in the study, to illustrate the degree of allele-sharing among assigned parents and offspring.

Table S5 Overview of the number of offspring and parents successfully sampled and identified as genetic relatives between 2006 and 2015 at the Hubbard Brook Experimental Forest.

Table S6 Comparison of the parentage assignments for 109 offspring using CERVUS to MasterBayes genetic (G) and MasterBayes phenotypic (P) models.

Table S7 Sample sizes used in selection gradients.

Table S8 Standardized selection gradients relating morphological variables to territory quality in three study plots.

Table S9 Standardized selection gradients relating morphological variables to reproductive success in three study plots.

Data deposited at Dryad: https://doi.org/10.5061/dryad.r5nj3

Received 11 October 2016; revised 19 August 2017; accepted 27 September 2017