Ⓔ

Natural Selection on Testosterone Production in a Wild Songbird Population

Joel W. McGlothlin,[1,*] Danielle J. Whittaker,[2] Sara E. Schrock,[2] Nicole M. Gerlach,[2] Jodie M. Jawor,[3] Eric A. Snajdr,[2] and Ellen D. Ketterson[2]

1. Department of Biology, University of Virginia, Charlottesville, Virginia 22904;   2. Department of Biology and Center for the Integrative Study of Animal Behavior, Indiana University, Bloomington, Indiana 47405;   3. Department of Biological Sciences, University of Southern Mississippi, Hattiesburg, Mississippi 39406

ABSTRACT: Because of their role in mediating life-history trade-offs, hormones are expected to be strongly associated with components of fitness; however, few studies have examined how natural selection acts on hormonal variation in the wild. In a songbird, the dark-eyed junco (*Junco hyemalis*), field experiments have shown that exogenous testosterone alters individuals' resolution of the survival-reproduction trade-off, enhancing reproduction at the expense of survival. Here we used standardized injections of gonadotropin-releasing hormone (GnRH) to assay variation in the testosterone production of males. Using measurements of annual survival and reproduction, we found evidence of strong natural selection acting on GnRH-induced increases in testosterone. Opposite to what would be predicted from the survival-reproduction trade-off, patterns of selection via survival and reproduction were remarkably similar. Males with GnRH-induced testosterone production levels that were slightly above the population mean were more likely to survive and also produced more offspring, leading to strong stabilizing selection. Partitioning reproduction into separate components revealed positive directional selection via within-pair siring success and stabilizing selection via extrapair mating success. Our data represent the most complete demonstration of natural selection on hormones via multiple fitness components, and they complement previous experiments to illuminate testosterone's role in the evolution of life-history trade-offs.

Keywords: natural selection, sexual selection, testosterone, life-history trade-offs, GnRH challenge.

Introduction

Hormones are often crucial for the translation of genotype to phenotype, regulating key steps in development and integrating the expression of suites of functionally important traits (Moore 1991; Finch and Rose 1995; Ketterson and Nolan 1999; Wingfield et al. 2000; Nijhout 2003;

* Corresponding author; e-mail: jmcgloth@virginia.edu.

Adkins-Regan 2005). Hormones are also expected to be intimately related to fitness, and they often mechanistically underlie life-history trade-offs among fitness components such as survival and reproduction (Ketterson and Nolan 1992; Stearns 1992; Sinervo and Svensson 1998; Zera and Harshman 2001; Ricklefs and Wikelski 2002; Adkins-Regan 2005; Breuner et al. 2008; Lessells 2008; Mills et al. 2008, 2009; Bonier et al. 2009*a*). Despite the expected relationship between hormones and fitness, we have little information about how selection acts on individual variation in hormone levels and hormone-mediated traits in the wild (Kingsolver et al. 2001; Adkins-Regan 2005). Testosterone has been linked to mating success in a number of species (Borgia and Wingfield 1991; Alatalo et al. 1996; Mills et al. 2007), and several recent studies have shown relationships between corticosterone, a glucocorticoid stress hormone, and survival or reproductive success (Brown et al. 2005; Blas et al. 2007; Bonier et al. 2007, 2009*a*, 2009*b*; Cabezas et al. 2007; Breuner et al. 2008; Angelier et al. 2009; John-Alder et al. 2009; MacDougall-Shackleton et al. 2009). However, there is very little evidence indicating how selection on hormones varies across multiple components of fitness. This information is crucial to an understanding of the role of hormones in mediating life-history trade-offs (Arnold and Wade 1984; Schluter et al. 1991).

Much of our understanding of the role of hormones in the evolution of natural populations derives from "phenotypic engineering" studies, in which hormone levels are experimentally altered in order to test for their effects on phenotype and fitness (Ketterson and Nolan 1992, 1999; Ketterson et al. 1996; Sinervo and Svensson 1998; Zera and Harshman 2001; Adkins-Regan 2005; Reed et al. 2006). Despite their power to demonstrate causality, such experiments provide only limited information about evolutionary processes because the act of hormonal manip-

ulation alters the phenotypic variation on which selection operates. For this reason, manipulative experiments should be complemented by studies that focus on the causes and consequences of individual hormonal variation (Kempenaers et al. 2008; McGlothlin and Ketterson 2008; Williams 2008; Ketterson et al. 2009). Hormonal manipulations can probe the mechanistic bases of trade-offs, but studying individual variation is necessary to reveal how these trade-offs translate into selection in the wild and to predict evolutionary change. Furthermore, manipulative studies that measure components of fitness do not generate straightforward predictions about natural selection because selection acting via fitness components that trade off with one another does not always resemble the individual trade-off function (Roff and Fairbairn 2007). Rather, the pattern of natural selection via multiple components of fitness is expected to depend on how individuals vary in allocation (i.e., the resolution of trade-offs) versus individual quality or resource acquisition (i.e., the amount of currency, such as energy, available to allocate between fitness components; van Noordwijk and de Jong 1986; Roff 2002; Roff and Fairbairn 2007). Depending on how allocation and quality vary, patterns of selection on the hormonal basis of trade-offs may range from mirroring the individual trade-off function to having no resemblance to it.

One of the reasons measuring selection on hormones and hormone-mediated traits has proved to be difficult is the ubiquity of within-individual variation (Williams 2008). For example, in most male songbirds, levels of testosterone (the most common androgen in birds) vary markedly over the breeding season, usually decreasing after an early breeding season peak (Wingfield et al. 1990; Goymann et al. 2007). Although such seasonal patterns are often of intrinsic interest, they may obscure the ability to detect variation at the individual level. Short-term variation in testosterone levels is also common. In many species, males transiently increase testosterone levels in response to social stimuli such as male competitors or potential mates (Harding 1981; Moore 1983; Wingfield 1985; Wingfield et al. 1990; Hirschenhauser et al. 2003; Hirschenhauser and Oliveira 2006; Goymann et al. 2007; Landys et al. 2007; Pinxten et al. 2007). This socially modulated elevation in testosterone (or "androgen responsiveness") is likely to be particularly important when considering the evolution of testosterone-mediated traits, including their roles in the evolution of mating systems and life histories (Wingfield et al. 1990; Hirschenhauser et al. 2003; Hirschenhauser and Oliveira 2006; Goymann et al. 2007; Landys et al. 2007; Goymann 2009). Although the functions of short-term testosterone changes have not been completely described, these changes have long been associated with territoriality (Wingfield 1985; Wingfield et

al. 1987) and response to females (Moore 1983; Pinxten et al. 2007) and have recently been associated with parental behavior as well (McGlothlin et al. 2007). Our previous work suggests that measuring short-term testosterone increases in the same individuals multiple times across the breeding season is an effective way to assess individual variation in hormone profiles and should generate suitable measurements for measuring selection (Jawor et al. 2006).

In this study, we examine the relationships between individual variation in short-term testosterone elevation and fitness components of adult males in a breeding population of dark-eyed juncos (*Junco hyemalis*). A long-term study of this population found that experimentally elevated testosterone levels decreased survival (Reed et al. 2006). Testosterone-treated males more than compensated for reduced survival by siring more offspring via extrapair fertilizations than did controls, and as a result, they had higher fitness as measured by λ, the projected relative rate of population growth (Raouf et al. 1997; Reed et al. 2006). These results strongly suggest a role for testosterone in mediating the trade-off between survival and reproduction. However, the demonstration of this trade-off does not lead directly to predictions for how selection should act on testosterone in the wild. For example, if males differ primarily in the allocation of resources to survival and reproduction, we would expect to find that testosterone decreases survival but increases reproductive success, similar to that which has been demonstrated by experimental studies. Alternatively, if males differ primarily in quality, selection may act similarly via survival and reproduction because high-quality males are able to expend more effort on mating without diminishing their survival.

To test how natural variation in testosterone is related to survival and reproduction, we measured selection acting on both circulating testosterone levels and the ability of males to produce short-term testosterone increases. Testosterone was measured at multiple points in the breeding season in order to assess a male's average testosterone production. Short-term testosterone increases were measured using standardized injections of gonadotropin-releasing hormone (GnRH challenges; Wingfield and Farner 1993; Meddle et al. 2002; Millesi et al. 2002; Moore et al. 2002; Jawor et al. 2006). In vivo, GnRH is produced by the hypothalamus and regulates testosterone production by stimulating the hypothalamo-pituitary-gonadal (HPG) axis. Our GnRH challenge protocol is designed to measure a male's ability to elevate testosterone levels (i.e., the responsiveness of a male's HPG axis), and it generates repeatable short-term testosterone increases ($r = 0.36$; Jawor et al. 2006). Levels of testosterone produced in response to exogenous GnRH are strongly correlated with those produced in response to social stimuli (male territorial intruders) in the wild ($r = 0.68$; McGlothlin et al.

2008). In addition, GnRH-induced increases are positively correlated with attractive plumage and negatively correlated with parental behavior, and absolute levels of post-GnRH-challenge testosterone are positively correlated with aggressive behavior (McGlothlin et al. 2007, 2008).

Over two breeding seasons, we measured selection acting through annual survival, which was measured by recapture in the following breeding season, and annual offspring production, which was quantified using DNA paternity analysis. These selective episodes were added to estimate total annual selection (Arnold and Wade 1984; Wade and Kalisz 1989; McGlothlin 2010). We also partitioned reproductive selection in two different ways in order to examine potential trade-offs between components of reproduction. Juncos are socially monogamous breeders, meaning that they may achieve reproductive success either within or outside of a social pair (Ketterson et al. 1997; Nolan et al. 2002). Therefore, we asked whether selection differed when acting via mating success versus number of offspring per mate and via within-pair versus extrapair reproduction.

Material and Methods

Study Species and General Methods

We studied a population of the Carolina subspecies of the dark-eyed junco (*Junco hyemalis carolinensis*) that breeds at and around Mountain Lake Biological Station in Giles County, Virginia (37°22′N, 80°32′W), during the breeding seasons of 2003 and 2004. Males in this population had last been implanted with testosterone in 2000. In March and April, male juncos establish breeding territories that they defend throughout the season (Nolan et al. 2002). Typically, a single female nests on a male's territory. Females build nests, incubate clutches (usually of four eggs), and brood nestlings alone. Both parents feed the offspring after hatching. Mating often occurs outside the pair, and experimentally elevated testosterone has been shown to increase extrapair mating success (Ketterson et al. 1997; Raouf et al. 1997; Nolan et al. 2002; Reed et al. 2006).

In April–August 2003–2004, males ($n = 90$) were captured using mist nets or Potter traps. On capture, birds were transported to a central laboratory in a holding bag. If previously uncaptured, birds were fitted with a numbered aluminum leg band and a unique combination of colored plastic leg bands for identification. We determined age (yearling or older adult [≥ 2 years]) by examining first the color of the primary wing coverts and second the iris, which are both lighter in yearlings (Nolan et al. 2002). Age in years was determined by appearances in our capture records from previous years. If a bird was first captured and banded as an older adult, it was conservatively as-

sumed to be 2 years old in that year. Mass (g) was measured using a spring balance.

GnRH Challenges and Testosterone Assays

Each time a bird was captured, a blood sample was obtained from the wing vein (initial or prechallenge sample). Sampling details are reported in the appendix in the online edition of the *American Naturalist*. Handling time was recorded as the time in minutes from capture to collection of this blood sample, averaging 48 min (range 2–217 min; Jawor et al. 2006). Previous analyses have shown that increased handling time weakly affects our measurements of testosterone (Jawor et al. 2006). However, we were able to control for handling time statistically, and analyzing a reduced data set did not affect our results (see appendix). A solution of 1.25 μg chicken GnRH-I (Sigma L0637) in 50 μL of 0.1 M phosphate-buffered saline (PBS) was injected into an individual's pectoral muscle. The bird was returned to its holding bag, and after exactly 30 min, a second blood sample was collected (postchallenge sample). After collection of this sample, the bird was released at the site of capture. Plasma was separated and frozen ($-20°C$) for later hormone analysis. Red blood cells were saved and stored in lysis buffer (Longmire et al. 1992) for paternity analysis. Males were exposed to up to four GnRH challenges over the course of a breeding season.

Our GnRH challenge method stimulates a large increase in testosterone levels after 30 min, and levels return to baseline in less than 2 h (Jawor et al. 2006). Individual male juncos show repeatable differences in the magnitude of testosterone increases above initial levels, despite significant seasonal variation (Jawor et al. 2006).

We determined testosterone concentrations using an enzyme immunoassay kit (Assay Designs 901–065; Clotfelter et al. 2004) and summarized measurements to generate a single measurement for each male in each year. Details of the testosterone assay and a summary of measurements are reported in the appendix.

Paternity Analysis

We attempted to find all nests of pairs that nested on our study site. When nests were found during egg laying, one egg was collected as a part of another study (Jawor et al. 2007). Putative mothers and fathers of each nest were identified by behavior at the nest. We used genetic paternity analysis in order to assign genetic sires to offspring. For adults, we extracted DNA from red blood cells collected during GnRH challenges on males and from blood samples from adult females that had been collected concurrently. Blood samples were obtained from nestlings 6 days after hatching and stored in lysis buffer. We genotyped

a total of 265 6-day-old nestlings from the 2 years, as well as 57 dams and 108 potential sires, using five microsatellite loci, and we assigned paternity using Cervus 3.0 (Marshall et al. 1998; Kalinowski et al. 2007). Details are reported in the appendix.

Selection Analysis

To test how selection may act differently through different components of fitness, we measured selection acting via annual survival and via offspring production. We also partitioned annual reproduction in two different ways in order to ask about trade-offs between different paths to reproductive success. First, offspring production was partitioned into number of mates and number of offspring per mate. Second, offspring production was partitioned into within-pair and extrapair offspring. Details of fitness component assignment and correlations between pairs of fitness components are reported in the appendix.

Selection gradients for initial testosterone levels and GnRH-induced testosterone increases were estimated as regression slopes (Lande and Arnold 1983; Brodie et al. 1995). We conducted selection analyses for both years combined. To correct for repeated measures of a single male across the 2 years, we used PASW Statistics 17.0 (SPSS) to fit linear mixed models, with individual as a random factor in each. To generate measurements of relative fitness, before analysis each fitness component was divided by its average (survival, 0.553; offspring, 3.45; number of mates, 1.29; offspring per mate, 2.84; within-pair offspring, 2.23; extrapair offspring, 1.15). We measured both linear selection gradients (β), which indicate positive or negative directional selection, and nonlinear (quadratic) selection gradients (γ), which arise from the curvature of the relationship between fitness and phenotype and indicate either stabilizing/disruptive selection (γ_{ii}) or correlational selection (γ_{ij}; Lande and Arnold 1983; Brodie et al. 1995). Directional selection gradients were estimated in a model that included the two hormone measurements (least squares individual means [see appendix], standardized to zero mean and unit variance) as well as a fixed effect of year and age in years as a covariate. Nonlinear terms were estimated by adding the squared terms (initial testosterone^2 and GnRH-induced testosterone increase^2) and the cross product (initial testosterone × GnRH-induced testosterone increase) to this model. The squared-term regression coefficients (and their standard errors) were doubled to generate the stabilizing/disruptive selection gradients (Brodie et al. 1995; Stinchcombe et al. 2008). Standardization of traits was performed separately for each fitness component.

Statistical significance of selection gradients was tested in ASReml 2.0 (Gilmour et al. 2006) using generalized linear mixed models with appropriate error structures (binomial for annual survival, normal for offspring per mate, and Poisson for all other fitness components). These models used absolute (not relative) fitness components as dependent variables and a structure of independent variables that was identical to the linear mixed models described above. We report conditional values for F and P (Type III sums of squares). Standard errors were calculated using regression slopes from the linear mixed models and F values from generalized linear mixed models using the formula SE $= \beta$ (or $\gamma)/(F)^{1/2}$. To visualize the form of selection, we fitted univariate cubic splines using glms version 4.0/glmsWIN 1.0 (Schluter 1988). The smoothing parameter (λ) was chosen by minimizing generalized cross-validation scores. We also added selection gradients for survival and reproduction to estimate the strength of total annual selection. Details of this calculation are reported in the appendix.

Results

We found statistically significant relationships between fitness components and the testosterone increase induced by GnRH, but there were no significant relationships between fitness components and initial testosterone levels (table 1). There was no significant directional selection on GnRH-induced testosterone increases via annual survival or total offspring production. The relationships with both fitness components were curvilinear, as evidenced by the negative quadratic selection gradients that indicate stabilizing selection (table 1). Examination of the fitness function indicated a pattern of stabilizing selection acting via both survival and reproduction (fig. 1). In both cases, the fitness optimum was shifted slightly to the right, reflecting the weakly positive directional selection gradient for each fitness component. Stated another way, males with slightly higher than average GnRH-induced testosterone increases were most successful in both survival and reproduction, but those with very high increases were less successful along both dimensions.

When total offspring production was split into number of genetic mates and offspring per mate, we found no significant selection acting via mating success but relatively strong directional selection acting via offspring per mate (table 1). The fitness function for mating success was relatively flat, with a weak stabilizing component (fig. 1). Males that produced intermediate GnRH-induced testosterone increases tended to have the highest mating success. In contrast, the fitness function for offspring per mate was strongly directional: males with higher GnRH-induced testosterone increases produced more offspring with each mate (fig. 1).

A similar pattern was observed when total offspring

production was split into within-pair and extrapair components. We found significant positive directional selection via within-pair success (table 1), and this remained significant when we controlled for apparent within-pair offspring production (the total number of 6-day-old nestlings produced by a male's social partner, including the offspring he did not sire; table 1). This effect indicates that males with higher GnRH-induced testosterone increases were more successful primarily because they sired a higher percentage of the offspring of their social mate that survived to be nestlings, and not because they were able to attract a more fecund social mate. The fitness function for within-pair offspring was mostly linear with a weak stabilizing component that was probably driven by a single male that produced a very high testosterone increase but sired no offspring (fig. 1). Controlling for apparent offspring production removed this stabilizing component (fig. 1). In contrast, selection acting via extrapair offspring production was strongly stabilizing, with no significant evidence of directional selection (table 1). The fitness function confirmed that the fitness optimum was situated near the population mean (fig. 1).

When we combined survival and reproductive selection, total annual selection on GnRH-induced testosterone increase was strongly stabilizing (table 2). Directional selection via the two fitness components was reinforcing, leading to net positive directional selection via annual fitness (table 2).

Discussion

Natural selection acting on the magnitude of GnRH-induced testosterone production in our population was primarily stabilizing, with some evidence of a directional component. Stabilizing selection was strong, with a magnitude that was well above the median quadratic selection ($|\gamma| = 0.10$) reported by Kingsolver et al. (2001). Patterns of selection via annual survival and offspring production were very similar, and these two episodes of selection reinforced each other to generate stronger total annual selection. The location of the fitness optimum did not differ substantially for survival and reproduction, suggesting that selection on natural variation in GnRH-induced testosterone increases does not mirror the within-individual trade-off between survival and reproduction that is indicated by experimental studies. Some evidence for different trade-offs was found between components of annual offspring production; directional selection favored greater GnRH-induced testosterone increases via offspring per mate and within-pair siring success, but selection was mostly stabilizing (with a nonsignificant negative directional selection component) via mating success and extrapair offspring production.

Targets of Selection

Traditionally, measurements of selection on natural variation in hormones and other physiological traits have been much less common than those on morphological traits (Kingsolver et al. 2001; Adkins-Regan 2005). A number of studies, most of them very recent, have demonstrated selection acting on natural variation in animal hormone levels (Borgia and Wingfield 1991; Alatalo et al. 1996; Brown et al. 2005; Blas et al. 2007; Bonier et al. 2007, 2009*a*, 2009*b*; Cabezas et al. 2007; Mills et al. 2007; Breuner et al. 2008; Angelier et al. 2009; John-Alder et al. 2009; MacDougall-Shackleton et al. 2009). By far, most of this work has been conducted on glucocorticoid stress hormones. Significant directional selection on baseline glucocorticoid levels has been found in a number of studies, with some studies finding a negative relationship and others finding a positive one (Bonier et al. 2009*b*; John-Alder et al. 2009). Relationships between the acute glucocorticoid stress response and fitness also seem to be quite variable (Breuner et al. 2008; Angelier et al. 2009; MacDougall-Shackleton et al. 2009). Fewer studies have reported selection on testosterone. Three studies found evidence of strong positive relationships between circulating testosterone and mating success (Borgia and Wingfield 1991; Alatalo et al. 1996; Mills et al. 2007), whereas other studies found no evidence of selection acting on circulating testosterone (Brown et al. 2005; John-Alder et al. 2009). Our study adds to the growing body of literature demonstrating selection acting on natural variation in hormones.

We found strong evidence of selection acting on GnRH-induced testosterone increases but no evidence of selection acting on pre-GnRH challenge levels. This difference is perhaps unsurprising because GnRH-induced testosterone increases are fairly repeatable within individuals, whereas initial testosterone levels do not show similar individual consistency (Jawor et al. 2006). Our measurements of initial testosterone are not necessarily identical to measurements of breeding baseline testosterone because of the handling time involved in obtaining the blood sample (see appendix); therefore, additional study is required to examine potential relationships between baseline testosterone and fitness.

Testosterone levels produced in response to GnRH injections have been shown to predict those produced in response to territorial intruders, and responses to GnRH challenges are also associated with variation in behavior and plumage (McGlothlin et al. 2007, 2008). Specifically, GnRH-induced testosterone increases were positively correlated with tail white, an attractive plumage trait, and were negatively correlated with nestling feeding rate, while absolute levels of GnRH-induced testosterone were posi-

Table 1: Linear (β) and quadratic (γ) selection gradients measuring relationships between testosterone and components of fitness

Trait	β or γ	SE	F	df	P
Annual survival:					
Initial testosterone	.02	.110	.02	1, 107	.88
GnRH-induced increase	.12	.106	1.27	1, 107	.27
Initial testosterone^2	−.08	.067	1.55	1, 104	.47
GnRH-induced increase^2	**−.30**	**.133**	**4.65**	**1, 104**	**.03**
Initial × increase	.06	.078	.51	1, 104	.48
Year			.59	1, 107	.45
Age			.02	1, 89.8	.89
Number of offspring:					
Initial testosterone	.01	.038	.03	1, 43	.85
GnRH-induced increase	.15	.095	2.58	1, 43	.12
Initial testosterone^2	−.03	.118	.05	1, 40	.82
GnRH-induced increase^2	**−.31**	**.127**	**5.85**	**1, 40**	**.02**
Initial × increase	−.09	.126	.56	1, 40	.46
Year			*2.63*	*1, 43*	*.10*
Age			1.11	1, 43	.26
Number of mates:					
Initial testosterone	.02	.112	.02	1, 49	.90
GnRH-induced increase	−.07	.200	.14	1, 49	.71
Initial testosterone^2	−.08	.123	.38	1, 46	.54
GnRH-induced increase^2	−.24	.161	2.19	1, 46	.15
Initial × increase	−.04	.131	.08	1, 46	.78
Year			1.57	1, 49	.22
Age			.41	1, 49	.53
Offspring per mate:					
Initial testosterone	.02	.098	.06	1, 39.5	.80
GnRH-induced increase	**.26**	**.098**	**7.24**	**1, 38.9**	**.01**
Initial testosterone^2	.01	.114	.02	1, 35.6	.90
GnRH-induced increase^2	.02	.161	.02	1, 36.9	.90
Initial × increase	−.03	.145	.03	1, 33.7	.86
Year				1, 28.9	.31
Age				1, 39.8	.88
Within-pair offspring:					
Initial testosterone	.02	.060	.11	1, 43	.75
GnRH-induced increase	**.23**	**.106**	**4.57**	**1, 43**	**.04**
Initial testosterone^2	.00	.045	.01	1, 40	.91
GnRH-induced increase^2	−.21	.130	2.61	1, 40	.11
Initial × increase	−.13	.187	.49	1, 40	.49
Year			1.45	1, 43	.24
Age			.18	1, 43	.67
Within-pair offspring (controlling for apparent within-pair offspring):					
Initial testosterone	.06	.087	.51	1, 40	.48
GnRH-induced increase	**.18**	**.085**	**4.76**	**1, 40**	**.04**
Initial testosterone^2	.01	.026	.32	1, 37	.58
GnRH-induced increase^2	−.09	.067	1.81	1, 37	.19
Initial × increase	−.11	.296	.15	1, 37	.70
Year			.89	1, 40	.35
Age			.33	1, 40	.57
Apparent within-pair offspring			**40.81**	**1, 40**	**<.001**
Extrapair offspring:					
Initial testosterone	.05	.346	.02	1, 43	.90
GnRH-induced increase	−.01	.011	.40	1, 43	.53
Initial testosterone^2	−.10	.124	.64	1, 40	.43
GnRH-induced increase^2	**−.71**	**.333**	**4.52**	**1, 40**	**.04**

Table 1 (*Continued*)

Trait	β or γ	SE	*F*	df	*P*
Initial × increase	.06	.625	.01	1, 40	.91
Year			*2.96*	*1, 43*	*.09*
Age			2.29	1, 43	.14

Note: The first two effects in each analysis measure directional selection (β). The superscript 2 indicates a squared term. Squared effects, as well as the cross product, represent quadratic selection (γ). Selection gradients are from linear mixed models with normal error variance and a random effect of individual. Significance testing was performed using generalized linear mixed models with binomial (annual survival), normal (offspring per mate), or Poisson (all other fitness components) error structure. Effects with $P < .05$ are shown in bolded type; factors with $P < .1$ are shown in italics. GnRH = gonadotropin-releasing hormone.

tively correlated with territorial aggression (McGlothlin et al. 2007, 2008). Thus, our results demonstrate selection acting on ecologically relevant hormonal changes that are known to be related to a suite of phenotypes involved in territoriality, mate attraction, and parental behavior.

It is likely that our measurements of selection on GnRH-induced increases in testosterone reflect the relative fitness costs and benefits of a large number of behaviors and other traits that interact with the physical and the social environment (McGlothlin and Ketterson 2008). Because hormones often mediate suites of functionally related traits, such traits are expected to experience correlational selection, which occurs when traits interact in their effects on fitness and is predicted to maintain trait integration over time (Lande 1980; Cheverud 1982; Lande and Arnold 1983; Phillips and Arnold 1989; Brodie 1992; Sinervo and Svensson 2002; McGlothlin et al. 2005). Stabilizing and disruptive selection are the univariate analogs of correlational selection (Cheverud 1982; Schluter and Nychka 1994; Blows and Brooks 2003; Blows 2007), and the stabilizing selection observed to be acting on short-term testosterone increases may be an indicator of underlying correlational selection that acts to maintain integration of the traits that testosterone mediates (McGlothlin and Ketterson 2008).
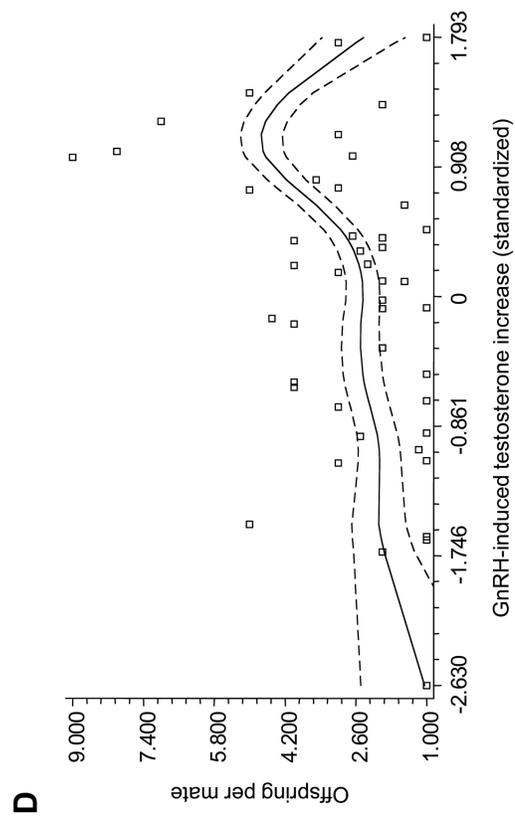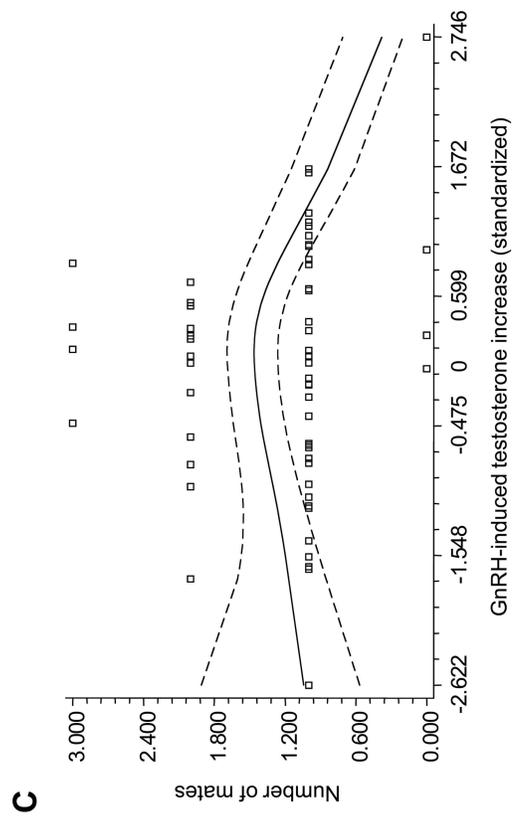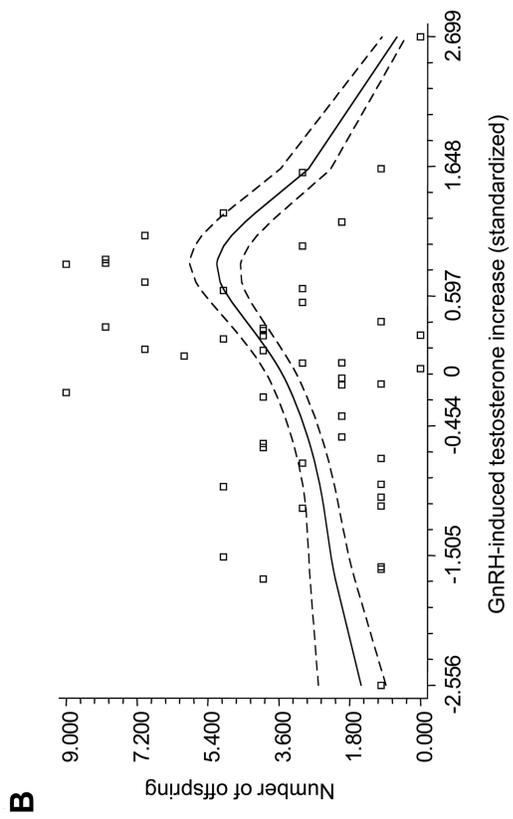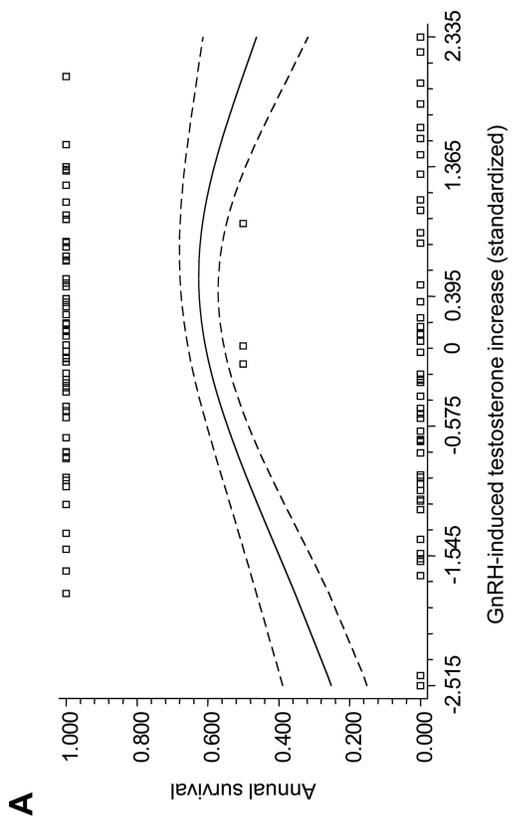
Trade-Offs between Survival and Reproduction

In many species, testosterone has been implicated as a physiological mediator of trade-offs between survival and reproduction, specifically, mating effort (Sinervo and Svensson 1998; Ketterson and Nolan 1999; Wingfield et al. 2001; Adkins-Regan 2005; Reed et al. 2006; Hau 2007; Mills et al. 2009). Most evidence regarding the role of testosterone in trade-off regulation has come from studies that experimentally altered testosterone levels. Specifically, in juncos, testosterone implants have been shown to increase mating success at the expense of survival (Reed et al. 2006). Our data show that this testosterone-mediated trade-off between mating effort and survival does not nec-

essarily translate to opposing selection on natural variation in testosterone. Instead, both survival and reproductive fitness functions showed a pattern of stabilizing selection with a fitness peak slightly to the right of the population mean, indicating that a given testosterone level was favored equally via both components of fitness. Several possible explanations exist for this initially surprising difference between experimental and correlational studies of the fitness effects of testosterone.

First, experimental and observational studies tend to measure subtly different aspects of trade-offs. Several authors have pointed out that trade-offs may be considered in two different senses, functional and statistical (van Noordwijk and de Jong 1986; Houle 1991; de Jong and van Noordwijk 1992; Roff 2002; Roff and Fairbairn 2007). Functional trade-offs represent the decision rule (or "trade-off function") that individuals use to allocate resources to two conflicting traits or activities, whereas statistical trade-offs represent the phenotypic or genetic correlations between traits or fitness components in a population. Individual variation in quality, condition, or resource acquisition may cause the sign of the statistical trade-off to be different from that of the functional trade-off (van Noordwijk and de Jong 1986). This variation in quality or condition may also alter the relationship between phenotype and fitness (Schluter et al. 1991). Studies using hormonal manipulation explore functional trade-offs by experimentally altering the allocation of resources by individuals, but correlational studies necessarily consider statistical relationships in a population. As a consequence, individual variation in quality may mask the signal of functional trade-offs within individuals.

This difference between functional and statistical trade-offs is perhaps the most likely explanation for the differences between our results and those of previous studies using hormonal manipulation. If testosterone production were influenced by a male's quality, we would expect similar patterns of selection acting via survival and reproduction and a positive correlation between survival and reproduction. Indeed, survival and reproductive fitness
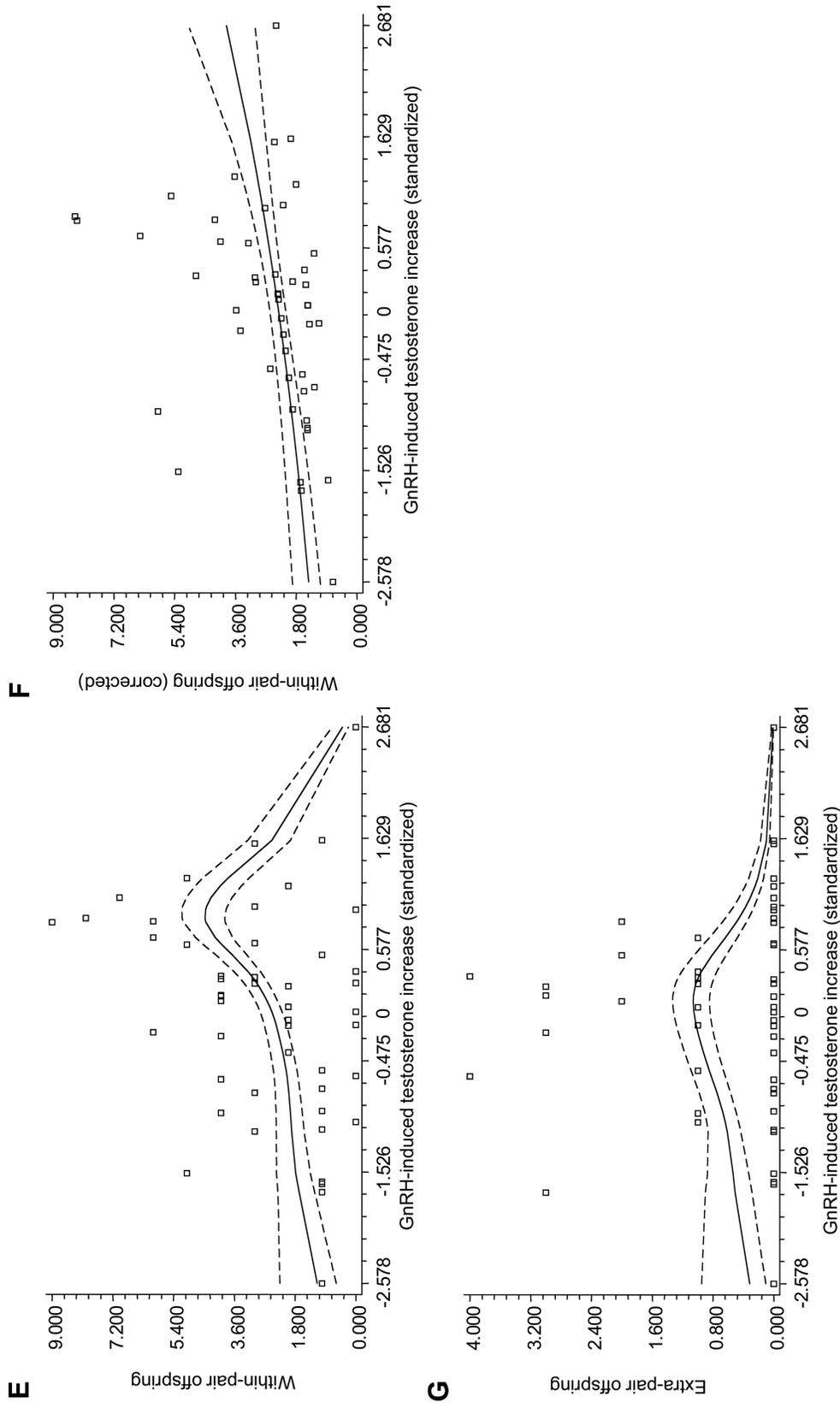
A Annual survival — GnRH-induced testosterone increase (standardized)

B Number of offspring — GnRH-induced testosterone increase (standardized)

C Number of mates — GnRH-induced testosterone increase (standardized)

D Offspring per mate — GnRH-induced testosterone increase (standardized)

694

Figure 1: Cubic spline representations of the relationships between gonadotropin-releasing hormone (GnRH)–induced testosterone change and components of fitness. Testosterone change was standardized to zero mean and unit variance. Points represent the average fitness of males with given levels of testosterone change. Solid lines represent the predicted relationship, and dashed lines represent 1 SE in either direction as calculated by bootstrapping. *A*, Annual offspring production. *C*, Number of mates. *D*, Offspring per mate. *E*, Within-pair offspring. *F*, Within-pair offspring, controlling for the total apparent number of offspring produced in the home nest. *G*, Extrapair offspring. All fitness functions except number of mates consist of one statistically significant selection gradient (table 1).

Table 2: Total annual linear and quadratic selection gradients

Trait	β or γ
Initial testosterone	.03
GnRH-induced increase	.27
Initial testosterone^2	−.13
GnRH-induced increase^2	−.59
Initial × increase	−.06

Note: Total gradients were calculated by adding the selection gradients in table 1 (annual survival and number of offspring) while correcting for changes in the phenotypic (co)variance matrix across selective episodes (see "Material and Methods" for details). As in table 1, the first two effects in each analysis measure directional selection (β). The superscripted 2 indicates a squared term. Squared effects, as well as the cross product, represent quadratic selection (γ). GnRH = gonadotropin-releasing hormone.

functions were almost identical, and there was a positive correlation between survival and reproduction (see appendix), suggesting the males that were successful at survival were also those that were successful at reproduction. The most successful males (those with medium-high GnRH-induced testosterone increases) may be high-quality males that are able to achieve an optimal allocation between mating effort and survival effort that allows for both. Many of the males to the left of the fitness peak are likely to be low-quality males that are physiologically unable to produce or maintain the cost of higher testosterone levels. Poor condition or genetic quality (e.g., deleterious mutations) may account for their decreased fitness independently of testosterone level.

If individuals varied only in quality, we might expect selection on GnRH-induced testosterone production to be primarily directional, acting in the same direction via both survival and reproduction. Variation in allocation between mating effort and survival effort caused by testosterone is also likely to be important, and it may contribute to the stabilizing aspects of the fitness functions. For example, males with very high testosterone increases to the right of the survival optimum may suffer reduced survival because they allocate too much to mating effort at the expense of self-maintenance. Males to the right of the reproductive optimum may be investing highly in mating effort but not receiving fitness benefits, perhaps because females find very intense testosterone-mediated behaviors unattractive or because these males do not strike the right balance between courtship and competition (Qvarnström and Forsgren 1998; Hunt et al. 2009). Allocation may also contribute to the lower reproductive fitness of males with low GnRH-induced testosterone production. For example, some of the males to the left of the reproductive fitness peak may have underinvested in mating effort, leading to lower mating success. Such underinvestment may reduce

survival as well, perhaps due to a correlated overinvestment in costly parental care (McGlothlin et al. 2007).

A second, not mutually exclusive explanation for the difference between our results and those of experimental studies of testosterone is that the fitness effects of the production of short-term testosterone increases may differ from those of the constitutively elevated testosterone levels produced by implants. Although testosterone implants increase testosterone to the levels naturally produced during short-term testosterone increases (Ketterson et al. 1992; McGlothlin et al. 2008), they remove the ability to facultatively lower testosterone levels, which may impose additional survival costs or reproductive benefits. When testosterone levels were elevated experimentally, the average survival of testosterone-treated males was 38% (Reed et al. 2006), which is somewhat lower than that predicted for males with the very highest GnRH-induced testosterone changes in this study (∼46%). This pattern suggests that the ability to modulate testosterone levels when needed may ameliorate some of the costs of testosterone, and this ability may be favored over maintaining constitutively high testosterone levels (Wingfield et al. 2001).

A third explanation is that the fitness effects of testosterone may depend on environmental variation over time. For example, our population of juncos experiences cycling due to acorn mast and predator abundance, which is associated with variation in adult population density, demography, and nest success (Clotfelter et al. 2007). In 2003–2004, the years during which this study was conducted, the population was characterized by low breeding density, a demography shifted toward older adults, and high nest success (J. W. McGlothlin, personal observation). In contrast, the experimental studies summarized by Reed et al. (2006) were conducted from 1993 to 2001, a time span representing a wide range of environments. To further complicate matters, many of the costs and benefits of testosterone depend on social interactions, whose fitness effects are likely to be frequency dependent (Maynard Smith 1982; Sinervo and Calsbeek 2006). In previous experimental studies, the social environment consisted of a bimodal distribution of testosterone-treated and control males, whereas in our study, the distribution of testosterone levels in the population was more continuous. The absence of testosterone-treated males is likely to be the most important difference in the social environment between studies and is likely to have affected relationships between testosterone and fitness, adding to the difficulty of comparing implanted and unimplanted populations. The social environment may have differed between the studies due to natural fluctuations arising from the factors discussed above.

Within-Pair versus Extrapair Reproduction

Another striking difference between results obtained from implant studies and natural variation regards the production of within-pair versus extrapair offspring. Experimental testosterone enhancement led to the production of more extrapair offspring (Raouf et al. 1997; Reed et al. 2006). Testosterone-treated males lost significantly more paternity from their home nest in an early study (Raouf et al. [1997] analyzed data from 1990 to 1993) but not in a larger data set (Reed et al. [2006] analyzed data from 1993 to 2000). In contrast, we found that greater GnRH-induced testosterone increases led to greater success in siring within-pair offspring, but it did not lead to greater extrapair success. Rather, selection via extrapair offspring was stabilizing, with males with levels slightly below the population mean having the highest success. The correlation between the two fitness components was very small (see appendix), suggesting that increased within-pair success did not necessarily lead to decreased extrapair success.

Factors similar to those considered above may also help to explain this difference. First, short-term testosterone increases may not affect behavior in the same way as do testosterone implants. Parental care, song, and aggression are associated with exogenous and natural variation in testosterone (as assessed via GnRH challenges) in similar ways (McGlothlin et al. 2007), but as of yet we have no data on other behaviors thought to be important to extrapair mating, such as home range and courtship (Chandler et al. 1994; Enstrom et al. 1997). As evidenced by their relationship with territorial intrusions, short-term testosterone increases may indeed favor mating effort related to within-pair success, such as mate guarding, but not behavior directed toward extrapair mating. Second, temporal variation in the physical or the social environment across studies may have altered relationships between testosterone and extrapair mating. The rate of extrapair fertilizations in our population varies as a result of unknown causes that may include predator-driven demographic cycling (Nolan et al. 2002; Clotfelter et al. 2007). Although the extrapair fertilization rate in 2003–2004 was above average for our population (Nolan et al. 2002), this higher rate does not necessarily translate into stronger selection, which should depend on the variance among males in extrapair mating success (Webster et al. 1995; Shuster and Wade 2003; Whittingham and Dunn 2005). It is possible that the males in this study that had very high GnRH-induced testosterone production spent a great deal of time searching for extrapair mating opportunities but could find only few, or they were unable to successfully court the females they did locate. If, for example, these males court females too intensely, females may be threatened, leading to fewer successful extrapair fertilizations (Patri-

celli et al. 2002). As discussed above, the social environment may have frequency-dependent effects on extrapair mating success. In particular, demographic factors may be at play; in our sample, 75% of males were second-year breeders or older. Previous studies have indicated that second-year and older males are more likely to gain extrapair fertilizations (Reed et al. 2006). Although we did not detect a significant age effect, only two first-year males sired any extrapair young, each achieving one extrapair fertilization. It is likely that many of the factors listed here interact to produce the curvilinear relationship between GnRH-induced testosterone and extrapair fertilization success, suggesting that optimal testosterone production may reflect the interaction of numerous selective factors and may vary over time.

Interpretation of GnRH-Induced Testosterone Production

Evidence presented here and elsewhere indicates that GnRH challenging may be used to estimate variation among individual male juncos in their ability to produce short-term testosterone increases and that such variation can be related to phenotype and fitness (Jawor et al. 2006; McGlothlin et al. 2007, 2008). The method is also useful in female juncos, where GnRH induces an elevation in testosterone (although only when a female is producing eggs), and GnRH-induced testosterone change predicts the amount of testosterone deposited in the yolk (Jawor et al. 2007). Proper caution must be taken, however, when interpreting these results and when applying GnRH challenges in other systems, because much remains to be learned.

First, some species do not produce short-term testosterone increases, and thus GnRH challenges may not be informative about individual variation in them. For example, male green anoles (*Anolis carolinensis*) do not increase testosterone when challenged with GnRH (Husak et al. 2009). Second, higher-level effects on the HPG axis may reduce the ability of GnRH challenges to predict socially induced changes in testosterone or variation in behavior. For example, male Gambel's white-crowned sparrows (*Zonotrichia leucophrys gambelii*) do not socially modulate testosterone, despite increases in testosterone when GnRH challenges are administered (Meddle et al. 2002). This subspecies also shows behavioral insensitivity to exogenous testosterone (Meddle et al. 2002). This pattern has been demonstrated in a number of other species as well (Lynn 2008).

Third, it is crucial to understand that no single measurement of a hormonal pathway is likely to capture all the ecologically and behaviorally important variation, and the links between GnRH and testosterone production represent only a small portion of the many variables that

influence the expression of testosterone-mediated phenotypes. Thus, although GnRH-induced testosterone level is related to behavior in our population (McGlothlin et al. 2007), much variation in the behavior remains unexplained. Individuals are also likely to vary in other aspects of the hormonal pathway that cannot be captured by the GnRH challenge, such as high-level neural regulation of the HPG axis, effects of other interacting pathways (such as the adrenocortical stress axis), and downstream effects such as receptor expression (Adkins-Regan 2005; Hau 2007; Ball and Balthazart 2008; McGlothlin and Ketterson 2008). Natural selection is likely to shape these aspects of the testosterone pathway as well, and future studies should investigate this complexity (McGlothlin and Ketterson 2008; Ketterson et al. 2009).

Finally, we stress that even when GnRH challenges provide a robust estimate of individual variation in testosterone production, studies using this method must not be blindly interpreted as demonstrating mechanisms of testosterone action. As a case in point, although males with larger tail white patches respond more strongly to GnRH (McGlothlin et al. 2008), we do not expect this to reflect a direct mechanistic effect because exogenous testosterone inhibits molt in this species (Nolan et al. 1992), even in very low doses (J. M. Jawor, J. W. McGlothlin, and E. D. Ketterson, unpublished data). The most profitable approach is pluralistic, combining phenotypic engineering and other mechanistic approaches to uncover physiological causation with careful studies of patterns of covariation between physiology, morphology, and behavior to assess the ecological relevance of individual variation in hormone levels.

Conclusion

Our results indicate that testosterone is associated with multiple components of fitness in the wild. The strong stabilizing selection acting on GnRH-induced testosterone increases via both survival and reproduction suggests that selection acts to maintain the integration of the suite of traits mediated by testosterone. Despite this strong stabilizing component, we also found evidence that males with higher testosterone production may be favored by selection, primarily through within-pair fertilization success. However, we did not find evidence that patterns of directional selection directly mirrored a functional trade-off between survival and reproduction; rather, individual variation in condition or quality appeared to obscure the predicted trade-off, such that individuals that were successful at reproduction were also more likely to survive. These results suggest that a full understanding of the relationships between hormones and fitness components will require both experimental and correlational approaches,

illustrating the utility of combining manipulative experiments with studies of natural variation (Kempenaers et al. 2008; McGlothlin and Ketterson 2008; Williams 2008). If we are to understand the evolutionary consequences of the patterns of selection in the wild, future work should explore the genetic and environmental causes of individual variation in testosterone production. Furthermore, as our study measured GnRH-induced testosterone production averaged across the breeding season and assessed only annual selection, future work should also examine seasonal variation in testosterone and measure lifetime fitness in order to more fully characterize how selection shapes testosterone-mediated life-history trade-offs. Finally, if we are to predict evolution from patterns of selection, we must know far more about the inheritance of short-term and long-term patterns of testosterone production than we currently do.

Literature Cited

Adkins-Regan, E. 2005, Hormones and animal social behavior. Princeton University Press, Princeton, NJ.

Alatalo, R. V., J. Höglund, A. Lundberg, P. T. Rintamäki, and B. Silverin. 1996. Testosterone and male mating success on the black

grouse leks. Proceedings of the Royal Society B: Biological Sciences 263:1697–1702.

Angelier, F., R. L. Holberton, and P. P. Marra. 2009. Does stress response predict return rate in a migratory bird species? a study of American redstarts and their non-breeding habitat. Proceedings of the Royal Society B: Biological Sciences 276:3545–3551.

Arnold, S. J., and M. J. Wade. 1984. On the measurement of natural and sexual selection: theory. Evolution 38:709–719.

Ball, G. F., and J. Balthazart. 2008. Individual variation and the endocrine regulation of behaviour and physiology in birds and other vertebrates: a cellular/molecular perspective. Philosophical Transactions of the Royal Society B: Biological Sciences 363:1699–1710.

Blas, J., G. R. Bortolotti, J. L. Tella, R. Baos, and T. A. Marchant. 2007. Stress response during development predicts fitness in a wild, long lived vertebrate. Proceedings of the National Academy of Sciences of the USA 104:8880–8884.

Blows, M. W. 2007. A tale of two matrices: multivariate approaches in evolutionary biology. Journal of Evolutionary Biology 20:1–8.

Blows, M. W., and R. Brooks. 2003. Measuring nonlinear selection. American Naturalist 162:815–820.

Bonier, F., P. R. Martin, K. S. Sheldon, J. P. Jensen, S. L. Foltz, and J. C. Wingfield. 2007. Sex-specific consequences of life in the city. Behavioral Ecology 18:121–129.

Bonier, F., P. R. Martin, I. T. Moore, and J. C. Wingfield. 2009*a*. Do baseline glucocorticoids predict fitness? Trends in Ecology & Evolution 24:634–642.

Bonier, F., I. T. Moore, P. R. Martin, and R. J. Robertson. 2009*b*. The relationship between fitness and baseline glucocorticoids in a passerine bird. General and Comparative Endocrinology 163:208–213.

Borgia, G., and J. C. Wingfield. 1991. Hormonal correlates of bower decoration and sexual display in the satin bowerbird (*Ptilonorhynchus violaceus*). Condor 93:935–942.

Breuner, C. W., S. H. Patterson, and T. P. Hahn. 2008. In search of relationships between the acute adrenocortical response and fitness. General and Comparative Endocrinology 157:288–295.

Brodie, E. D., III. 1992. Correlational selection for color pattern and antipredator behavior in the garter snake *Thamnophis ordinoides*. Evolution 46:1284–1298.

Brodie, E. D., III, A. J. Moore, and F. J. Janzen. 1995. Visualizing and quantifying natural selection. Trends in Ecology & Evolution 10:313–318.

Brown, C. R., M. B. Brown, S. A. Raouf, L. C. Smith, and J. C. Wingfield. 2005. Effects of endogenous steroid hormone levels on annual survival in cliff swallows. Ecology 86:1034–1046.

Cabezas, S., J. Blas, T. A. Marchant, and S. Moreno. 2007. Physiological stress levels predict survival probabilities in wild rabbits. Hormones and Behavior 51:313–320.

Chandler, C. R., E. D. Ketterson, V. Nolan Jr., and C. Ziegenfus. 1994. Effects of testosterone on spatial activity in free-ranging male dark-eyed juncos, *Junco hyemalis*. Animal Behaviour 47:1445–1455.

Cheverud, J. M. 1982. Phenotypic, genetic, and environmental morphological integration in the cranium. Evolution 36:499–516.

Clotfelter, E. D., D. M. O'Neal, J. M. Gaudioso, J. M. Casto, I. M. Parker-Renga, E. A. Snajdr, D. L. Duffy, V. Nolan Jr., and E. D. Ketterson. 2004. Consequences of elevating plasma testosterone in females of a socially monogamous songbird: evidence of constraints on male evolution? Hormones and Behavior 46:171–178.

Clotfelter, E. D., A. B. Pedersen, J. A. Cranford, N. Ram, E. A. Snajdr,

V. Nolan Jr., and E. D. Ketterson. 2007. Acorn mast drives long-term dynamics of rodent and songbird populations. Oecologia (Berlin) 154:493–503.

de Jong, G., and A. J. van Noordwijk. 1992. Acquisition and allocation of resources: genetic (co)variances, selection, and life histories. American Naturalist 139:749–770.

Enstrom, D. A., E. D. Ketterson, and V. Nolan Jr. 1997. Testosterone and mate choice in the dark-eyed junco. Animal Behaviour 54:1135–1146.

Finch, C. E., and M. R. Rose. 1995. Hormones and the physiological architecture of life-history evolution. Quarterly Review of Biology 70:1–52.

Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson. 2006. ASReml user guide, release 2.0. VSN International, Hemel Hempstead.

Goymann, W. 2009. Social modulation of androgens in male birds. General and Comparative Endocrinology 163:149–157.

Goymann, W., M. M. Landys, and J. C. Wingfield. 2007. Distinguishing seasonal androgen responses from male-male androgen responsiveness: revisiting the challenge hypothesis. Hormones and Behavior 51:463–476.

Harding, C. F. 1981. Social modulation of circulating hormone levels in the male. American Zoologist 21:223–231.

Hau, M. 2007. Regulation of male traits by testosterone: implications for the evolution of vertebrate life histories. BioEssays 29:133–144.

Hirschenhauser, K., and R. F. Oliveira. 2006. Social modulation of androgens in male vertebrates: meta-analyses of the challenge hypothesis. Animal Behaviour 71:265–277.

Hirschenhauser, K., H. Winkler, and R. F. Oliveira. 2003. Comparative analysis of male androgen responsiveness to social environment in birds: the effects of mating system and paternal incubation. Hormones and Behavior 43:508–519.

Houle, D. 1991. Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters. Evolution 45:630–648.

Hunt, J., C. J. Breuker, J. A. Sadowski, and A. J. Moore. 2009. Male-male competition, female mate choice, and their interaction: determining total sexual selection. Journal of Evolutionary Biology 22:13–26.

Husak, J. F., D. J. Irschick, J. P. Henningsen, K. S. Kirkbride, S. P. Lailvaux, and I. T. Moore. 2009. Hormonal response of male green anole lizards (*Anolis carolinensis*) to GnRH challenge. Journal of Experimental Zoology A: Ecological Genetics and Physiology 311A: 105–114.

Jawor, J. M., J. W. McGlothlin, J. M. Casto, T. J. Greives, E. A. Snajdr, G. E. Bentley, and E. D. Ketterson. 2006. Seasonal and individual variation in response to GnRH challenge in male dark-eyed juncos (*Junco hyemalis*). General and Comparative Endocrinology 149: 182–189.

———. 2007. Testosterone response to GnRH in a female songbird varies with stage of reproduction: implications for adult behaviour and maternal effects. Functional Ecology 21:767–775.

John-Alder, H. B., R. M. Cox, G. J. Haenel, and L. C. Smith. 2009. Hormones, performance and fitness: natural history and endocrine experiments on a lizard (*Sceloporus undulatus*). Integrative and Comparative Biology 49:393–407.

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

Kempenaers, B., A. Peters, and K. Foerster. 2008. Sources of individual variation in plasma testosterone levels. Philosophical Transactions of the Royal Society B: Biological Sciences 363:1711–1723.

Ketterson, E. D., and V. Nolan Jr. 1992. Hormones and life histories: an integrative approach. American Naturalist 140(suppl.):S33–S62.

———. 1999. Adaptation, exaptation, and constraint: a hormonal perspective. American Naturalist 154(suppl.):S4–S25.

Ketterson, E. D., V. Nolan Jr., L. Wolf, and C. Ziegenfus. 1992. Testosterone and avian life histories: effects of experimentally elevated testosterone on behavior and correlates of fitness in the dark-eyed junco (*Junco hyemalis*). American Naturalist 140:980–999.

Ketterson, E. D., V. Nolan Jr., M. J. Cawthorn, P. G. Parker, and C. Ziegenfus. 1996. Phenotypic engineering: using hormones to explore the mechanistic and functional bases of phenotypic variation in nature. Ibis 138:70–86.

Ketterson, E. D., P. G. Parker, S. A. Raouf, V. Nolan Jr., C. Ziegenfus, and C. R. Chandler. 1997. The relative impact of extra-pair fertilizations on variation in male and female reproductive success in dark-eyed juncos (*Junco hyemalis*). Pages 81–101 *in* P. G. Parker and N. T. Burley, eds. Avian reproductive tactics: female and male perspectives. Ornithological Monographs 49. Allen, Lawrence, KS.

Ketterson, E. D., J. W. Atwell, and J. W. McGlothlin. 2009. Phenotypic integration and independence: hormones, performance, and response to environmental change. Integrative and Comparative Biology 49:365–379.

Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, A. Hoang, P. Gibert, and P. Beerli. 2001. The strength of phenotypic selection in natural populations. American Naturalist 157:245–261.

Lande, R. 1980. The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94:203–215.

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

Landys, M. M., W. Goymann, M. Raess, and T. Slagsvold. 2007. Hormonal responses to male-male social challenge in the blue tit *Cyanistes caeruleus*: single-broodedness as an explanatory variable. Physiological and Biochemical Zoology 80:228–240.

Lessells, C. M. 2008. Neuroendocrine control of life histories: what do we need to know to understand the evolution of phenotypic plasticity? Philosophical Transactions of the Royal Society B: Biological Sciences 363:1589–1598.

Longmire, J. L., G. F. Gee, O. L. Handenkipf, and G. A. Mark. 1992. Establishing paternity in whooping cranes, *Grus americana*, by DNA analysis. Auk 109:522–529.

Lynn, S. E. 2008. Behavioral insensitivity to testosterone: why and how does testosterone alter paternal and aggressive behavior in some avian species but not others? General and Comparative Endocrinology 157:233–240.

MacDougall-Shackleton, S. A., L. Dindia, A. E. M. Newman, D. A. Potvin, K. A. Stewart, and E. A. MacDougall-Shackleton. 2009. Stress, song and survival in sparrows. Biology Letters 5:746–748.

Marshall, T. C., J. Slate, L. E. B. Kruuk, and J. M. Pemberton. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7:639–655.

Maynard Smith, J. 1982. Evolution and the theory of games. Cambridge University Press, Cambridge.

McGlothlin, J. W. 2010. Combining selective episodes to estimate lifetime nonlinear selection. Evolution, doi:10.1111/j.1558-5646.2009.00914.x.

McGlothlin, J. W., and E. D. Ketterson. 2008. Hormone-mediated suites as adaptations and evolutionary constraints. Philosophical Transactions of the Royal Society B: Biological Sciences 363:1611–1620.

McGlothlin, J. W., P. G. Parker, V. Nolan Jr., and E. D. Ketterson. 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., J. M. Jawor, and E. D. Ketterson. 2007. Natural variation in a testosterone-mediated trade-off between mating effort and parental effort. American Naturalist 170:864–875.

McGlothlin, J. W., J. M. Jawor, T. J. Greives, J. M. Casto, J. L. Phillips, and E. D. Ketterson. 2008. Hormones and honest signals: males with larger ornaments elevate testosterone more when challenged. Journal of Evolutionary Biology 21:39–48.

Meddle, S. L., L. M. Romero, L. B. Astheimer, W. A. Buttemer, I. T. Moore, and J. C. Wingfield. 2002. Steroid hormone interrelationships with territorial aggression in an arctic-breeding songbird, Gambel's white-crowned sparrow, *Zonotrichia leucophrys gambelii*. Hormones and Behavior 42:212–221.

Millesi, E., I. E. Hoffmann, S. Steurer, M. Metwaly, and J. P. Dittami. 2002. Vernal changes in the behavior and endocrine responses to GnRH application in male European ground squirrels. Hormones and Behavior 41:51–58.

Mills, S. C., A. Grapputo, E. Koskela, and T. Mappes. 2007. Quantitative measure of sexual selection with respect to the operational sex ratio: a comparison of selection indices. Proceedings of the Royal Society B: Biological Sciences 274:143–150.

Mills, S. C., L. Hazard, L. Lancaster, T. Mappes, D. Miles, T. A. Oksanen, and B. Sinervo. 2008. Gonadotropin hormone modulation of testosterone, immune function, performance, and behavioral trade-offs among male morphs of the lizard *Uta stansburiana*. American Naturalist 171:339–357.

Mills, S. C., A. Grapputo, I. Jokinen, E. Koskela, T. Mappes, T. A. Oksanen, and T. Poikonen. 2009. Testosterone-mediated effects on fitness-related phenotypic traits and fitness. American Naturalist 173:475–487.

Moore, I. T., N. Perfito, H. Wada, T. S. Sperry, and J. C. Wingfield. 2002. Latitudinal variation in plasma testosterone levels in birds of the genus *Zonotrichia*. General and Comparative Endocrinology 129:13–19.

Moore, M. C. 1983. Effect of female sexual displays on the endocrine physiology and behavior of male white-crowned sparrows, *Zonotrichia leucophrys*. Journal of Zoology 199:137–148.

———. 1991. Application of organization-activation theory to male reproductive strategies: a review. Hormones and Behavior 25:154–179.

Nijhout, H. F. 2003. Development and evolution of adaptive polyphenisms. Evolution and Development 5:9–18.

Nolan, V., Jr., E. D. Ketterson, C. Ziegenfus, D. P. Cullen, and C. R. Chandler. 1992. Testosterone and avian life histories: effects of experimentally elevated testosterone on prebasic molt and survival in male dark-eyed juncos. Condor 94:364–370.

Nolan, V., Jr., E. D. Ketterson, D. A. Cristol, C. M. Rogers, E. D. Clotfelter, R. C. Titus, S. J. Schoech, and E. Snajdr. 2002. Dark-eyed junco (*Junco hyemalis*). *In* A. Poole and F. Gill, eds. The birds of North America, no. 716. http://bna.birds.cornell.edu/bna/species/716. Birds of North America, Philadelphia.

Patricelli, G. L., J. A. C. Uy, G. Walsh, and G. Borgia. 2002. Male displays adjusted to female's response. Nature 415:279–280.

Phillips, P. C., and S. J. Arnold. 1989. Visualizing multivariate selection. Evolution 43:1209–1222.

Pinxten, R., E. de Ridder, L. Arckens, V. M. Darras, and M. Eens. 2007. Plasma testosterone levels of male European starlings (*Sturnus vulgaris*) during the breeding cycle and in relation to song and paternal care. Behaviour 144:393–410.

Qvarnström, A., and E. Forsgren. 1998. Should females prefer dominant males? Trends in Ecology & Evolution 13:498–501.

Raouf, S. A., P. G. Parker, E. D. Ketterson, V. Nolan Jr., and C. Ziegenfus. 1997. Testosterone affects reproductive success by influencing extra-pair fertilizations in male dark-eyed juncos (Aves: *Junco hyemalis*). Proceedings of the Royal Society B: Biological Sciences 264:1599–1603.

Reed, W. L., M. E. Clark, P. G. Parker, S. A. Raouf, N. Arguedas, D. S. Monk, E. Snajdr, V. Nolan Jr., and E. D. Ketterson. 2006. Physiological effects on demography: a long-term experimental study of testosterone's effects on fitness. American Naturalist 167:667–683.

Ricklefs, R. E., and M. Wikelski. 2002. The physiology/life-history nexus. Trends in Ecology & Evolution 17:462–467.

Roff, D. A. 2002. Life history evolution. Sinauer, Sunderland, MA.

Roff, D. A., and D. J. Fairbairn. 2007. The evolution of trade-offs: where are we? Journal of Evolutionary Biology 20:433–447.

Schluter, D. 1988. Estimating the form of natural selection on a quantitative trait. Evolution 42:849–861.

Schluter, D., and D. Nychka. 1994. Exploring fitness surfaces. American Naturalist 143:597–616.

Schluter, D., T. D. Price, and L. Rowe. 1991. Conflicting selection pressures and life history trade-offs. Proceedings of the Royal Society B: Biological Sciences 246:117–122.

Shuster, S. M., and M. J. Wade. 2003. Mating systems and strategies. Princeton University Press, Princeton, NJ.

Sinervo, B., and R. Calsbeek. 2006. The developmental, physiological, neural, and genetical causes and consequences of frequency-dependent selection in the wild. Annual Review of Ecology, Evolution, and Systematics 37:581–610.

Sinervo, B., and E. Svensson. 1998. Mechanistic and selective causes of life history trade-offs and plasticity. Oikos 83:432–442.

———. 2002. Correlational selection and the evolution of genomic architecture. Heredity 89:329–338.

Stearns, S. C. 1992. The evolution of life histories. Oxford University Press, New York.

Stinchcombe, J. R., A. F. Agrawal, P. A. Hohenlohe, S. J. Arnold, and M. W. Blows. 2008. Estimating nonlinear selection gradients using quadratic regression coefficients: double or nothing? Evolution 62:2435–2440.

van Noordwijk, A. J., and G. de Jong. 1986. Acquisition and allocation of resources: their influence on variation in life-history tactics. American Naturalist 128:137–142.

Wade, M. J., and S. Kalisz. 1989. The additive partitioning of selection gradients. Evolution 43:1567–1569.

Webster, M. S., S. Pruett-Jones, D. F. Westneat, and S. J. Arnold. 1995. Measuring the effects of pairing success, extra-pair copulations and mate quality on the opportunity for sexual selection. Evolution 49:1147–1157.

Whittingham, L. A., and P. O. Dunn. 2005. Effects of extra-pair and within-pair reproductive success on the opportunity for selection in birds. Behavioral Ecology 16:138–144.

Williams, T. D. 2008. Individual variation in endocrine systems: moving beyond the "tyranny of the Golden Mean." Philosophical Transactions of the Royal Society B: Biological Sciences 363:1699–1710.

Wingfield, J. C. 1985. Short-term changes in plasma levels of hormones during establishment and defense of a breeding territory in male song sparrows, *Melospiza melodia*. Hormones and Behavior 19:174–187.

Wingfield, J. C., and D. S. Farner. 1993. Endocrinology of reproduction in wild species. Pages 163–327 *in* D. S. Farner, J. R. King, and K. C. Parkes, eds. Avian biology. Academic Press, London.

Wingfield, J. C., G. F. Ball, A. M. Dufty Jr., R. E. Hegner, and M. Ramenofsky. 1987. Testosterone and aggression in birds. American Scientist 75:602–608.

Wingfield, J. C., R. E. Hegner, A. M. Dufty Jr., and G. F. Ball. 1990. The "challenge hypothesis": theoretical implications for patterns of testosterone secretion, mating systems, and breeding systems. American Naturalist 136:829–846.

Wingfield, J. C., J. D. Jacobs, A. D. Tramontin, N. Perfito, S. Meddle, D. L. Maney, and K. Soma. 2000. Toward an ecological basis of hormone-behavior interactions in reproduction of birds. Pages 86–128 *in* K. Wallen and J. E. Schneider, eds. Reproduction in context. MIT Press, Cambridge, MA.

Wingfield, J. C., S. E. Lynn, and K. K. Soma. 2001. Avoiding the "costs" of testosterone: ecological bases of hormone-behavior interactions. Brain, Behavior and Evolution 57:239–251.

Zera, A. J., and L. G. Harshman. 2001. The physiology of life-history trade-offs in animals. Annual Review of Ecology and Systematics 32:95–126.