Phenotypic Integration of Sexually
Selected Traits in a Songbird

Joel W. McGlothlin

Presented to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
In the Department of Biology
Indiana University
June 2007
To Mary-Louise
Acknowledgments

The completion of this dissertation would have been impossible without the support, hard work, and general good will of a large number of people. First, I thank my advisor Ellen Ketterson, with whom it has been a joy to work for the past six years. Actually, it has been seven. I first worked with Ellen in the summer before my senior year at Mountain Lake Biological Station (MLBS) as part of the Research Experience for Undergraduates (REU) program. My experience that summer was wonderful (for which Diane Neudorf also gets my thanks) as well as instrumental in determining my future scientific trajectory. Ellen not only taught me how to handle birds and take blood samples, she also gave me my first reading assignment on quantitative genetics. During my Ph. D. work, Ellen has been a consistently supportive advisor. She allowed me the freedom to pursue my own divergent project my first year, which turned out to be a data-free learning experience thanks to corvid stubbornness and my own naiveté. Luckily, and wisely, Ellen had suggested I pursue a “side project,” which developed into something not so peripheral after all (see Chapter 2). In the years that followed, Ellen taught me how to be a good scientist, from conceiving a project to writing up the findings, and for that, I am extremely grateful. If the work here has any merit (and I hope it does), Ellen deserves a great deal of the credit.

I want to thank the other members of my thesis committee, who have provided me with the right mix of encouragement and challenge. Butch Brodie gets triple thanks for being a stellar committee member, for teaching me everything I know about quantitative genetics, and for giving me a job. As is obvious by the number of times his name appears
in my bibliography, Butch has been extremely influential in shaping my scientific worldview. Emília Martins has provided invaluable statistical advice and has stimulated me to think comparatively and historically. Dale Sengelaub has been the consistent voice of pragmatism on my committee, reminding me to always keep mechanism in mind.

Although my name appears alone on the title page, I could not have done any of this work by myself. My dissertation work has been largely the result of being in the right place, surrounded by the right people, at the right time. Jodie Jawor not only did two years worth of hormone assays that form a major basis for Chapters 3-5, she also helped me in the field in two different summers, taught me how to do my own hormone assays, and collaborated with me on numerous other projects. Joe Casto (Chapter 3) was the first to figure out how to inject a junco with GnRH. Joe also worked with me in the field during my first full season at Mountain Lake, helping me to get the ball rolling on my fieldwork. Tim Greives worked with me in the field for two years, helped me develop Chapter 3, and has been great to talk science with in general. I thank Val Nolan for his work on Chapter 2. Val also helped me write my very first scientific paper. I have a vivid memory of receiving the manuscript after he had read it. It was covered in pencil marks which managed to transform my pedestrian prose into something quite elegant. I thank Patty Parker, at University of Missouri, St. Louis, for orchestrating the molecular work that was crucial to Chapter 2, and Jenny Phillips for her work as an REU student that was included in Chapter 3. Finally, I want to thank Danielle Whittaker and Sara Schrock for continuing to collaborate with me on a project that didn’t quite make it before deadline.

Conducting a large field project like the one at MLBS requires a lot of help from a lot of people. I owe my greatest thanks to Eric Snajdr, who has been working on the
The success of the junco project as a whole is largely due to his efforts. Beyond those listed above, Amanda Bessler, Jackie Gaudioso, Nicki Gerlach, Katie Hill, Carrington Knox, Annie Lindsay, Dawn O’Neal, Ian Parker-Renga, Katie Pavlis, Peter Stevens, Charles Ziegenfus, and Devin Zysling have helped immensely in the field. (Chapter 2 also benefitted from the labors of dozens of junco workers of yore. I don’t have space to list them all again here.) Other non-junco people helped me in many ways at MLBS, including Henry Wilbur (director), Eric Nagy (associate director), Julian McCroskey, Stephanie Held, Patrice Ludwig, and Volker Rudolf.

I have been lucky to be surrounded by an amazing group of people in the Ketterson Lab for the past few years. Britt Heidinger, in particular, has talked with me about every idea I have had, even the bad ones, and read almost every word I have written (often more than once). Both Britt and Jen Grindstaff were extremely kind to show me the graduate school ropes and made the early years much easier. I thank my other lab mates Eva Allen, Jonathan Atwell, Christy Bergeon, Kristal Cain, Joe Casto, Deb Duffy, Nicki Gerlach, Jodie Jawor, Ryan Kiley, Dawn O’Neal, Wendy Reed, Heather Rupp, Sara Schrock, Beth Schultz, Eric Snajdr, Lynn Siefferman, Brandi Van Roo, Danielle Whittaker, and Wendy Wolf for all you have done, and for making the Ketterson Lab such an intellectually stimulating place to be.

I am grateful to many others at Indiana University who have helped me in various ways, including Heather Bleakley, Gretchen Clearwater, Greg Demas, Erin Kelso, Terry Ord, Amy Poehlman, and Mike Wade. My thanks also go to a number of non-Indiana
Acknowledgments

folks, including Dave McCauley (my undergraduate advisor who first introduced me to MLBS), Jeff Conner, Trevor Price, and John Wingfield.

The funding for this work was provided by a National Science Foundation Doctoral Dissertation Improvement Grant, a Louis Agassiz Fuertes Award from the Wilson Ornithological Society, an American Ornithologists Union Research Award, and awards from Mountain Lake Biological Station, Sigma Xi, Indiana University Graduate School, IU Department of Biology, and the Center for the Integrative Study of Animal Behavior. Also, I have been fortunate enough to receive fellowships to allow me to be in the field when I would have otherwise been in the classroom. These have been provided by the NSF Graduate Research Fellowship Program, a National Institutes of Health Research Training Grant to IU (“Common Themes in Reproductive Diversity”), and a summer fellowship from CISAB.

Finally, I am extremely grateful for the love and encouragement shown by my parents, Chris and Nancy McGlothlin, and my wife Mary-Louise Maas. My parents, who are both retired teachers, created an environment early on that set me down the path of perpetual education. Mary-Louise has been there to see me through all the vicissitudes of graduate school and of life in general. I truly could not have made it without her.
Preface

Phenotypic integration, defined at the most basic level, refers to how different parts of an organism fit together. Olson and Miller (1958) defined morphological integration (the title of their book) as “the summation of the totality of characters which, in their interdependency of form, produce an organism.” While traditional evolutionary biology had focused (rightly so) on the mean and how it changes over time and across populations, Olson and Miller focused on correlations among traits, attempting to associate $P$-sets (groups of correlated traits) with $F$-sets (groups of traits with a common function). This task was taken up again by Cheverud (1982), who described how selection may act to integrate traits over time. At the same time, evolutionary theory was becoming explicitly multivariate, and empiricists received the statistical tools to examine how selection acted on multiple traits in nature (Lande 1979; Lande and Arnold 1983). The study of integration appears to be experiencing somewhat of a renaissance today, with the appearance of a recent edited volume (Pigliucci and Preston 2004), the continuing integration of developmental and evolutionary biology, and publication of synthetic works such as West-Eberhard’s (2003) tome.

Phenotypic integration may be approached from both externalist and internalist perspectives (Schwenk and Wagner 2004). The former focuses on how forces outside the organism—primarily natural selection—act to favor certain phenotypic arrangements. The latter focuses on the processes within the organism—development and physiology—that underlie phenotypic expression on a shorter time scale. These viewpoints are often dichotomized into “ultimate” and “proximate” causes. We are often issued warnings
about confusing the two levels of analysis. Although the dichotomization of ultimate and proximate may prevent certain logical errors, it is important not to let it prevent us from asking interesting scientific questions. We often obtain a much richer understanding of evolution when we ask questions that blend the ultimate and the proximate, the external and the internal.

This dissertation is an attempt to examine phenotypic integration from both the outside and the inside. Throughout, I maintain a strong focus on correlational selection, which is thought to be the primary evolutionary mechanism by which adaptive suites of co-expressed traits are assembled. However, my work also strongly relies on hormones, which often serve as the physiological glue keeping such suites together.

In Chapter 1, I present a literature review that explores the interaction of the ultimate and the proximate. I ask how selection may be involved in shaping suites of hormonally mediated traits, and in turn, how hormonal mediation may influence the course of adaptive evolution.

Beginning in Chapter 2, I describe studies of the integration of the “mating phenotype” in a common songbird, the dark-eyed junco. One of the most important aspects of the life history of almost every multicellular organism is finding a mate. Two facts about sexual organisms give rise to sexual selection, which is among the strongest evolutionary forces: males can often increase their fitness immensely by mating more than once, and males cannot be equally successful in this pursuit. In fact, because every offspring produced has but a single father, the more successful some males are, the more unsuccessful males there must be (Shuster and Wade 2003). Sexual selection leads to the evolution of the morphological, physiological, and behavioral traits that constitute the
mating phenotype, and which are often elaborate. Moreover, selection may favor
different traits for different males. This should give rise to interesting patterns of
variation and covariation among traits, making the study of the phenotypic integration of
sexually selected traits especially rewarding.

In Chapter 2, I use long-term data from a natural pedigree to measure genetic
relationships between morphological traits that are expected to be sexually selected,
plumage and body size. I found that a plumage ornament used in male-male competition
and female choice, the white patch on the tail ("tail white"), was genetically correlated
with body size. That is, the two traits were genetically integrated. Further, I used DNA
paternity data to show that correlational sexual selection acted on the two traits,
suggesting that selection may act to maintain this integration.

In Chapter 3, I continued to study tail white, while adding a focus on natural
variation in testosterone that extends into the final two chapters. A long-term study by
Ellen Ketterson, Val Nolan, and colleagues has shown that testosterone is likely to be a
central component of the mating phenotype. Experimentally increasing a male’s
testosterone levels caused him to increase his mating effort—song, mate searching, etc.—
at the expense of parental effort and survival. Sexual selection predicts that high quality
males should have larger sexual signals and should produce higher levels of mating
effort, because they can better afford the survival costs and reap the fecundity benefits of
such an investment. In accordance with this prediction, I found that males with more tail
white were able to produce higher testosterone levels. Interestingly, however, this
relationship was not with "baseline" (constitutively circulating) levels of testosterone, but
with the magnitude of testosterone increases produced in response to an injection of
GnRH, a hypothalamic hormone. In the same study, I showed that such changes were also correlated with a male’s hormonal response to a territorial intruder, suggesting that attractive males possess greater “androgen responsiveness.” Many songbirds produce bursts of testosterone in response to social stimuli that coincide with episodes of territorial or courtship behavior. These results suggested that flexible androgen responsiveness could underlie variation of behavior within the mating phenotype.

I tested this idea in Chapter 4 by relating variation in androgen responsiveness to natural variation in relevant behavior. I found that males that produced higher absolute levels of testosterone in response to GnRH also showed more aggressive behavior in response to territorial intruders. Additionally, the change in testosterone induced by GnRH was negatively correlated with the intensity of parental behavior. These results suggested that androgen responsiveness was indeed associated with relative investment in mating effort and parental effort. Combined with the results of Chapter 3, this study indicated that tail white may act as a reliable signal of a male’s behavioral tendencies, and perhaps, his quality.

In Chapter 5, I began to explore the fitness consequences of physiological and behavioral variation within the mating phenotype. That is, how does selection shape the variation in the proximate mechanism that appears to underlie the integration? I found that males with very high or very low androgen responsiveness were less likely to survive until the next breeding season. In other words, stabilizing selection acted on the mechanism found to be associated with the trade-off between mating effort and parental effort. This result may have arisen because both mating effort and parental effort are energetically expensive, and thus may impose survival costs. Survival is but a single
component of fitness, and I have not yet been able to measure sexual selection (which acts through differences in mating success, not survival). Sexual selection is probably the most important selective mechanism shaping the mating phenotype, so there is obviously much more work to be done.

Nevertheless, I hope the studies present here represent a small step forward in understanding how the mating phenotype evolves, and in general, how ultimate and proximate factors interact in evolution.
Abstract

Joel W. McGlothlin

PHENOTYPIC INTEGRATION OF SEXUALLY SELECTED TRAITS IN A SONGBIRD

Natural selection favors traits that fit not only the external environment, but also the internal environment of the organism. As a consequence, traits often show a pattern of correlation, or phenotypic integration. In this dissertation, I examined both the evolutionary processes and the physiological mechanisms that generate phenotypic integration. I studied a natural population of a songbird, the dark-eyed junco (Junco hyemalis), focusing on the male “mating phenotype,” the suite of morphology, physiology, and behavior used to attract and compete for mates. In Chapter 1, I review literature suggesting that correlational selection, which occurs when traits interact in their effects on fitness, may have effects on the physiological mechanisms that underlie integrated suites of traits. In Chapter 2, I found that correlational sexual selection favored an association between body size and a white patch on the tail feathers (“tail white”), an ornament used both in courtship and male-male competition. I also found that body size and tail white were genetically correlated. These results suggest that correlational selection may maintain the integration of the two traits. In Chapters 3-5, I focus on the role of the steroid hormone testosterone in the mating phenotype. In Chapter 3, I measured natural variation in testosterone levels and found that more attractive males had higher androgen responsiveness. That is, males with more tail white produced more
Abstract
testosterone in response to an injection of GnRH, a hypothalamic hormone. This suggests
that investment in mating behavior (which seems to be controlled by testosterone) may
covary with attractiveness. Indeed, in Chapter 4, I found that androgen responsiveness
naturally covaries with both mating and parental behavior. Males that produced more
testosterone defended their territories more vigorously and fed their offspring less often.
Finally, in Chapter 5, I examined how selection acts on androgen responsiveness, and
found that males with very high or very low responsiveness were less likely to survive.
Combined, these studies suggest that testosterone, on a physiological level, and
correlational selection, on an evolutionary level, act as integrators of the male mating
phenotype.
# Table of Contents

Preface ........................................................................................................... ix
Abstract ........................................................................................................ xiv

Chapter One ..................................................................................................... 1
   Hormones and the continuum between adaptation and constraint

Chapter Two .................................................................................................. 34
   Correlational selection leads to genetic integration of body size and an attractive plumage trait in dark-eyed juncos

Chapter Three ................................................................................................ 72
   Hormones and honest signals: males with larger ornaments elevate testosterone more when challenged

Chapter Four ................................................................................................102
   Natural variation in a testosterone-mediated trade-off between mating effort and parental effort

Chapter Five .................................................................................................134
   Stabilizing selection on natural variation in androgen responsiveness

Conclusions and future directions ...............................................................154
References ....................................................................................................165
CHAPTER ONE

Hormones and the continuum
between adaptation and constraint

JOEL W. MCGLOTHLIN
ELLEN D. KETTERSON

Submitted to Philosophical Transactions of the Royal Society of London B
Summary

Hormones have been proposed both to facilitate and to constrain adaptive evolution because of their role in mediating and coordinating the expression of suites of correlated traits. Selection that leads to a change in the hormone signal may lead to simultaneous changes in the expression of multiple aspects of the hormonally regulated suite, which may be either adaptive or maladaptive. The field of evolutionary quantitative genetics has developed both theory and empirical methods for studying the evolution of correlated traits. Here we argue that the application of quantitative genetics to the study of hormonally mediated suites may prove fruitful. We first briefly review the evidence for the evolutionary potential of hormonally mediated suites. We then examine how selection may shape their evolution and argue that correlational selection, which is defined as selection that arises when traits interact in their effects on fitness, may have both constructive and conservative roles. Finally, we present examples from our own work on dark-eyed juncos that explore both ends of the adaptation-constraint continuum. We explore the potential roles of correlational selection in adaptive integration of the male mating phenotype, and the role of female testosterone as a potential constraint on the independent evolution of the sexes. We suggest that future work on hormonally mediated suites that is motivated by quantitative genetic theory may provide insight into the dual role of such suites as products of and constraints on adaptive evolution.
Correlations among traits are ubiquitous in nature. Darwin (1859) noted the generality of “correlations of growth,” speculating that when changes in one character are “accumulated through natural selection, other parts become modified” (p. 143). Lande (1979; Lande and Arnold 1983) incorporated this observation into a general mathematical model of the evolution of correlated traits in response to natural selection. Lande’s “multivariate breeders’ equation,” $\Delta \mathbf{z} = \mathbf{G} \beta$, describes how the evolutionary response of an interrelated set of traits ($\Delta \mathbf{z}$, a vector) to directional selection ($\mathbf{\beta}$, a vector of “selection gradients”) is influenced both by the amount of additive genetic variance in a particular focal trait (diagonal elements of the matrix $\mathbf{G}$) and by additive genetic covariance among traits (off-diagonal elements of $\mathbf{G}$). Additive genetic variance represents the amount of variation in a trait that is heritable, and is often presented in its standardized form, heritability (which ranges from 0 to 1). Additive genetic covariance represents the degree to which traits are co-inherited due to pleiotropy, the common effect of loci on multiple traits, and/or linkage disequilibrium, the physical or statistical linkage of loci affecting one trait to loci affecting a second trait (Lande 1980a, 1984), and is also often written in its standardized form, the genetic correlation.

Genetic correlations can cause multiple traits to respond to selection as a unit, and depending on the direction of natural selection, they may accelerate or impede its effects. The latter effect has received perhaps the most attention, and genetic correlations among traits are often seen as measures of genetic constraint on phenotypic evolution (Cheverud 1984; Maynard Smith et al. 1985; Arnold 1992). Certain patterns of genetic covariance...
Hormones, adaptation, and constraint

may slow or prevent a population’s response to selection or may lead to maladaptive changes in correlated traits. This constraining effect is expected to be strongest when the genetic correlation between two traits is perfect (that is, approximately ±1), in which case independent response to selection may be completely restricted. However, certain patterns of weak genetic correlations among multiple traits may constrain evolutionary response just as strongly (Blows and Hoffmann 2005).

Interestingly, genetic correlations may be seen as an outcome of adaptive evolution as well a hindrance to it, i.e. as adaptations as well as constraints (Merilä and Björklund 2004). Patterns of covariation across multiple traits may arise because selection favors their integration to perform a common function (Olson and Miller 1958; Cheverud 1982). Theory predicts that selection, particularly a form known as correlational selection, should be able to change the magnitude of genetic covariances over time (Lande 1980a; Lande and Arnold 1983; Phillips and Arnold 1989; Rice 2000; Sinervo and Svensson 2002; Jones et al. 2003; Phillips and McGuigan 2006). Correlational selection describes situations in which the fitness effects of two traits are interactive (or epistatic) rather than additive; that is, the selective advantage of one trait depends on the value of another. When correlational selection favors a certain combination of traits, the genetic covariance between them is predicted to increase over time (Lande 1980a; Phillips and Arnold 1989; Phillips and McGuigan 2006). Indeed, correlational selection has been shown to coincide with existing genetic correlations in a number of cases, as one would predict if correlational selection in the past accounted for the present existence of the genetic correlations (Chapter 2, Brodie 1989, 1992; Conner and Via 1993; Morgan and Conner 2001, but see Blows et al. 2004).
Recent work in evolutionary genetics promises to shed light on the relative importance of genetic covariances as constraints and adaptations. Specifically, a number of studies have focused on comparing the pattern of genetic covariance, as represented by $G$, across related populations or species to determine whether and how $G$ evolves (reviewed in Steppan et al. 2002). Cases of both stability (usually within species) and lability (usually among species) have been reported. This pattern suggests that on the short term, $G$ may act as a constraint and/or that correlational selection may act conservatively to stabilize $G$. However, over longer periods of time, $G$ may evolve as selection pressures change. Importantly, the extent to which a genetic correlation persists as a constraint, and the extent to which it may be modified by selection, is likely to depend on its genetic architecture (Agrawal et al. 2001). Pleiotropy is predicted to lead to stricter constraints on evolution than linkage disequilibrium, which is subject to erosion by recombination (Lande 1980a, 1984). Correlational selection may lead to an increase in both linkage disequilibrium and pleiotropy, depending on the genetic variation that is available. Currently, we have little information about the relative importance of pleiotropy and linkage for the architecture of genetic correlations in natural populations (Hawthorne and Via 2001; Conner 2002).

Because hormones coordinate the expression of multiple aspects of physiology, morphology, and behavior, they are often considered physiological analogues of genes with pleiotropic effects (Finch and Rose 1995; Ketterson and Nolan 1999). Hormones may often serve as a link between an organism’s environment and the expression of an appropriate phenotype. We refer to a set of interconnections among environment, hormones, and phenotypic expression as a hormonal system. Hormonal systems often
involve a great deal of complexity, and many aspects of such systems are likely to vary among individuals. Individuals may differ in the rate of hormone synthesis, release, and degradation, leading to variation in circulating hormone levels. There may be variation in how a hormonal system responds to an environmental stimulus as well as how hormones interact with each other or with target tissues to cause phenotypic effects.

To the extent that it is associated with genetic variation, a hormonal system may create pleiotropy between traits that might otherwise be genetically independent. Like genetic correlations, mechanistic correlations arising from common hormonal mediation of multiple aspects of the phenotype may be seen both as adaptive products of selection and as constraints on future evolution. One of the most obvious cases of common hormonal mediation as an adaptation is when hormones coordinate the simultaneous expression of traits at a transition between life-history stages (Wingfield et al. 2000). At the beginning of the breeding season, for example, seasonally reproducing animals must simultaneously increase the production of gametes, grow primary or secondary sexual structures, and produce courtship behavior. Mediation by a common hormonal system often facilitates the concurrent expression of these changes. Other examples of adaptive hormonal coordination include the regulation of metamorphosis and phenotypic plasticity (Robinson 1992; Denver et al. 2002).

The potential for hormonal systems to have a dual nature as adaptations and constraints may be seen most clearly in their roles as physiological mediators of life history trade-offs, or negative associations between components of fitness (Ketterson and Nolan 1992; Stearns 1992; Sinervo and Svensson 1998; Zera and Harshman 2001; Ricklefs and Wikelski 2002). In some cases, hormonal mediation of a trade-off may be an

Functional constraints may derive from the laws of physics or chemistry, or the facts that an organism cannot be in two places at the same time and has limited energy to allocate among activities. In this case, certain combinations of traits may be impossible, and hormonal mediation may have arisen as an adaptation to permit optimal allocation to competing traits. For example, the glucocorticoid stress response has been proposed as a mechanism that allows vertebrates to optimally allocate limited energy between current and future reproduction (Ricklefs and Wikelski 2002; Heidinger et al. 2006). However, even when common hormonal mediation has arisen as an adaptive solution to a trade-off, it may act as a genetic constraint to an evolutionary response to changing selection pressures. Even if selection favors (and the laws of physics and chemistry allow) a change in the way traits trade off, common hormonal mediation may limit the variation upon which selection may act. If hormonally mediated suites cannot evolve when selection pressures change, they may act as true genetic constraints, limiting adaptive evolution by maintaining maladaptive correlations between traits. To use our earlier example, an novel environment may induce a stress response, triggering individuals to allocate energy towards survival and delay reproduction. However, early reproduction may be imperative in order for a population to persist in this new environment. By restricting the combinations of traits available to selection, the glucocorticoid system may limit this population’s ability to adapt quickly. Whether hormonal systems have such a constraining role in natural populations is an important question that deserves greater study.
In this paper, we explore how hormonally regulated suites of traits are likely to evolve, focusing on the interplay between adaptation and constraint. The definitions of both of these terms are many and varied. For the purposes of our discussion, we define adaptations as the products of past selection for a given function that result in closer matches between the organism and its environment (Gould and Vrba 1982). Our consideration of constraints is primarily in the genetic sense, i.e. patterns of genetic covariance that may prevent or delay an adaptive response to selection (Arnold 1992). First, we briefly review the evidence that hormonally mediated suites of traits can and do evolve, which suggests that constraints imposed by hormones may be evolutionarily labile. Second, we examine selective mechanisms that may affect the evolution of hormonally mediated suites. We suggest that the origin and maintenance of hormonally mediated suites may often be favored by correlational selection, and that changes in the strength and direction of correlational selection should lead to changes in these suites. Finally, we present examples from our own work on dark-eyed juncos that explore both ends of the adaptation-constraint continuum.

**Can hormonally mediated suites evolve?**

When we take a broad view of hormonally mediated suites of traits, there is an abundance of evidence that these suites can and do evolve. Adkins-Regan (2005) presents a comprehensive review of the interplay between hormones and behavior, considering the topic from multiple levels of analysis ranging from biochemistry to phylogeny, and we
Hormones, adaptation, and constraint

will address this topic only briefly. Like much of what is known about the workings of nature, the evidence shows that hormonal systems are an amalgam of conservation and divergence. On the side of conservation, hormones often have identical or very similar structures across taxa. Further, the basic structures of many hormonal systems, such as the hypothalamo-pituitary-gonadal (HPG) axis, which regulates the production of sex steroids, and the hypothalamo-pituitary-adrenal (HPA) axis, which regulates the glucocorticoid stress response, are remarkably conserved across vertebrates. Similarly, while regulation of “maleness” by androgens and “femaleness” by estrogens are stereotypes, they are stereotypes for a reason. These sex steroids play a significant role in the development and evolution of behavioral and morphological differences between the sexes, again in most vertebrates.

On the side of divergence, and thus evolutionary lability, of hormonal systems, there is clearly variation among taxa in the hormonal mediation of suites of traits. A hormone that commonly mediates a particular trait may come to regulate different traits or different combinations of traits in different species. For example, testosterone is often associated with a trade-off between mating effort and parental effort (Ketterson and Nolan 1992, 1994). If males of many species (particularly the well-studied songbirds) are given a dose of testosterone, they spend more time singing or searching for mates and less time feeding the offspring they have already sired. This relationship does not hold true for all species, however. Male three-spined sticklebacks (Gasterosteus aculeatus) show decreased 11-ketotestosterone during parental care, but experimental enhancement has no effect on parental care (Páll et al. 2002a, b). Testosterone may even enhance parental behavior in some species, including the California mouse (Peromyscus
In a review of the effects of testosterone on male behavior in birds, Hau (2007) shows that the negative relationship between testosterone and male parental care does not even hold in all songbirds. For example, testosterone does not suppress parental behavior in chestnut-collared longspurs (*Calcarius ornatus*) and great tits (*Parus major*) (Lynn et al. 2002; Van Duyse et al. 2002).

Hau (2007) proposes two hypotheses for the relationship between suites of hormonally mediated traits (hereafter hormonally mediated suites) and adaptive evolution. Hormonal mediation may be more likely to constrain evolution if the components of the system (hormone, receptors, etc.) are tightly linked and cannot be disentangled by selection (the “evolutionary constraint hypothesis”). If the components of the system are largely independent of each other, then the suite may be free to evolve in response to changing selection pressures (the “evolutionary potential hypothesis”). So far, comparative evidence seems to favor the latter hypothesis, but as Hau emphasizes, both may come into play, and the prevailing hypothesis may depend on a number of factors, including when and where the hormone acts. Future comparative work, both across higher taxa and closely related species that differ ecologically should provide evidence as to how hormonally mediated suites evolve.

One limitation of comparative studies is that they necessarily include only evolutionary survivors. Populations that have become extinct, perhaps because their ability to respond to a new selection pressure was constrained, cannot be observed. Therefore, studies within a species or population are a necessary complement to comparative work. Such studies can help us understand the microevolutionary processes...
that may be responsible for the patterns seen in comparative studies. Work examining variation among individuals in hormonal systems, and thus their ability to respond to selection, is in its infancy. There are some cases in which variation in hormonally mediated traits and components of hormonal systems (e.g., plasma levels of hormone) have been shown to be heritable, but it is unclear whether hormonal systems should be expected to commonly respond to selection on the traits they regulate (Adkins-Regan 2005). Hormonal systems seem to be sufficiently complex in their biochemistry and gene regulation that genetic variation, and hence the ability to respond to selection, should be substantial. Indeed, a few recent studies have shown that hormonal systems may respond to selection on hormonally mediated traits (Adkins-Regan 2005). In birds, for example, selection for personality in great tits has led to evolution of the hypothalamo-pituitary-adrenal axis that coordinates the stress response (Carere et al. 2003; Drent et al. 2003). Similarly, artificial selection directly on a hormonal system (the glucocorticoid stress response) in rainbow trout (Oncorhynchus mykiss) has led to concomitant behavioral and physiological changes (Øverli et al. 2005).

**How do hormonally mediated suites evolve?**

Despite the mounting evidence that hormonal systems can and do evolve, few studies have examined the microevolutionary processes at work in natural populations. In this section, we employ the analogy between hormonally mediated suites and genetic covariation and use the concept of phenotypic selection from evolutionary quantitative
genetics (Lande and Arnold 1983) to provide a framework in which to pursue such studies.

Natural selection acts directly on phenotypes, and leads indirectly to genetic evolutionary change. The evolution of hormonally mediated suites should occur in an analogous way, with selection acting directly on the traits mediated by the hormone and leading to evolutionary change in hormonal systems and the genes that underlie them. Thus, we consider selection acting on the traits that make up hormonally mediated suites—the phenotypes that directly interact with the biotic and abiotic environment—rather than on the components of the hormonal system itself.

The effects of selection on hormonally mediated suites, and the extent to which hormonal regulation acts as a constraint, is likely to depend on the mode of selection (the shape of the relationship between phenotype and fitness) as well its strength and consistency (whether it fluctuates or continues in the same direction for many generations). Directional selection is the most commonly studied mode of selection. It seems to be relatively common in natural populations as well, likely accounting for much adaptive evolutionary change (Hoekstra et al. 2001; Kingsolver et al. 2001; Rieseberg et al. 2002; Estes and Arnold 2007). Directional selection acting via differential survival is expected to be strongest when the environment changes (e.g. during a colonization event) and to weaken as a fitness optimum is reached (Arnold et al. 2001). Directional sexual selection, which operates through variation in mating success, may be open-ended and consistently strong (Shuster and Wade 2003). The extent to which $G$ constrains the response to directional selection depends on the direction of the vector of selection gradients $\boldsymbol{\beta}$. The response to directional selection is predicted to be strongest when $\boldsymbol{\beta}$ is
aligned with the direction of maximal genetic variance ($g_{\text{max}}$, as measured by the dominant eigenvector of $G$, and to be most constrained when $\beta$ points in a direction with little genetic variance (Schluter 1996; Blows and Hoffmann 2005). Simplistically, adaptive evolution should be most rapid when selection acts on combinations of traits that, collectively, have the most genetic variance.

This is easiest to envision in the two-dimensional case (Figure 1.1). Genetic correlations should facilitate evolution when the direction of selection is aligned with the major axis of the correlation between two traits. In the case of two positively correlated traits, selection should be most efficient when it favors a simultaneous increase in both traits. Returning to the analogy of a hormonally mediated suite, selection could lead to simultaneous evolution of two traits mediated in the same way by a common hormone by favoring an increase in the circulating signal or a coordinated change in receptor expression. A hypothetical example might be selection for greater repertoire size and song frequency in a songbird, traits that might require more highly developed vocal control nuclei and greater development of the syringeal musculature. Both of these targets are androgen-sensitive, and selection might lead to an increase in testosterone levels or androgen receptor expression. An evolutionary response should be most constrained when selection is aligned with the minor axis of the correlation (Figure 1.1). For two hormonally mediated traits that are positively related, this would occur when selection favored an increase in one trait and a decrease in the other. Hypothetically, a response to conflicting selection pressures to increase repertoire size (favored by open-
Figure 1.1. Schematic representation of the effect of genetic covariance (as represented by the off-diagonal element of $G$) on selection response. Response to selection acting in the same direction as the correlation, i.e. on the major axis of the correlation $(a)$ is facilitated by genetic covariance. Response to selection acting in opposite directions, i.e. on the minor axis $(b)$ is constrained. Intensity of directional selection is the median value from studies of natural populations (Kingsolver et al. 2001). Elements of $G$ are chosen arbitrarily.

ended female choice) and to decrease song frequency (favored by predation) may be slowed by common dependence on androgens.

Artificial selection experiments on hormonally mediated suites have been similar to selection on the major axis; they have selected on a correlated suite of traits such as “personality” (Adkins-Regan 2005). This sort of selection may be common in the wild. For example, increased male mating effort—courtship, territorial aggression, etc.—
may be favored when the opportunity for sexual selection (Shuster and Wade 2003) increases, leading to coevolution of the traits mediated by testosterone. Selection along the minor axis to uncouple correlations ought to be constrained by hormonal mediation, assuming that the pattern of hormonal mediation remains stable. In a stable population, where the direction and intensity of selection may fluctuate, this resistance to uncoupling may hold. That is, hormonal mediation may have a stabilizing role in the face of such short-term selection associated with variable environments. However, we predict that when such selection is strong and consistent, as might be expected when a population colonizes a new environment, their power as constraints may be overcome because the correlations themselves may evolve.

To understand how this may occur, we must consider correlational selection, which as described above, acts when traits interact epistatically to affect fitness; that is, the effect of one trait on fitness depends on its co-expression with another trait (Lande and Arnold 1983). Correlational selection has rarely been detected in natural populations, probably because of the paucity of studies that have set out to measure it (and the large sample size required) (Kingsolver et al. 2001). Despite the lack of measurements, Kingsolver et al. (2001) suggest that it may be the most common mode of selection. The classic example of correlational selection comes from an experiment on garter snakes (Thamnophis ordinoides) by Brodie (1992). After releasing juvenile garter snakes into the wild, Brodie found that those that exhibited certain combinations of anti-predator behavior and color pattern had higher survival rates. Specifically, spotted snakes were more likely to survive if they performed evasive behavior, whereas striped snakes
survived better if they did not perform the behavior. In short, correlational selection favored a negative relationship between the two traits.

Correlational selection can cause evolutionary change in the genetic covariance between traits. The change in the $G$ matrix in one generation due to selection may be predicted using the equation $\Delta G = G(\gamma - \beta \beta^T)G$, where $\gamma$ is a matrix representing non-linear selection (quadratic—stabilizing and disruptive—selection is on the diagonal, and correlational selection is on the off-diagonal), $\beta$ is the vector of directional selection gradients, and $T$ denotes matrix transposition (Phillips and Arnold 1989, derived from Lande 1980). This equation is mathematically complicated, but certain generalizations can be made. Directional and negative quadratic (stabilizing) selection tend to degrade genetic variances and covariances, while positive quadratic (disruptive) selection tends to build them up. Correlational selection has the most direct effect on genetic covariance. When correlational selection is positive, genetic covariance tends to increase, whereas negative correlational selection causes covariance to decrease (or become more negative). Under certain assumptions, and adding estimates of mutation (which may either increase or decrease covariance) and recombination (which decreases covariance), these conclusions can be extended to across-generation changes (Jones et al. 2003; Phillips and McGuigan 2006).

In the case of garter snakes, negative correlational selection coincided with a negative genetic correlation between color pattern and behavior, suggesting a causal link between the two (Brodie 1989, 1992). Similar patterns of correlational selection and genetic correlations have been detected for floral morphology in wild radish ($Raphanus raphanistrum$) (Conner and Via 1993; Morgan and Conner 2001) and ornamental
plumage and body size in dark-eyed juncos (*Junco hyemalis*) (Chapter 2). Correlational selection has also been implicated in the evolution of female polymorphism in side-blotched lizards (*Uta stansburiana*) (Svensson et al. 2001; Sinervo and Svensson 2002). These studies suggest that correlational selection may be a common ultimate cause of genetic covariance in nature.

Applying this approach to hormonally mediated suites, we predict that correlational selection should also be important in the evolution of the composition of such suites. When correlational selection acts positively by favoring variants with particular combinations of traits, it inherently generates linkage disequilibrium—statistical association between loci—by differential survival or reproduction of individuals with the “right” combinations of alleles. However, it also favors alleles that generate pleiotropy between the traits under selection, which provides more stability in the face of recombination. To the extent that these pleiotropic variants are generated by a hormonal system, the association of traits within a hormonally mediated suite may be strengthened or weakened by correlational selection. Stated another way, correlational selection arising from the effects of coordinated co-expression of traits may have effects on the co-sensitivity of target tissues to a hormonal signal by favoring or disfavoring coordinated expression of traits.

Rice (2000) presents a mathematical and visual model that shows how correlational selection can alter developmental relationships, increasing the integration of the phenotype by associating traits with common “underlying factors.” Thus, if those factors are hormones (or other aspects of a hormonal system), and correlational selection acts consistently enough, it may favor the addition of traits to hormonally mediated
suites. Because of the complexity of hormonal systems, evolutionary change may occur at many different steps along the hormonal pathway (Nijhout 2003). The most likely type of pleiotropic mutation may be one that affects the expression of hormone receptors. Hormones have no effect on a tissue unless it expresses the appropriate receptor, and the expression of a receptor in a novel location or at a novel timing could allow a trait to be co-opted into a hormonally regulated suite. Co-option of existing physiological mechanisms is probably common because it requires fewer evolutionary steps than building a new pathway de novo. Nijhout (2003) suggests such a scenario for the evolution of horn polymorphism in Onthophagus beetles.

Other forces, such as directional selection, univariate stabilizing selection, recombination, and mutation may break up correlations among traits. Although hormonal mediation (and pleiotropy in general) may act as a buffer against such change, especially on the short term, the pattern of covariance among traits may change significantly if these forces act consistently over a long period of time. Correlational selection, as the multivariate analogue of stabilizing selection, is expected to provide stability to groups of correlated traits when it is aligned with G (Blows and Brooks 2003; Estes and Arnold 2007). A recent study has shown that divergence of single traits can best be explained by a model that assumes long-term stabilizing selection combined with occasional movements of optima, which together create a pattern of stasis punctuated by occasional divergent evolution (Estes and Arnold 2007). Estes & Arnold suggest that a similar mechanism, acting via correlational selection, may be responsible for the long-term maintenance of phenotypic integration.
By analogy, we expect that once hormonally mediated suites have evolved, owing perhaps to correlational selection, they should also be reinforced over the long term by correlational selection, but that traits may be lost or gained when the selective landscape changes. Figure 1.2 illustrates this idea. (The graphs are fitness surfaces, which are three-dimensional representations of selection in a population [Brodie et al. 1995]. Trait combinations with high fitness \( w \), plotted on the vertical axis) are favored by selection. The surfaces are generated using selection gradients, which, when measured in a natural population, represent the best quadratic fit between traits and fitness.) In Figure 2.2a, individuals that have high values of two traits have the higher fitness, and the traits’ effects on fitness interact; that is, they are under both directional and correlational selection, creating a rising fitness ridge. Proximately, the correlation between the traits is generated by a common hormonal mechanism, while ultimately this hormonal mechanism is maintained by correlational selection, which acts to strongly disfavor variants that lead to a loss of hormonal regulation of one of the traits (e.g. by turning off the expression of the receptor in a specific tissue). Figures 2.2b & 2.2c show how traits may be decoupled from a hormonal suite. Imagine the population has colonized a new environment where one of the traits is now disfavored. Correlational selection is also reduced, because the two traits are no longer favorable when expressed together. In this case, the genetic covariance between the traits is predicted to decrease. Mutations that decouple one of the traits from hormonal regulation may now invade, and the composition of the hormonal suite may change in a number of generations. Epistatic fitness effects may also arise in the new environment, causing negative correlational selection, which would further accelerate the loss of a trait from hormonal mediation.
Figure 1.2. Hypothetical example of the effect of correlational selection on hormonally regulated trait suites. The plots at left are individual fitness surfaces, with two traits, \( z_1 \) and \( z_2 \), on the horizontal axes and relative fitness, \( w \), on the vertical axis (Brodie et al. 1995). In (a), the two traits are regulated by a common hormone, as represented by the circle at the right. As in Figure 2.1, each has a genetic variance \( G = 0.5 \), and the genetic covariance is \( G_{12} = 0.75 \). The fitness surface shows natural selection on the two traits. Both traits are under moderate directional selection, \( \beta = 0.16 \), and stabilizing selection, \( \gamma = 0.1 \) (medians from Kingsolver et al. 2001), and are affected by relatively strong correlational selection, \( \gamma_{12} = 0.3 \). Using the equation \( \Delta G = G(\gamma - \beta \beta^T) \), this selective regime is predicted to maintain the correlation between the traits, and hence, their common hormonal basis. In (b), the direction of selection on \( z_2 \) is reversed (\( \beta = -0.16 \)), and there is no correlational selection. Within one generation, a decrease in the genetic correlation is predicted. Over several generations, \( z_2 \) may become disassociated from hormonal regulation. In (c), negative correlational selection (\( \gamma_{12} = -0.3 \)) also occurs, accelerating the dissociation of \( z_2 \).
Rice (2000) shows that such selection for “deintegration” is possible in theory when considering three or more traits. Such a change in the fitness surface may be responsible for a disassociation of traits from hormonal regulation, such as parental care and testosterone in chestnut-collared longspurs (Lynn et al. 2002).

We discuss this scenario in order to show how a hormonally mediated suite might evolve in response to an environmental change, but we do not mean to imply that all organisms will be able to do so. Some hormonally mediated suites systems may respond readily to changing selection pressures, whereas others will be so central to the organisms’ development that even small changes to their expression would be lethal or clearly inferior. If a change in the latter type of system is necessary to adapt to a new environment, extinction is likely to be the result, making these cases difficult to study.

So far, our discussion has neglected one of the primary properties of hormones—the ability to mediate phenotypic plasticity (Dufty et al. 2002). Traditionally, plasticity is thought to slow genetic evolution because if plastic response to environmental change is large enough, it may eliminate the “need” for an evolutionary response. However, Price et al. (2003) show that intermediate levels of plasticity may actually facilitate evolutionary change (see also West-Eberhard 2003). This is because plasticity allows a population to establish a foothold in a new environment before it becomes extinct. Once a population becomes established, selection may alter the phenotype further. In the process, selection may alter the relationship between the genotype and phenotype, stabilizing plastic changes with genetic ones. Hormonal mechanisms are likely contributors to this process. For example, in some populations of the white-crowned sparrow (*Zonotrichia leucophrys*), males modulate testosterone levels during aggressive behavior, while males
in other populations do not (Wingfield and Hahn 1994; Meddle et al. 2002; Lynn et al. 2007). It is unknown whether these differences represent plastic or genetic changes; however, one might predict that a plastic response (perhaps to decreased competition from other males) may have preceded a genetic one.

Application of quantitative genetics to hormone studies

To date, most of the studies examining potential effects of hormones on adaptation and constraint have used “phenotypic engineering” (e.g. Ketterson and Nolan 1999; Clotfelter et al. 2004; Reed et al. 2006). These studies typically examine the effects of experimentally enhanced hormone levels on behavior or components of fitness. Although these studies are quite powerful in their ability to infer mechanisms of hormone action and allow experimental tests of adaptive (or maladaptive) hypotheses, they provide only limited insight into the evolutionary processes that shape hormonally mediated suites. Phenotypic engineering should be accompanied by studies that are motivated by quantitative genetic theory if we are to understand the evolution of hormonally mediated suites more fully.

One of the first steps is to quantify patterns of natural variation and covariation within the population of interest. Phenotypic engineering studies may show that multiple traits are affected by a hormone, but this does not necessarily translate to natural covariation. Ideally, we would be able to estimate $G$ in order to predict response to selection by hormones and hormonally mediated traits. However, this is often quite
Hormones, adaptation, and constraint
difficult (though not impossible) to achieve in the field, and estimates of phenotypic
variances and covariances ($P$ in quantitative genetic parlance) may provide reasonable
estimates of $G$. Even measuring $P$ may be difficult, however, because both behavior and
hormones are notoriously variable within individuals. Using multiple measurements and
standardized behavioral and physiological assays (such as simulated territorial intrusions,
stress series, and GnRH challenges, see below) may alleviate this problem to some
degree. Promisingly, studies measuring individual (co)variation for multiple behavioral
and hormonal traits are becoming more common (Bell 2007).

Here, we have speculated as to how selection may act on hormonally mediated
suites. There have been few measurements of selection on physiology, behavior, and life
history in the wild, presumably because such traits are more difficult to measure than
morphology (Kingsolver et al. 2001). Measurements of selection using the methods
described by Lande & Arnold (1983; see also Arnold and Wade 1984a; Phillips and
Arnold 1989; Brodie et al. 1995) are needed to test the hypotheses that we have advanced
here. The effort required to obtain such measures is substantial. One needs a fairly large
sample of multiple traits and fitness estimates measured on the same individuals. Larger
sample sizes are required to measure correlational selection than directional selection,
which probably accounts for the paucity of measurements in the literature (Kingsolver et
al. 2001).

In the absence of such large-scale studies, measurements of selection on
individual variation in hormones (as opposed to the entire suite of hormone-mediated
traits) may provide some insight. Blows & Brooks (2003; see also Blows 2007) argue
that correlational selection may be understood by rotation of $\gamma$ to generate measurements
of quadratic (stabilizing or disruptive) selection on linear combinations of traits. This suggests that a measurement of quadratic selection on physiological measurements such as hormone levels, though not providing information about which traits contribute to fitness differences, may indicate the action of correlational selection on the multivariate suite. In one nice example, corticosterone levels were found to be under stabilizing selection in cliff swallows (*Petrochelidon pyrrhonota*) (Brown et al. 2005).

In addition to studies of selection in natural populations, artificial selection studies can provide experimental confirmation of the malleability or rigidity of constraints (Conner 2003; Fuller et al. 2005). Artificial selection experiments in a butterfly (*Bicyclus anynana*) (Frankino et al. 2005) and wild radish (Conner et al. in prep.) have demonstrated selection response despite very strong genetic correlations between morphological characters. In these experiments, selection was applied in both directions along the minor axis of the correlation (the direction in which evolutionary response should be most constrained), creating both large/small and small/large lines. In both cases, the means of both phenotypes evolved independently, but the genetic correlation remained, that is, the intercept but not the slope of the relationship was altered. To our knowledge, no such studies have been attempted on hormonally mediated traits. This might be accomplished by applying simultaneous selection for decreased song rate and increased nestling feeding rate (and vice versa) in a male songbird. Another important avenue of research would be to apply artificial correlational selection to genetically correlated traits in an attempt to strengthen or weaken the correlation. Such studies, which have not yet been carried out for any traits, hormonally mediated or not, should provide insight into the persistence of genetic constraints.
Another interesting approach that may also have a hormonal parallel is to apply artificial selection in one sex and measure a response in the other. Genetic correlations between male and female traits may cause coevolution of the sexes in response to selection on one sex. Lande (1980b) suggested that such correlations may constrain the evolution of sexual dimorphism, but further modeling suggested that this constraint was not absolute (Reeve and Fairbairn 2001). However, when the trait in question trades off with another trait related to fitness, as is so often the case for hormonally mediated traits, such a constraint may be more difficult to surmount. In a dioecious plant (*Silene latifolia*) Delph et al. (2004) showed that selection for larger female flower size in females caused a decrease in flower number in males, and vice versa. These two traits are negatively genetically correlated and represent a life-history trade-off, with flower size more important for female fitness and flower number more important for male fitness (Steven et al. 2007).

**Adaptation and constraint in dark-eyed juncos**

From 1987-2001, Ketterson, Nolan, and colleagues studied the effects of testosterone in a natural population of songbirds, the dark-eyed junco (*Junco hyemalis*) using the methods of “phenotypic engineering” (reviewed in Ketterson and Nolan 1992, 1999; Ketterson et al. 1996, 2001; Reed et al. 2006). At the beginning of the breeding season, subcutaneous implants were used to experimentally elevate testosterone levels in half the breeding males. These “T-males,” which experienced prolonged exposure to breeding-season peak
testosterone levels, could then be compared to controls ("C-males") that received empty implants and thus maintained normal testosterone levels. The general conclusions of these experiments were that experimentally enhanced testosterone levels led to an increase in mating effort, as measured by song, courtship behavior, home range size, and success at obtaining extra-pair fertilizations (Ketterson et al. 1992; Chandler et al. 1994; Enstrom et al. 1997; Raouf et al. 1997; Reed et al. 2006). This increase in mating effort, however, came at a cost to both parental behavior, as measured by nestling feeding rate and nest defense, and self-maintenance, as measured by body mass, immune function, and survival (Ketterson et al. 1991, 1992; Cawthorn et al. 1998; Schoech et al. 1998; Casto et al. 2001; Reed et al. 2006). When summing these results, T-males had higher lifetime fitness than C-males, because the benefits of obtaining more mates outweighed the costs of reduced survival and parental care (Reed et al. 2006). Because males with constitutively elevated testosterone do exist in nature, this result led us to ask two questions. First, does natural variation in male testosterone represent an adaptive resolution of the trade-off between mating effort versus parental effort and survival? Second, to what extent do testosterone levels in females constrain their evolution in males?

As the first step toward answering the first question, we assessed natural variation in male testosterone levels during two breeding seasons (2003-2004). Previous studies of natural variation measured "baseline" levels of testosterone, that is, the amount of testosterone circulating in the plasma of undisturbed individuals. However, in other songbirds (including closely related sparrows) natural testosterone levels may fluctuate rapidly in response to social stimuli such as competing males or potential mates (e.g.
Wingfield 1985; Pinxten et al. 2003). This observation is the foundation of the “challenge hypothesis,” and has been suggested to act as a mechanism that allows males to produce testosterone-mediated behavior when circumstances call for it, while also lessening the costs that would accompany constitutively expressed high testosterone levels (Wingfield et al. 1987, 1990, 2001). Thus, we attempted to measure males’ capacity to produce short-term testosterone increases as well as their undisturbed circulating levels. To do this, we used injections of gonadotropin-releasing hormone (“GnRH challenges”), which stimulated the HPG axis to produce transient increases of plasma testosterone. We found that males were able to produce these short-term testosterone increases across the breeding season, even when feeding nestlings (Jawor et al. 2006). There was great individual variation in the response to GnRH challenges; importantly, however, when males were challenged multiple times across the breeding season, the magnitude of the short-term testosterone increase (that is, the difference between post-challenge and pre-challenge levels) was found to be repeatable (Jawor et al. 2006). Although we have not yet measured heritability due to the difficulty of obtaining a large sample of relatives, this individual consistency suggests that some genetic variance may underlie variation in short-term testosterone increases. This GnRH-challenge response also seems to be ecologically relevant, as the testosterone levels produced by GnRH challenges predicted the levels produced in response to a male territorial intruder (Chapter 3).

Next, we asked to what extent natural variation in testosterone predicted individual variation in behavior, specifically, behavior related to the trade-off between mating effort and parental effort. Such covariation at the individual level is important, because in its absence we cannot use implantation studies to predict how hormonally
mediated suites might evolve. Prior attempts to correlate natural variation in hormones and behavior have provided mixed results, possibly because so many environmental factors may influence both measurements (Adkins-Regan 2005). Therefore, we again used GnRH challenges, allowing us to measure hormone levels in response to standardized physiological stimulus and to estimate the relative importance of short-term increases versus constitutively maintained levels. We also used standardized protocols to elicit behavior. As a measure of mating effort, we assessed territorial aggression in response to simulated territorial intrusions (Wingfield 1985) and found that the peak testosterone levels produced in response to GnRH predicted a male’s level of aggression toward the intruder (Chapter 4). This suggests that males with more responsive HPG axes invest more effort into a component of mating effort (the primary purpose of territorial defense is to maintain an area in which to breed); however, further studies measuring behaviors such as courtship are needed. We also found that males with larger changes in testosterone levels in response to GnRH made fewer trips to the nest during nestling feeding (Chapter 4). In this study, a male’s mate was temporarily removed in order to control for potential interactions between members of a pair. These correlations provide us with compelling evidence that natural variation in testosterone production affects life-history trade-offs in the same way as experimental testosterone elevation. However, because these studies were not always conducted on the same individuals at the same time, more evidence is needed.

The results raise the question of why individual males should vary in the resolution of mating-effort/parental-effort trade-off, and thus in testosterone production, at all. Sexual selection theory provides us with a possible answer. When males vary in the
number of mates they obtain, some males necessarily succeed while others fail; in fact, as variation in mating success increases, the number of males that are “losers” must increase as well (Shuster and Wade 2003). In iteroparous or paternal species, it may benefit males that are unlikely to be successful at obtaining extra mates to divert some of their energy from mating effort to self-maintenance or parental effort. It has been noted that this pattern should give rise to correlational sexual selection acting on attractive signals used in courtship and the resolution of trade-offs involving mating effort ( Getty 1998, 2006). This is because the effects of a sexual signal and mating effort on fitness are expected to be multiplicative, e.g., it is useless for a male to be attractive if he does not attempt to obtain mates, but attractive males that expend considerable effort may be highly successful at obtaining mates. Such selection should associate mechanisms underlying trade-off resolution with attractive signals, potentially maintaining the honesty of the signal (its covariation with male quality) and variation in both the signal and the resolution of the trade-off.

Consistent with such a scenario, we have found covariation between the testosterone response to GnRH-challenge and the size of a white plumage patch on the tail (Chapter 3). This patch, referred to as tail white, is a morphological signal used by females in mate choice (Hill et al. 1999). In our population, males with more tail white show more intense response to GnRH challenges (as measured by the increase above pre-challenge levels). Interestingly, this pattern is more apparent in young males, where decisions about whether to delay mating effort to a later breeding season may be more crucial. To date we have not measured correlational sexual selection acting on tail white and testosterone or testosterone-mediated behaviors, but a consideration of how tail white
is used makes interactive fitness effects likely. A male’s tail white patch is hidden by the gray interior tail feathers at rest, but it becomes visible when a male spreads its tail during courtship displays (See Figure 3.1). In addition to increasing general mating effort, testosterone implants increase the frequency of such displays (Enstrom et al. 1997). Taken together, these findings suggest that testosterone-mediated behavior and tail white may interact to produce frequent, attractive displays that may enhance mating success.

Combined, the results of these studies suggest that variation in the testosterone-mediated suite of behaviors may be associated with male attractiveness. This suite of behaviors may act as an adaptation allowing males to produce optimal levels of mating and parental effort depending on their quality. However, other work from our research group suggests another factor that may affect the evolution of testosterone-mediated suites. Across species, male and female mean testosterone levels tend to be correlated, suggesting that they may have coevolved (Ketterson et al. 2005; Møller et al. 2005; Mank 2007). Within species, a genetic correlation across the sexes could act as a genetic constraint on the evolution of testosterone-mediated traits in males due to a correlated response to selection in females. To study this possibility, we have examined the behavioral and fitness effects of experimentally elevated testosterone in females, and have begun preliminary studies of individual variation in females.

An early study manipulating testosterone levels in captive females found that increased testosterone led to decreased choosiness when females were presented with two males (one T-male and one C-male) (McGlothlin et al. 2004). This suggested a potential constraining effect of females on the evolution of males. Because the benefits of testosterone-mediated traits are related to mating, reduced choosiness by females may
lead to reduced sexual selection on males. A reduction in female fitness might lead to antagonistic selection that creates a constraining effect on males. Studies of captive birds showed that females, like males, have decreased immune function when implanted with testosterone (Zysling et al. 2006). Whereas down-regulation of the immune system by testosterone may be adaptive for males (because it allows them to divert energy to mate acquisition), it may represent a net cost for females (Zuk 1990; Stoehr and Kokko 2006). Also in captive females, testosterone seems to act as an inhibitor of brood patch formation (Clotfelter et al. 2004). In the wild, testosterone implantation seemed to interfere with nest initiation, as time to first egg was longer in T-females (Clotfelter et al. 2004). Incubation consistency and nest defense during the egg stage were unaffected (Clotfelter et al. 2004).

It is not yet clear whether testosterone in females may act as a constraint on the evolution of testosterone-mediated traits in males. The extent to which the sexes are genetically correlated in testosterone production is unknown. Our initial investigation of individual variation indicates that males and females may regulate testosterone in different ways. Whereas males respond to GnRH challenges throughout the breeding season, females seem to do so only when producing eggs (Jawor et al. 2007). This effect has interesting implications for the role of female testosterone in maternal effects, as the magnitude of this GnRH-challenge response showed a strong correlation with testosterone deposited in the yolk, but it suggests that the evolution of the testosterone-mediated suite may be somewhat decoupled across sexes. Further work, especially applying the approaches described above to both sexes, is necessary to determine the importance of cross-sexual interactions in the evolution of hormonally mediated suites.
Conclusion

Although there is a wealth of knowledge about variation in hormone profiles in natural populations and about the impact of experimentally altered hormones on suites of phenotypic characters, we are only just beginning to dissect the mechanisms responsible for evolutionary change in hormonal systems and hormonally mediated suites (Adkins-Regan 2005). Here, we have advocated integrating the methods and theory of quantitative genetics with traditional endocrinological approaches as a promising way to address this issue. When combined with other approaches, including molecular and developmental genetics, the synthesis appears likely to provide important insights as we strive to understand how the inside world of organisms becomes adapted to the outside world.
Acknowledgments

Discussions with other participants in the E-Bird symposia, particularly Elizabeth Adkins-Regan, Creagh Breuner, Michaela Hau, Sharon Lynn, Trevor Price, and John Wingfield were crucial to the development of the ideas in this paper. We thank Marcel Lambrechts, Marcel Visser, Tony Williams, and John Wingfield for their work in organizing the symposia, and the NSF/ESF/NSERC network for providing funds to attend. We also thank Butch Brodie III, Britt Heidinger, Jodie Jawor, Emília Martins, and Dale Sengelaub for helpful discussions and suggestions, and Jeff Conner for providing access to an unpublished manuscript. During the preparation of this paper, JWM was funded by a training grant from the National Institutes of Health (Common Themes in Reproductive Diversity, T32HD049336-01) and a National Science Foundation Doctoral Dissertation Improvement Grant (DEB 0508692), and EDK was funded by a NSF grant (IOB 0519211).
Chapter Two

Correlational selection leads to genetic integration of body size and an attractive plumage trait in dark-eyed juncos

Joel W. McGlothlin

Patricia G. Parker *

Val Nolan Jr.

Ellen D. Ketterson

Evolution 59: 658-671

* Department of Biology, University of Missouri, St. Louis, MO
Summary

When a trait’s effect on fitness depends on its interaction with other traits, the resultant selection is correlational and may lead to the integration of functionally related traits. In relation to sexual selection, when an ornamental trait interacts with phenotypic quality to determine mating success, correlational sexual selection should generate genetic correlations between the ornament and quality, leading to the evolution of honest signals. Despite its potential importance in the evolution of signal honesty, correlational sexual selection has rarely been measured in natural populations. In the dark-eyed junco (*Junco hyemalis*), males with experimentally elevated values of a plumage trait (whiteness in the tail or “tail white”) are more attractive to females and dominant in aggressive encounters over resources. We used restricted maximum-likelihood analysis of a long-term dataset to measure the heritability of tail white and two components of body size (wing length and tail length), as well as genetic correlations between pairs of these traits. We then used multiple regression to assess directional, quadratic, and correlational selection as they acted on tail white and body size via four components of lifetime fitness (juvenile and adult survival, mating success, and fecundity). We found a positive genetic correlation between tail white and body size (as measured by wing length), which indicates past correlational selection. Correlational selection, which was largely due to sexual selection on males, was also found to be currently acting on the same pair of traits. Larger males with whiter tails sired young with more females, most likely due to a combination of female choice, which favors males with whiter tails, and male-male competition, which favors both tail white and larger body size. To our knowledge, this is the first study to show both genetic correlations between sexually selected traits and currently acting correlational sexual selection, and we suggest that correlational sexual selection frequently may be an important mechanism for maintaining the honesty of sexual signals.
Sexual selection is one of the strongest evolutionary forces and may lead to rapid evolutionary change and striking sexual dimorphism (Andersson 1994; Hoekstra et al. 2001; Kingsolver et al. 2001; Shuster and Wade 2003). Sexually selected traits (ornaments) often function as signals of benefits to a potential mate that are either phenotypic or genetic (Andersson 1994; Møller and Alatalo 1999; Møller and Jennions 2001; Kokko et al. 2003). The evolution of such honest signals requires that the optimum value of the trait differs for individuals of differing phenotypic quality (Nur and Hasson 1984; Grafen 1990), which should lead to fitness surfaces that are shaped like rising ridges (Getty 1998). This fitness surface occurs because high quality males with highly developed ornaments will have the highest fitness. Low-quality males with ornaments that are too large will have lower fitness because of factors such as predation or male-male interactions, and high-quality males with ornaments that are too small will have lower fitness because they attract fewer females.

Such a rising fitness ridge is an example of correlational selection, which occurs when a trait’s fitness effect depends on its interaction with another trait (Cheverud 1982, 1984; Lande and Arnold 1983; Phillips and Arnold 1989; Brodie 1992; Schluter and Nychka 1994; Sinervo and Svensson 2002). Positive correlational selection creates linkage disequilibrium and favors covariance due to pleiotropy; over many generations, it can lead to trait integration, or the evolution of common inheritance of functionally related traits (Cheverud 1982, 1984; Lande and Arnold 1983; Brodie 1989, 1992; Phillips and Arnold 1989; Schluter and Nychka 1994; Sinervo and Svensson 2002). Correlational
sexual selection, in particular, may generate genetic correlations between ornamental traits and traits that reliably predict dominance or condition, leading to the evolution of signal honesty (LeBas et al. 2003). Such sexual selection may arise as the result of interactions between intrasexual competition and intersexual choice. A male ornamental trait, for example, may attract the attention not only of females but also of other males (Berglund et al. 1996; Ligon 1999). Male-male interactions may “enforce” the relationship of an ornamental trait with a quality-related trait such as body size because males with attractive signals are repeatedly challenged by other males.

Although many studies have demonstrated that sexual selection acts on multiple characters (Andersson 1994; Candolin 2003), we know very little about the importance of correlational sexual selection in natural populations. The regression-based method developed by Lande and Arnold (1983) is a useful way to measure sexual selection on multiple characters, but few studies using this method have reported measurements of the off-diagonal components of gamma (\( \gamma_{ij} \)), known as correlational selection gradients (Moore 1990; Fairbairn and Preziosi 1996; Rodriguero et al. 2002; LeBas et al. 2003; 2004, reviewed in Kingsolver et al. 2001).

In the dark-eyed junco (Junco hyemalis, Passeriformes: Emberizidae), both males and females vary in the relative size of the area of white found on their otherwise gray outer rectrices ("tail white," Hill et al. 1999; Wolf et al. 2004; Yeh 2004). Males with experimentally enlarged areas of tail white are more attractive to females (Hill et al. 1999); however, such enhancement of females does not affect their attractiveness to males (Wolf et al. 2004). During courtship, male juncos exhibit their tail feathers to females in a display known as “tail spreading” (Enstrom et al. 1997; Nolan et al. 2002).
Tail white is also displayed in dominance contests, which tend to be won by males with whiter tails (Balph et al. 1979; Holberton et al. 1989). Body size, as measured by wing length, is also a predictor of dominance in juncos (Ketterson 1979). If the interaction between female choice and male-male competition has led to the association of tail white with male quality, we should find evidence of the integration of whiteness with body size.

In this study, we examined the evolution of tail white and two components of body size, wing length and tail length. To determine whether selection favors the integration of tail white with body size, we used data from a long-term field study of a natural population of juncos to estimate the strength of correlational (as well as directional and quadratic) selection. We measured selection using four different components of lifetime fitness: juvenile survival, adult survival, mating success, and fecundity, which allowed us to detect the specific episode during which selection for integration occurs and to identify potential conflicting selection pressures (Arnold and Wade 1984a, b; Schluter et al. 1991). Because all three traits are expressed in both sexes, we measured selection separately on males and females; this permitted us to consider whether selection on females might constrain the integration of male traits (Lande and Arnold 1985).

Using a maximum-likelihood pedigree analysis, we estimated the $G$ matrix, which includes measurements of additive genetic variance and covariance (Lynch and Walsh 1998) in order to detect the results of correlational selection (Brodie 1989, 1992, 1993b; Phillips and Arnold 1989; Sinervo and Svensson 2002) and to make evolutionary predictions based on our measurements of selection (Lande 1979; Grant and Grant 1995). We also measured between-sex genetic correlations, to examine the potential for further
evolution of sexual dimorphism (Lande 1980a; Price and Burley 1993; Merilä et al. 1998).

**Methods**

*Study species and general methods*

We studied the Carolina subspecies of the dark-eyed junco (*J. h. carolinensis*), which breeds at high elevations in the Southern Appalachians (Nolan et al. 2002). Juncos are socially monogamous, but extra-pair fertilizations (EPFs) occur commonly (Ketterson et al. 1997; Raouf et al. 1997; Nolan et al. 2002). Females build nests and incubate eggs, and both sexes defend eggs and young against predators and feed nestlings and fledglings. The population used in this study breeds at and around Mountain Lake Biological Station in Giles County, Virginia (37°22'N, 80°32'W).

At the beginning of each breeding season (April and May), we censused the population by capturing birds using mist nets and traps at baited locations that remained the same from year to year. All individuals were marked with aluminum U. S. Fish and Wildlife Service leg bands and a unique combination of plastic color bands. We measured wing length, tail length, and tail white at each capture, and blood samples for DNA parentage analysis were collected once a year. Upon first capture (1987-2000), adult males were implanted subcutaneously with silastic tubes that were either filled with crystalline testosterone (T-males) or left empty (C-males). Effects of testosterone
Correlational selection

treatment on male juncos are reviewed elsewhere (Ketterson and Nolan 1992, 1999; Ketterson et al. 1996). From April to July, we monitored the nesting attempts of all birds on the study area (usually 50–60 pairs). On day 6 after hatching, we marked nestlings with aluminum and color bands and collected a blood sample for DNA parentage analysis. Juncos in this population may raise two (rarely three) broods each summer in the absence of nest predation and usually attempt to renest after nest loss (Nolan et al. 2002). We censused the population a second time at the end of each summer (July and August), capturing adults and newly independent young that had reached adult size (juveniles). The outer rectrices of juveniles tend to have less white than those of adults, and the two age classes are readily distinguishable by their body plumage (Nolan et al. 2002; Wolf et al. 2004; Yeh 2004). We also removed implants from males caught at the end of the summer.

Trait measurement

We measured wing length as the distance from the wrist joint to the tip of the longest primary when the wing was flattened with the thumb, and tail length as the distance between the tip of the longest pair of rectrices and their point of insertion on the body (Pyle et al. 1987). Both measurements were taken with a ruler to an accuracy of 0.5 mm. When multiple measurements existed for an individual bird, we selected those taken by
Figure 2.1. Three outer right rectrices (R6 is the outermost) of a male junco. Tail white values for these feathers are R6 = 0.85, R5 = 0.65, and R4 = 0.15, giving this individual a tail white score of 1.65.

more experienced observers and summarized multiple measurements, using the mode, if any. In the absence of a clear mode, measurements were averaged.

We measured the tail white value of a rectrix as the percentage of its area that was white; an individual’s score was the sum of the tail-white values on the right side of the tail (Figure 3.1). Juncos in our population may have white on two, three, or four of the outer pairs of rectrices, and tail white scores tend to fall between 1 and 3.5 (Wolf et al. 2004). Tail white values were estimated by eye in increments of 5%; values obtained by this method were highly correlated with values obtained using computer image analysis, a more precise method ($r = 0.96, n = 74$; W. L. Wolf, J. M. Casto, E. D. Ketterson, unpublished data). Multiple measurements were summarized separately for each rectrix,
and these values were summed to give the tail white score for each individual. Again, we used the modal score if available.

Because each of the three focal traits may change with age (McGlothlin et al. unpublished data, see also Nolan et al. 2002; Wolf et al. 2004; Yeh 2004), our analysis was based on measurements taken from the juvenile plumage, unless otherwise noted. The wings and tail of the juvenile plumage are retained until the end of the individual’s second summer (its first breeding season, Nolan et al. 2002). Therefore, when measurements from juveniles were unavailable, we used tail white measurements from first-year adults. However, we did not use wing length and tail length measurements from first-year adults because feathers will have been shortened by a year’s abrasion. Individual differences remain consistent across age classes for each trait in both males and females (McGlothlin et al. unpublished data), which is a requirement for making evolutionary predictions for traits that change with age (Brodie 1993a).

Wing length, tail length, and tail white, all show modest sexual dimorphism in juncos (Table 2.1, see also Nolan et al. 2002; Wolf et al. 2004; Yeh 2004). Wing length had the highest loading (0.91) on the first principal component in an analysis that included four components of body size (wing length, tail length, mass, and tarsus length), so wing length was used as a correlate for overall size (McGlothlin et al., unpublished data).
Table 2.1. Descriptive statistics for wing length, tail length, and tail white (see Methods for definitions). Values are calculated from measurements taken from 258 females and 313 males, measured as juveniles. All traits are significantly sexually dimorphic (MANOVA, \( P < 0.001 \)). A sexual dimorphism index (SDI) for each trait was calculated by dividing the male mean by the female mean.

<table>
<thead>
<tr>
<th></th>
<th>MALES</th>
<th>FEMALES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. dev.</td>
</tr>
<tr>
<td>Wing length (mm)</td>
<td>82.2</td>
<td>1.50</td>
</tr>
<tr>
<td>Tail length (mm)</td>
<td>71.3</td>
<td>1.91</td>
</tr>
<tr>
<td>Tail white</td>
<td>2.18</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Identification of the sexes

The sex of juvenile juncos cannot be reliably determined in the field using a single diagnostic measurement, and during the period of this study we did not collect blood samples from juveniles, hence we cannot use genetic markers to assign sex. However, adults caught during breeding may be sexed by the presence of a brood patch (females) or a cloacal protuberance (males), so juveniles caught as adults could be sexed retroactively. In order to assign sex to juveniles that were not recaptured as adults, we created a discriminant function using 571 juveniles (258 females and 313 males) that survived to adulthood. The function
Correlational selection

\[ y = 111.1 \log (\text{wing length}) + 6.3 \log (\text{tail length}) + 1.5 \text{ tail white} - 226.3, \quad (1) \]

which correctly classified 95.3% of these juveniles (cutting point \( y = -0.16 \); Wilks' \( \lambda = 0.262 \)), was used to classify individuals of unknown sex.

Parentage analysis

Parentage analysis was performed using DNA extracted from blood collected from nestlings and adults during the breeding seasons of 1990-1996. Methods used to determine genetic parentage are presented in detail elsewhere (Ketterson et al. 1997; Raouf et al. 1997). Briefly, established multilocus minisatellite DNA fingerprinting methods (Rabenold et al. 1990; Piper and Parker Rabenold 1992) were used to either include or exclude putative parents (i.e., adults whose behavior at or near the nest appeared parental). When this method indicated that one putative parent (in all cases except one, the male) was not the genetic parent of the nestling under consideration, additional gels were run to determine the identity of the extra-pair parent (repeated use of minisatellites for data from 1990-1993, microsatellites for data from 1994-1996).
Estimation of genetic parameters and variance components

**PEDIGREE.** We assembled a pedigree of 643 birds, including juveniles that had been banded as nestlings during 1990-1996 and their genetic parents. There were 611 birds with records for at least one of the three traits (wing length, tail length, and tail white), and 490 had records for all three. Before analysis, 10 individuals without phenotypic records and a single familial link were removed, resulting in a pedigree consisting of 633 birds: 397 progeny, 109 sires, and 155 dams (28 birds were both progeny, and later, parents). In cases of unknown parentage, the female observed at the nest was assumed to be the dam, and the sire was left unknown, resulting in a pedigree that contained more dams than sires.

**GENETIC MODELS.** We used a restricted maximum-likelihood (REML) method (DXMUX procedure of DFREML v. 3.1, Meyer 1998, 2000) to estimate $G$, genetic parameters (heritabilities, $h^2$, and genetic correlations, $r_g$), and their standard errors. The DFREML program uses a derivative-free REML method to estimate additive genetic parameters given a pedigree and individual phenotypic values, while correcting for the influences of fixed or random effects (Meyer 1991, 1998). Unlike conventional ANOVA-based methods, multivariate REML does not require balanced data sets and can account for missing trait values and a complex pedigree (Meyer 1991, 1998). Our model incorporated two fixed effects: sex (to correct for sexual dimorphism) and hormonal treatment of the male associated with the nest that produced the offspring (to correct for potential effects of differential parental investment by T- and C-males, Ketterson et al.)
Correlational selection

1992). Models that incorporated parental effects, effects of shared nest environment, or birth-year effects did not generate significantly different genetic parameters and are not presented here (J. W. McGlothlin, P. G. Parker, V. Nolan, Jr., and E. D. Ketterson, unpublished data). Significance of genetic parameters was tested using two-tailed, one-sample $t$-tests ($H_0: \mu = 0$), with degrees of freedom equal to the number of individuals in the model minus one.

**Between-sex genetic correlations.** To determine whether male and female traits should evolve independently or in concert, we also estimated between-sex genetic correlations for the three traits. We used a REML model similar to the previous model, except traits were assigned as sex-specific, resulting in a model that included six traits: male wing length, male tail length, male tail white, female wing length, female tail length, and female tail white. The REML values of the genetic correlations calculated between male and female traits provide estimates of the between-sex genetic correlations. This method of estimating genetic correlations does not generate estimates greater than unity or negative standard errors, as is possible using the typical, regression-based method (Price and Burley 1993; Lynch and Walsh 1998; Merilä et al. 1998). We tested for a significant difference from both zero and unity using one-tailed, one-sample $t$-tests, with degrees of freedom equal to the number of individuals in the model minus one.

**Assortative mating.** Because estimates of heritability and genetic correlations may be biased by assortative mating, we tested for nonrandom mating by examining correlations
between phenotypic values of dams with the average phenotypic values of their mates, weighted by the number of offspring sired.

Selection analysis

**DATASET.** We measured selection using birds of known age (i.e. they were first captured as juveniles; 1431 males, 1329 females) hatched between 1989 and 1996. The analysis included only individuals that had measurements available for all three traits. Because none of the individuals in the analysis was still alive at the time the analysis was performed, our sample consisted purely of individuals tracked over their entire lifespan.

**FITNESS COMPONENTS.** To partition selection into different episodes, we measured selection using each of four fitness components (juvenile survival, adult survival, average mating success, and average fecundity per mate) chosen such that they would multiply to give an estimate of lifetime fitness (Arnold and Wade 1984a, b). Each fitness component was transformed to relative fitness \((w)\) in each regression by dividing by mean fitness. We estimated selection separately for males and females.

**SURVIVAL.** Juvenile and adult survival were considered separately because natal dispersal (Nolan et al. 2002) is more likely to inflate estimates of mortality at the juvenile stage. Juvenile survival was counted as 1 if an individual was captured or sighted as an adult in the year after hatching or in any subsequent year. Individuals that were not captured were
assigned a value of 0. Adult survival refers to the number of summers a bird was 
recaptured or re-sighted as an adult after having been banded as a juvenile. Adult survival 
ranged from 1 to 7. Males that had received testosterone implants that had not been 
removed at the end of a breeding season were excluded from the analysis of adult 
survival, because a prolonged exposure to testosterone may inhibit molt, decreasing 
over-winter survival (Nolan et al. 1992).

REPRODUCTIVE SUCCESS. Measurements of reproductive success (mating success and 
fecundity per mate) were taken only from those years for which DNA analysis was 
conducted (1990-1996) and were based on the number of nestlings that survived to day 6, 
the age at which we collected blood samples for genetic analysis. The three focal traits 
are not expressed until after the feathers are grown (after fledging), so they should not 
experience direct selection before day 6. Therefore, measuring reproductive success as 
counts of day-6 nestlings should not bias our measurements of selection by assigning 
offspring fitness to the parent (Wolf and Wade 2001).

Mating success was calculated by counting the number of mates with which an 
individual produced a day-6 offspring. During a given breeding season, mating success 
ranged from 0 to 3. We were not able to detect individuals, if any, that were unable to 
pair, so our estimates of sexual selection are conservative (Ketterson et al. 1997). 
Nevertheless, some individuals did receive a value of zero for a given year; these were 
males that sired none of the offspring of their social mate and achieved no detectable 
extra-pair fertilizations. Fecundity per mate ranged from 0 to 7 and was calculated by 
dividing the total number of day-6 offspring produced in a given year by the number of
mates in that year. Individuals that had no young surviving to day 6 were assigned a value of zero. If birds had records of mating success or fecundity from two or more years, we averaged them to generate a single score.

**Selection gradients.** We used multiple linear regression to calculate selection gradients, which are estimates of the direct force of selection on a given trait when considered independently of the effects of selection on correlated traits included in the analysis (Lande and Arnold 1983). Linear (directional) selection gradients (\( \beta \)) indicate selection that changes the population mean, and nonlinear selection gradients (\( \gamma \)) indicate selection that acts on either the phenotypic variance of a trait \( i \) (\( \gamma_i \), quadratic selection) or the phenotypic covariance between two traits \( i \) and \( j \) (\( \gamma_{ij} \), correlational selection).

Linear gradients were estimated from a regression model that excluded cross products and squared terms, while non-linear gradients were estimated from a full model (Lande and Arnold 1983; Mitchell-Olds and Shaw 1987; Brodie et al. 1995). Regression residuals were not normally distributed, so we could not use parametric tests of significance for selection gradients (Mitchell-Olds and Shaw 1987). We calculated standard errors of regression coefficients using a simple delete-one-individual jackknife technique and tested for significance using one-sample \( t \)-tests of the jackknifed estimates following the method described in Sokal and Rohlf (1995, pp. 820-823). Because of the small sample size in many of the selection analyses, we note trends \((P < 0.1)\) as well as statistical significance \((P < 0.05)\). We used a Bernoulli process to calculate the probability that the number of significant gradients measured was due to chance (Moran 2003). This method is more appropriate for the interpretation of tables with many small \( P \)
values than a sequential Bonferroni (Rice 1989), which is overly restrictive (Moran 2003).

All traits were standardized to zero mean and unit variance to facilitate comparisons among selection gradients (Arnold and Wade 1984b). Because the sample varied across analyses, this standardization was performed separately for each regression.

**MALE HORMONAL TREATMENT.** Interpretation of the selection gradients measured in the study may be complicated by hormonal treatment of males, because testosterone treatment is likely to affect some fitness components (e.g., T-males obtain more EPFs than C-males, Raouf et al. 1997; Reed et al. 2006). However, measurements of selection should not be affected unless there is a correlation between hormonal treatment and one of the traits under consideration, for example, if larger males were more likely to have T implants (Lande and Arnold 1983). All of the traits in this study were measured on juveniles, and birds were not implanted until the beginning of their first breeding season, so testosterone treatment cannot have affected the development of the traits. Hormone treatment was applied randomly to males; T-males, C-males, and untreated males did not differ with respect to any of the traits in this study (MANOVA, Wilks’ λ = 0.98, $F_{6,488} = 0.868$, $P = 0.52$). Nevertheless, we tested for an effect of hormone treatment on measures of selection on adult males (via adult survival, mating success, and fecundity) by running separate regressions that included hormone treatment as an independent variable. Because males did not always receive the same treatment during each breeding season, the covariate included in the analysis was the number of years a male received testosterone implants divided by the total number of years it was alive. The selection
Correlational selection

gradients from these analyses were not significantly different from those calculated in the original regressions, and hormone treatment did not significantly increase the fit of the models (partial $F$-test, $F_{1,9} > 3.65$, $P > 0.05$) so we report only the analyses that do not consider treatment.

**LIFETIME SELECTION.** Measurements of selection over the entire lifetime are necessary to make evolutionary predictions. Because our fitness components multiplied to lifetime fitness, we could add selection gradients measured at different episodes as an estimate of lifetime selection (Arnold and Wade 1984a, b). Selection gradients were summed separately for each sex. To estimate lifetime selection on both sexes combined, we averaged the sex-specific lifetime selection gradients. We calculated standard errors on these lifetime selection gradients by taking the square root of the sum of the squared standard errors estimated for each fitness component using jackknifing. This method assumes that the selection gradients from each analysis are independent (i.e., covariance between all pairs of selection gradients is zero). To test for significance, we performed one-sample $t$-tests using sums of jackknifed selection gradient estimates.

**FITNESS SURFACES.** Plotting fitness surfaces allows visualization of the form of selection simultaneously acting on two traits (Phillips and Arnold 1989; Brodie et al. 1995). We plotted a non-parametric representation of the fitness surface generated using a thin-plate spline fit, the three-dimensional analog of the cubic spline (Green and Silverman 1994; Blows et al. 2003). The smoothing parameter $\lambda$ for each spline was chosen by minimizing the generalized cross-validation score (Green and Silverman 1994). We used “R”
software (routine TPS, package FIELDS) to fit splines for fitness surfaces for each of the individual selection episodes. Where appropriate, we also used selection gradients (from both individual selection episodes and lifetime selection) to generate parametric fitness surfaces (Lande and Arnold 1983; Phillips and Arnold 1989; Brodie et al. 1995). Because parametric fitness surfaces are constrained to a limited number of shapes and because their interpretation may at times be misleading (e.g. because of extrapolation into areas based on few observations), the non-parametric splines were used as a guide to interpreting these fitness surfaces (Schluter 1988; Schluter and Nychka 1994; Brodie et al. 1995). Due to the composite nature of our measurements of lifetime selection as the product of individual episodes of selection, we were only able to generate parametric surfaces for lifetime fitness.

**Response to selection**

**CHANGE IN TRAIT MEANS.** We used lifetime selection gradients to predict evolutionary response to selection. The predicted response in trait means was calculated using the multivariate breeders’ equation,

\[ \Delta \bar{z} = G\beta \]  

(Lande and Arnold 1983). Although maternal effects may affect a predicted response to selection (Lande and Kirkpatrick 1990), to simplify the calculation we ignore them here.
Correlational selection

CHANGE IN VARIANCE. To estimate how the genetic variance-covariance matrix should change in response to selection, we used the equation

$$\Delta G = G(\gamma - \beta\beta^T)G \quad (3)$$

(Phillips and Arnold 1989). This equation describes change in the $G$ matrix that occurs within a generation; between-generation changes should be smaller due to the effects of recombination (Tallis and Leppard 1988; Tallis 1989; Wolf and Brodie 1998).

We used a standardized $G$ matrix ($h^2$ values on the diagonal, $r_{ij}h_ih_j$ off the diagonal) in both equations so that responses would be in standardized units.

Results

Quantitative inheritance of size and plumage traits

ADDITIVE GENETIC EFFECTS. Wing length, tail length, and tail white were all significantly heritable (i.e., showed significant additive genetic variance, Table 2.2). Wing length was genetically correlated with both tail length and tail white, but tail length and tail white were not genetically correlated (Table 2.2).
Table 2.2. Genetic parameters and variance components estimated using restricted maximum likelihood and an animal model. a. Heritabilities, $h^2 \pm SE$, are shown on the diagonal, genetic correlations, $r_g$, are below the diagonal, and phenotypic correlations are above the diagonal. b. Additive genetic (co)variance components ($G \pm SE$) are shown below the diagonal, and phenotypic (co)variance components are shown above the diagonal.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Wing length</th>
<th>Tail length</th>
<th>Tail white</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. $h^2$ and $r_g$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wing length</td>
<td>0.33 ± 0.100**</td>
<td>(0.46)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>Tail length</td>
<td>0.76 ± 0.118***</td>
<td>0.53 ± 0.095***</td>
<td>(0.09)</td>
</tr>
<tr>
<td>Tail white</td>
<td>0.41 ± 0.174*</td>
<td>0.04 ± 0.147</td>
<td>0.50 ± 0.082***</td>
</tr>
<tr>
<td>b. G</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wing length</td>
<td>0.67 ± 0.223**</td>
<td>(1.37)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Tail length</td>
<td>0.94 ± 0.264***</td>
<td>2.29 ± 0.492***</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Tail white</td>
<td>0.060 ± 0.027*</td>
<td>0.010 ± 0.040</td>
<td>0.032 ± 0.006***</td>
</tr>
</tbody>
</table>

Significance tested with two-tailed, one-sample $t$-tests ($H_0: \mu = 0$). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

**Between-sex genetic correlations.** Genetic correlations between the sexes were high for all traits ($r_g \pm S.E.:$ wing length $0.81 \pm 0.406$, tail length $0.89 \pm 0.231$, tail white $0.97 \pm 0.270$). All between-sex genetic correlations were significantly different from zero (one-tailed, one-sample $t$-tests, $df = 632$, $P < 0.05$), but not from unity ($P > 0.05$).

**Assortative mating.** Heritability measurements were not biased by assortative mating, as individuals did not mate assortatively based on any of the three traits ($0.006 \leq r \leq 0.076$, $P > 0.47$). Genetic correlations were probably also not affected by assortative mating.
Correlational selection

One trait pair (dam’s tail white and sire’s wing length) showed a trend toward assortative mating, but in the opposite direction of the observed genetic correlation ($r = -0.200$, $P = 0.079$). Other trait pairs were not correlated ($-0.120 \leq r \leq 0.122$, $P > 0.352$).

Selection

Juvenile survival. There was no directional selection via juvenile survival in either sex (Table 2.3). There was a trend toward pure quadratic selection acting to increase variance in male wing length (Table 2.3, Figure 2.2a), and the non-parametric fitness surface suggests the existence of two fitness peaks, one for very small males, and one for males that are only slightly larger than average (Figure 2.3a). There were no other significant non-linear selection gradients in either sex.

Adult survival. There was no significant evidence of selection based on adult survival on any of the three traits in males (Table 2.3). There was no directional selection acting on females, but we detected significant correlational selection acting on wing length and tail white (positive, Table 2.3, Figure 2.2c) and tail length and tail white (negative, Table 2.3). Non-parametric fitness surfaces suggest that larger females with more tail white
Table 2.3. Matrices of standardized directional (β) and quadratic (γ) selection gradients for wing length, tail length, and tail white. Diagonal elements in the quadratic selection matrix represent quadratic (γii) selection and off-diagonal elements represent correlational selection (γij). Selection gradients are partial regression slopes ± one standard error. Gradients are estimated separately for each sex, using four components of fitness. Sample size for each regression is reported below each set of matrices. Gradients with \( P < 0.10 \) are shown in boldface. The probability of obtaining nine gradients with \( P < 0.10 \) by chance is 0.12, and the probability of obtaining six gradients with \( P < 0.05 \) is 0.08 (Moran 2003).

<table>
<thead>
<tr>
<th>Fitness component</th>
<th>Trait</th>
<th></th>
<th></th>
<th>MALES</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>FEMALEs</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β</td>
<td>γ</td>
<td>γ</td>
<td>Wing length</td>
<td>Tail length</td>
<td>Tail white</td>
<td>β</td>
<td>γ</td>
<td>Wing length</td>
<td>Tail length</td>
<td>Tail white</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Juvenile survival</td>
<td>Wing length</td>
<td>-0.063</td>
<td>0.191</td>
<td>± 0.0593</td>
<td>± 0.1062</td>
<td>-0.050</td>
<td>0.092</td>
<td>± 0.0626</td>
<td>± 1.088</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tail length</td>
<td>0.010</td>
<td>-0.018</td>
<td>-0.114</td>
<td>± 0.0806</td>
<td>± 0.0868</td>
<td>0.071</td>
<td>0.046</td>
<td>0.078</td>
<td>± 0.0813</td>
<td>± 0.0878</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tail white</td>
<td>0.029</td>
<td>0.013</td>
<td>0.071</td>
<td>± 0.0590</td>
<td>± 0.0734</td>
<td>0.069</td>
<td>0.003</td>
<td>-0.006</td>
<td>-0.006</td>
<td>± 0.0974</td>
<td>± 0.0718</td>
<td>± 0.0724</td>
</tr>
<tr>
<td></td>
<td></td>
<td>± 0.0511</td>
<td>± 0.0620</td>
<td>± 0.0950</td>
<td>± 0.0734</td>
<td>± 0.0552</td>
<td>± 0.0629</td>
<td>± 0.0718</td>
<td>± 0.0724</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(N=1431)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(N=1329)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult survival</td>
<td>Wing length</td>
<td>0.009</td>
<td>-0.047</td>
<td>± 0.0512</td>
<td>± 0.0918</td>
<td>0.062</td>
<td>0.136</td>
<td>± 0.0496</td>
<td>± 0.0786</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tail length</td>
<td>-0.012</td>
<td>-0.021</td>
<td>0.034</td>
<td>± 0.0745</td>
<td>± 0.0976</td>
<td>-0.044</td>
<td>-0.074</td>
<td>0.082</td>
<td>± 0.0551</td>
<td>± 0.0619</td>
<td>± 0.0582</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tail white</td>
<td>-0.027</td>
<td>0.082</td>
<td>-0.038</td>
<td>0.072</td>
<td>± 0.0602</td>
<td>± 0.0580</td>
<td>± 0.0698</td>
<td>± 0.0399</td>
<td>± 0.0547</td>
<td>± 0.0504</td>
<td>± 0.0492</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>± 0.0433</td>
<td>± 0.0602</td>
<td>± 0.0580</td>
<td>± 0.0698</td>
<td>± 0.0551</td>
<td>± 0.0619</td>
<td>± 0.0582</td>
<td>± 0.0492</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(N=249)</td>
<td></td>
<td></td>
<td></td>
<td>(N=264)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mating success</td>
<td>Wing length</td>
<td>0.173**</td>
<td>-0.011</td>
<td>± 0.0669</td>
<td>± 0.1726</td>
<td>-0.027</td>
<td>-0.003</td>
<td>± 0.0344</td>
<td>± 0.0416</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tail length</td>
<td>-0.117†</td>
<td>0.125</td>
<td>-0.048</td>
<td>± 0.0693</td>
<td>± 0.2234</td>
<td>0.046</td>
<td>0.013</td>
<td>0.060</td>
<td>± 0.0689</td>
<td>± 0.0504</td>
<td>± 0.1402</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tail white</td>
<td>0.014</td>
<td>0.137†</td>
<td>-0.143</td>
<td>0.011</td>
<td>± 0.0456</td>
<td>± 0.0740</td>
<td>± 0.0868</td>
<td>± 0.0900</td>
<td>± 0.0334</td>
<td>± 0.0361</td>
<td>± 0.0706</td>
<td>± 0.0706</td>
</tr>
<tr>
<td></td>
<td></td>
<td>± 0.0456</td>
<td>± 0.0470</td>
<td>± 0.0868</td>
<td>± 0.0900</td>
<td>± 0.0551</td>
<td>± 0.0619</td>
<td>± 0.1402</td>
<td>± 0.0706</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(N=88)</td>
<td></td>
<td></td>
<td></td>
<td>(N=77)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fecundity per mate</td>
<td>Wing length</td>
<td>-0.010</td>
<td>-0.191</td>
<td>± 0.0995</td>
<td>± 0.2248</td>
<td>-0.143*</td>
<td>0.080</td>
<td>± 0.0685</td>
<td>± 0.1248</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tail length</td>
<td>0.073</td>
<td>0.100</td>
<td>0.055</td>
<td>± 0.1014</td>
<td>± 0.2496</td>
<td>0.196*</td>
<td>0.027</td>
<td>-0.037</td>
<td>± 0.0774</td>
<td>± 0.1331</td>
<td>± 0.1268</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tail white</td>
<td>0.072</td>
<td>0.115</td>
<td>-0.040</td>
<td>-0.122</td>
<td>± 0.0721</td>
<td>± 0.1110</td>
<td>± 0.1122</td>
<td>± 0.1306</td>
<td>± 0.0702</td>
<td>± 0.0945</td>
<td>± 0.1101</td>
<td>± 0.1150</td>
</tr>
<tr>
<td></td>
<td></td>
<td>± 0.0721</td>
<td>± 0.1110</td>
<td>± 0.1122</td>
<td>± 0.1306</td>
<td>± 0.0702</td>
<td>± 0.0945</td>
<td>± 0.1101</td>
<td>± 0.1150</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(N=106)</td>
<td></td>
<td></td>
<td></td>
<td>(N=95)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard errors are jackknife estimates; significance is estimated by t-tests.
† = \( P < 0.10 \); * = \( P < 0.05 \); ** = \( P < 0.01 \).
Figure 2.2. Parametric fitness surfaces for wing length and tail white, drawn using all gradients from Table 2.3. Trait values (x and y axes) are standardized to zero mean and unit variance, and w (z axis) represents relative fitness. Scales differ for each graph according to the range of trait and fitness values used in the analysis. (a) Male juvenile survival. (b) Male mating success. (c) Female adult survival. (d) Female mating success. (e) Female fecundity per mate.
Figure 2.3. Non-parametric fitness surfaces for wing length and tail white, fitted using a thin-plate spline. Trait values (x and y axes) are standardized to zero mean and unit variance, and \( w \) (z axis) represents relative fitness. Scales differ for each graph according to the range of trait and fitness values used in the analysis. (a) Male juvenile survival. (b) Male adult survival. (c) Male mating success. (d) Male fecundity per mate. (e) Female juvenile survival. (f) Female adult survival. (g) Female mating success. (h) Female fecundity per mate.
Correlational selection may have a survival advantage (Figure 2.3f), and that two fitness peaks—larger with more tail white and smaller with less tail white—may exist for males (Figure 2.3b).

MATING SUCCESS. Sexual selection was most evident in males; males with longer wings and shorter tails (n.s.) had higher mating success (Table 2.3). Although there was no directional sexual selection on tail white, there was a trend toward correlational sexual selection between tail white and wing length (Table 2.3). The combination of this correlational selection gradient and the significant linear selection gradient on wing length results in males with the longest wings and the whitest tails having the highest mating success (Figure 2.2b). The non-parametric fitness surface suggests the existence of a fitness ridge with one peak at relatively large values for each trait (Figure 2.3c).

Sexual selection was weak in females, but as in males, there was a significant negative correlational gradient between wing length and tail white (Table 2.3). In other words, females with matching tail white and wing length (less white with shorter wings or more white with longer wings) had fewer mates (Figure 2.2d). Both fitness surfaces are nearly flat, but suggest that small females with more tail white have slightly higher mating success (Figures 2.2d, 2.3g).

FECUNDITY. Fecundity selection was negligible in males, but relatively strong in females. Directional selection favored females with shorter wings and longer tails (Table 2.3, Figures 2.2e, 2.3h). There was no significant evidence of non-linear fecundity selection.
Table 2.4. Lifetime selection acting on males, females, and both sexes combined, calculated using selection gradients from Table 2.3 (see text for details). Diagonal elements in the quadratic selection matrix represent quadratic ($\gamma_{ii}$) selection and off-diagonal elements represent correlational selection ($\gamma_{ij}$). Selection gradients are partial regression slopes $\pm$ one standard error. Gradients with $P < 0.10$ are shown in boldface.

<table>
<thead>
<tr>
<th>Trait</th>
<th>MALES</th>
<th>FEMALES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta$ Wing length</td>
<td>$\gamma$ Tail length</td>
</tr>
<tr>
<td>Wing length</td>
<td>0.109 $\pm$ 0.143</td>
<td>-0.069 $\pm$ 0.316</td>
</tr>
<tr>
<td>Tail length</td>
<td>-0.046 $\pm$ 0.144</td>
<td>0.186 $\pm$ 0.287</td>
</tr>
<tr>
<td>Tail white</td>
<td>0.087 $\pm$ 0.108</td>
<td>$\pm$ 0.159$^*$</td>
</tr>
<tr>
<td></td>
<td>($N=1431$)</td>
<td></td>
</tr>
</tbody>
</table>

**SEXES COMBINED**

<table>
<thead>
<tr>
<th>Trait</th>
<th>$\beta$ Wing length</th>
<th>$\gamma$ Tail length</th>
<th>Tail white</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wing length</td>
<td>-0.024 $\pm$ 0.091</td>
<td>$\pm$ 0.065$^*$</td>
<td></td>
</tr>
<tr>
<td>Tail length</td>
<td>0.111</td>
<td>0.053</td>
<td>0.054</td>
</tr>
<tr>
<td>Tail white</td>
<td>0.079</td>
<td>$\pm$ 0.073$^{**}$</td>
<td>$\pm$ 0.080$^*$</td>
</tr>
<tr>
<td></td>
<td>($N=1380$)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^\dagger = P < 0.10; \ * = P < 0.05; \ ** = P < 0.01.$
LIFETIME SELECTION. When we considered selection combined over all episodes, we found that selection significantly favored long tails in females and a correlation between wing length and tail white in males (Table 2.4). This correlational selection gradient was also significant when selection was averaged across the sexes, and it was the main determinant of the shape of the wing-tail white fitness surfaces in both sexes (Table 2.4, Figure 2.4). There was also significant negative correlational selection between tail length and tail white when the sexes were combined, favoring a decreased relationship between the two traits (Table 2.4). None of the other lifetime selection gradients was significant.

Response to selection

Combining genetic data with measurements of lifetime selection led to predictions of very small increases in all three trait means over time (Table 2.5). Small increases in the genetic variance were also predicted (Table 2.5) due to the absence of strong directional selection and the weak positive quadratic selection acting on all traits (Table 2.4). Predicted changes in genetic covariance were in the same direction as the observed genetic correlations (Table 2.2, Table 2.5).
Figure 2.4. Fitness surfaces for wing length and tail white, drawn using all gradients from Table 2.4. Trait values (x and y axes) are on a standardized scale, and w (z axis) represents relative fitness. (a) Male lifetime selection. (b) Female lifetime selection. (c) Total lifetime selection.
Table 2.5. Predicted response to selection (see text for details of calculations). (a) Predicted between-generation change in the mean in standard deviation units. (b) Predicted within-generation change in standardized additive genetic variances (diagonal) and covariances (below the diagonal).

<table>
<thead>
<tr>
<th>Trait</th>
<th>a. ΔZ</th>
<th>Wing length</th>
<th>Tail length</th>
<th>Tail white</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wing length</td>
<td>0.040</td>
<td>0.036</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tail length</td>
<td>0.053</td>
<td>0.031</td>
<td>0.041</td>
<td></td>
</tr>
<tr>
<td>Tail white</td>
<td>0.037</td>
<td>0.029</td>
<td>0.001</td>
<td>0.052</td>
</tr>
</tbody>
</table>

Discussion

We found that the three traits examined in this study, wing length, tail length, and tail white, were heritable, and wing length, which is representative of overall body size, was positively genetically correlated with both tail length and tail white. While a positive genetic correlation between wing length and tail length is probably inevitable due to common developmental pathways, the genetic correlation between wing length and tail white may have evolved via correlational selection. We detected positive lifetime correlational selection acting on the two traits, which was largely the result of sexual selection on males, which favored large males with whiter tails. While directional sexual selection on males was largely opposed by other selection episodes and selection on
females, correlational sexual selection was reinforced. We discuss the implications of each of these results below.

*Correlational sexual selection*

In males, the sexual selection surface for wing length and tail white is a rising ridge, a pattern that can lead to the evolution of honest signals via the handicap principle (Getty 1998). If wing length, which is a good indicator of body size, is also indicative of overall phenotypic quality, then positive correlational selection on wing length and tail white may function to maintain tail white as an honest signal. Getty’s (1998) original formulation required an interaction between sexual selection and viability to create a fitness ridge, but other kinds of interaction can produce the same result. A rising fitness ridge could arise from pure sexual selection when mating success depends on an interaction between intrasexual and intersexual interactions (Berglund et al. 1996; Ligon 1999). Larger males with large ornaments should have the highest mating success, while “mismatched” males should fare more poorly due either to decreased attractiveness to females or decreased ability to compete with other males.

In juncos, larger males may be more successful in intrasexual competition for access to mates during the breeding season, as they are in contests over food during the winter (Balph et al. 1979; Ketterson 1979; Holberton et al. 1989). If so, they may be better able to defend against intrusions from rival males seeking extra-pair copulations or better able to invade neighboring territories where they might be more likely to
inseminate fertile females. Females may choose among the males they encounter based on tail white (Hill et al. 1999), causing the larger males with more tail white to have the highest mating success. Because a female that has settled on a male’s territory may choose only among him and a limited number of neighbors, and the males that she encounters may be restricted by male-male competition, her choices may be limited. This may explain why tail white, despite its attractiveness to females (Hill et al. 1999), is correlative, but not directionally, selected.

Consistent correlational selection acting on wing length and tail white should lead to an increase in the genetic correlation between the two traits. The positive genetic correlation we observed suggests that correlational selection may have integrated these two traits in the past. Furthermore, quantitative genetic models of evolutionary change predict that the genetic covariance between the two traits should continue to increase. Consequently, to the extent that body size is associated with quality, tail white is an honest signal that should be reliable across generations. Other studies have shown concordance between correlational selection and genetic correlations (Brodie 1989, 1992, 1993b; Conner and Via 1993; Morgan and Conner 2001), but to our knowledge, this is the first study to demonstrate such a relationship for a sexually selected trait in animals. LeBas et al. (2003) suggested that correlational sexual selection may lead to a genetic correlation that maintains signal honesty of a female ornament in dance flies, but they did not demonstrate a genetic correlation. In that study, correlational sexual selection favored a positive relationship between a female ornament (size of pinnate scales on the hind femur) and fecundity, and as a result, males that chose to mate with females with larger scales were able to sire more offspring. Because fecundity may also be related to nuptial
gifts received from males, however, the phenotypic correlation observed by LeBas et al. (2003) may be largely environmental.

**Strength of selection and interaction of selection episodes**

In addition to correlational selection, we found some evidence of directional selection, although it was quite weak; the strongest directional selection gradients were close to the median ($|\beta| = 0.16$) of those reported in a recent review (Kingsolver et al. 2001). In accordance with the findings of Hoekstra et al. (2001), the strongest directional selection occurred via sexual and fecundity selection, and as expected, males were selected most strongly due to variance in mating success, while females were selected most strongly due to variance in fecundity (Shuster and Wade 2003).

Although we found no evidence of directional survival selection in this study, it probably occurs over short periods of time in relation to short-term environmental changes. Indeed, survival selection tends to be stronger when it is measured over a shorter period of time (compare Hoekstra et al. 2001). If these episodes of selection balance each other out (e.g. because the environment fluctuates), no net selection would be detectable over a period of years as was the case in our study.

We did not detect quadratic selection in individual selection episodes, although when all episodes were combined, there was a trend toward quadratic selection for an increase in the variance of wing length. The strongest quadratic gradients were not much larger than the median value ($|\gamma| = 0.10$) found by Kingsolver et al. (2001). Correlational
selection, as indicated above, was pervasive but was not particularly strong at any one episode when compared to other published measurements (Sinervo and Svensson 2002).

Although the strength of selection differed among episodes, there was no strong evidence of opposing selection (i.e., significant selection gradients of opposite sign) within a sex. Directional selection tended to be important at only one episode in each sex (sexual selection in males and fecundity selection in females). The importance of these selection episodes is apparent when examining the lifetime fitness surfaces. The lifetime fitness surface for males, like the sexual selection surface, resembles a rising ridge. In females, the lifetime fitness surface is shallower and valley-shaped, similar in shape to the fecundity selection surface.

In most cases, quadratic and correlational selection measured at different episodes tended to be reinforcing. One notable case of this was correlational selection on male wing length and tail white. Although it was only detectable in one selection episode (sexual selection), this correlational selection was consistently positive across all episodes, leading to a fairly strong lifetime selection gradient.

**Selection and sexual dimorphism**

Opposing selection in males and females may lead to the evolution of sexual dimorphism (Lande 1980b; Slatkin 1984; Andersson 1994; Badyaev and Martin 2000; Shuster and Wade 2003). In this study, selection favored an increase in sexual dimorphism in wing length (larger males, smaller females favored), and a decrease in sexual dimorphism for
Correlational selection

tail length (females with longer tails favored). All lifetime selection gradients involving only wing length and tail length were of opposite sign in males and females, suggesting that size-related traits are under substantially different selection pressures in males and females. Despite the differences in male and female selection regimes, however, the high between-sex genetic correlations we measured may constrain further evolution of sexual dimorphism (Lande 1980b; Price and Burley 1993; Merilä et al. 1998). When both sexes were considered, selection gradients for size-related traits were nearly zero, leading to very little predicted evolutionary change in the trait means, despite heritability measures that were typical of morphological traits (Mousseau and Roff 1987).

In contrast to selection on size-related traits, selection gradients related to tail white were all of the same sign for males and females. This is especially notable for correlational selection acting on size-related traits and tail white. The correlational selection in males that may maintain the honesty of tail white as a sexual signal is actually reinforced by selection in females, making an evolutionary change in the $G$ matrix more likely.

The interaction of lifetime selection on males (a rising ridge shape) and lifetime selection on females (a flatter, valley-like shape) creates a total fitness surface more similar to a saddle. The defining feature of this surface is the strong correlational selection between wing length and tail white, which derives primarily from sexual selection on males. However, because directional selection on males is balanced by opposing selection on females, the two points of high fitness are roughly equivalent. This concave selection surface should maintain the correlation between wing length and tail
Correlational selection

white while also maintaining genetic variance in wing length (Brodie 1992; Blows et al. 2003).

Conclusion

This study provides one of the few reported examples of correlational sexual selection in a natural population (Moore 1990; Fairbairn and Preziosi 1996; Rodriguero et al. 2002; LeBas et al. 2003, 2004), and to our knowledge it is the first to show concordance between such selection and genetic correlations between sexually selected traits. This finding, along with those of other studies (Brodie 1989, 1992, 1993b; Conner and Via 1993; Morgan and Conner 2001), suggests an important role for correlational selection in generating and maintaining genetic integration between functionally related traits. However, such studies cannot directly implicate selection as the cause of genetic correlations, because genetic correlations are observed at only one point in time. Future work should use long-term field data or experimental systems to document how correlational selection affects changes in genetic correlations over time.

Correlational sexual selection may be of general importance in the evolution of sexually selected traits. As we have suggested, correlational sexual selection may lead to the evolution of signal honesty (see also LeBas et al. 2003). Correlational sexual selection may also serve to integrate multiple ornamental traits (Moore 1990; Rodriguero et al. 2002; Candolin 2003) or multiple aspects of a composite ornamental trait (Badyaev et al. 2001). This process of phenotypic integration may be opposed when there are
genetic constraints, such as tight linkage (Brooks and Endler 2001). In such cases, correlational sexual selection may create multiple, stable fitness peaks for combinations of attractive traits, which may lead to the preservation of genetic variance in individual attractive traits (Blows et al. 2003). More studies of correlational sexual selection, and correlational selection in general, should be conducted to allow us to evaluate its evolutionary importance (Kingsolver et al. 2001; Sinervo and Svensson 2002).

Although studies such as this one are useful because they allow quantification of the way selection acts on natural populations, experiments are necessary in order to understand the mechanisms of how selection occurs (Wade and Kalisz 1990). Although some such studies have already been conducted on male juncos (Balph et al. 1979; Ketterson 1979; Holberton et al. 1989; Hill et al. 1999), more studies are necessary to determine how male traits interact to generate the observed correlational sexual selection observed here. As we have demonstrated here, selection operating on females is important for understanding the evolution of male traits. Consequently, future studies should also explore the mechanisms underlying selection on correlated traits in females.
Acknowledgments

We thank Craig Benkman, Rob Brooks, and an anonymous reviewer for suggestions that greatly improved this manuscript. This work would have been impossible without the advice and guidance of Butch Brodie III and the work of N. Arguedas, D. Monk, S. Raouf, and T. Peare, who conducted most of the genetic parentage analysis. We thank the many collaborators, post-doctoral associates, graduate students, and research assistants who captured and measured thousands of juncos between 1989 and 1996: S. Bentz, K. Bruner, L. Callahan, D. Cullen, J. M. Cawthorn, M. Chambers, C. R. Chandler, L. Christensen, D. Enstrom, G. Gonzalez, J. Hill, S. Hudman, T. Kast, E. Kennedy, K. Kimber, L. Klukowski, S. Lynn, J. Martinez-Sanchez, G. McPeek, J. Mikesell, D. Morris, S. Radjy, S. Raouf, M. Rosenshield, L. Rowe, S. Peckham, M. Ragland, S. Schoech, M. Soensken, J. Steele, A. Stoehr, M. Tavel, R. Titus, M. Watson, L. Wolf, and especially E. Snajdr and C. Ziegenfus. We are grateful to H. Wilbur and Mountain Lake Biological Station of the University of Virginia for providing facilities and a stimulating scientific environment and the Mountain Lake Hotel and Wilderness Conservancy for providing much of our study area. We also thank E. S. Allen, J. Casto, D. Duffy, A. Eklund, N. Gerlach, J. Grindstaff, B. Heidinger, J. Kingsolver, M.-L. Maas, E. Martins, T. Price, W. Reed, B. Ridenhour, S. Schrock, M. Wade, W. Wolf, and D. Zysling for comments and helpful discussions. The authors were supported by a National Science Foundation (NSF) graduate research fellowship to JWM and NSF grants BSR 91-11498, BSR 87-18358, IBN 94-08061, and IBN 97-28384 to EDK and VN.
CHAPTER THREE

Hormones and honest signals:
males with larger ornaments elevate
testosterone more when challenged

JOEL W. McGlothlin
JODIE M. JAWOR*
TIMOTHY J. GREIVES
JOSEPH M. CASTO†
JENNIFER L. PHILLIPS‡
ELLEN D. KETTERSON

Submitted to Journal of Evolutionary Biology
*Department of Biological Sciences, University of Southern Mississippi, Hattiesburg, MS
†Department of Biological Sciences, Illinois State University, Normal, IL
‡Biology Department, College of William & Mary, Williamsburg, VA
Summary

When males invest differentially in mating depending on their quality, reliable sexual signals may evolve. In many songbirds, testosterone mediates mating investment, suggesting that signals should be linked to testosterone production. However, because testosterone may change rapidly during relevant behavior (such as territorial aggression and courtship), efforts to establish such a relationship have proved challenging. In a population of dark-eyed juncos, we measured short-term testosterone increases by injecting gonadotropin-releasing hormone (GnRH). We found a positive correlation between the magnitude of these increases and the size of a plumage ornament (“tail white”) used in both female choice and male-male competition. We then measured testosterone changes induced naturally by male-male competition and found that they covaried with GnRH-induced levels. Combined, these results suggest that selection may favor signal reliability by maintaining the association between tail white and testosterone increases. Consequently, conspecifics may use tail white to reliably assess potential mates and competitors.
Introduction

The origin and maintenance of reliability, or honesty, in sexual signals has been an issue of longstanding interest in evolutionary biology (Andersson 1994; Kokko et al. 2003; Maynard Smith and Harper 2003; Searcy and Nowicki 2005). Traits that reliably signal male quality are expected to evolve when high-quality males are able to achieve higher total fitness for a given level of signaling than are low-quality males (Nur and Hasson 1984; Grafen 1990; Michod and Hasson 1990; Getty 1998, 2006). This occurs because acquiring mates requires an investment of energy and time, usually at the expense of other components of fitness. High-quality males are predicted to invest more heavily in mating effort, whereas low-quality males may compensate for lower mating success by investing in self-maintenance or parental effort. The evolution of honest sexual signals may be thus viewed as arising from selection to optimize the resolution of life-history trade-offs (Getty 1998, 2006; Kokko 1998; Kokko et al. 2002). In many cases, selection should act to maintain the correlation between sexual signals and relative allocation to mating effort (Getty 2006).

Life-history trade-offs are often mediated by hormonal mechanisms (Stearns 1992; Sinervo and Svensson 1998; Ketterson and Nolan 1999; Zera and Harshman 2001; Ricklefs and Wikelski 2002; Adkins-Regan 2005). In songbirds, allocation of energy to mating effort is mediated in part by testosterone. In many species, experimentally elevated testosterone tends to increase behaviors related to mating, such as song and display, while decreasing self-maintenance and parental care (Ketterson and Nolan 1992, 1999; Adkins-Regan 2005; Hau 2007). Differential testosterone production has the
potential to act as a proximate mechanism allowing males to invest differentially in mating effort based on their quality. Thus, when the predictions of signaling theory are combined with empirical knowledge about the role of testosterone, the expectation arises that the expression of sexual signals, such as colorful plumage, should be correlated with testosterone levels.

Ornamental plumage is one of the most common types of sexual signal in songbirds (Hill 2006; Senar 2006). Although testosterone is generally not responsible for generating sexual dimorphism in plumage (as it often is for fleshy ornaments such as wattles and spurs), several studies have reported positive correlations between ornamental plumage and testosterone in males (Owens and Short 1995; Kimball 2006). The most well-supported case comes from house sparrows (Passer domesticus), where testosterone seems to have a direct effect on the size of a male’s bib (Evans et al. 2000; Buchanan et al. 2001; Gonzalez et al. 2001; Strasser and Schwabl 2004). Circulating plasma testosterone has been linked to carotenoid coloration in the house finch (Carpodacus mexicanus, Duckworth et al. 2004, but see Stoehr and Hill, 2001) and the crown of the blue tit (Cyanistes caeruleus, Peters et al. 2006), and fecal testosterone was correlated with the size of the forehead patch of the collared flycatcher (Ficedula albicollis, Garamszegi et al. 2004). Testosterone also induces the pre-nuptial molt to breeding plumage in superb fairy-wrens (Malurus cyaneus, Peters et al. 2000). Although these studies provide evidence for a possible role for testosterone in the evolution of honest plumage signals, all were focused on static measurements of testosterone.

Testosterone levels have the potential to vary greatly over time, and may show both seasonal patterns and short-term changes in response to social stimuli (Wingfield et al.
Importantly, temporary elevations of testosterone are produced during behaviors closely linked with mate acquisition, namely male-male aggression and courtship of females (Harding 1981; Wingfield 1985; Wingfield et al. 1990; Wingfield et al. 2001; Pinxten et al. 2003). Such social modulation of testosterone (and other hormones) is common in birds and has been demonstrated in other taxa (Wingfield et al. 1990; Wingfield et al. 2001; Oliveira 2004; Hirschenhauser and Oliveira 2006; Scott 2006). In songbirds, short-term modulation of testosterone is most evident in species that produce multiple broods each breeding season (Landys et al. 2007). In these species, males may need to make rapid shifts between mating effort and parental effort because opportunities to obtain additional matings (via extra-pair copulations) overlap with the need to care for offspring. Because of the close association between the production of socially modulated testosterone levels and mating-related behavior in such species, a relationship between plumage and transient testosterone changes would provide stronger evidence for the existence of links between ornamentation and allocation to mating effort than would a relationship with static testosterone levels.

We tested for a relationship between transient increases in testosterone and plumage in a free-living population of dark-eyed juncos (*Junco hyemalis*). A male’s ability to produce temporary testosterone increases may be assessed by administering intramuscular injections of gonadotropin-releasing hormone (“GnRH challenges,” Jawor et al. 2006). In nature, when GnRH is released from the hypothalamus, it stimulates the release of luteinizing hormone from the pituitary, which in turn temporarily increases production and release of testosterone from the gonad. In juncos, the testosterone response to GnRH challenges shows repeatable variation among individuals across the
Figure 3.1. Tail spreading behavior of a male dark-eyed junco. During courtship, males erect their body plumage and spread their tail to reveal the white patch on their outer tail feathers, which is mostly hidden at rest. Testosterone increases the frequency of this display, and increasing the size of the patch increases its attractiveness to females. Males also display their tails during intrasexual interactions, and males with more tail white are socially dominant. Photograph by Britt Heidinger.

breeding season (Jawor et al. 2006). Further, natural variation in transient testosterone elevations has been linked to natural variation in trade-off between mating effort and parental effort (Chapter 4). Specifically, maximum testosterone levels following GnRH challenges accurately predicted increased aggression in response to a territorial intruder, a measure of mating effort, whereas the magnitude of the GnRH-induced increase in testosterone predicted decreased nestling feeding, a measure of parental effort (Chapter 4).
In this study, we measured the phenotypic correlation between GnRH-challenge response and an attractive plumage ornament (a white patch on the tail, or “tail white,” Hill et al. 1999; Wolf et al. 2004). Males use tail white in both courtship and aggressive displays, and males with enhanced tail white become socially dominant and are more attractive to females (Balph et al. 1979; Holberton et al. 1989; Hill et al. 1999) (Figure 3.1). In nature, larger males with whiter tails have the highest mating success (Chapter 2). We predicted that males with more attractive plumage would also generate larger testosterone increases in response to GnRH. We further predicted that males that generated larger increases in testosterone in response to GnRH would also generate larger increases under natural circumstances and that these increases would be associated with higher levels of aggression. If our predictions were borne out, we would conclude that males earlier shown to be more attractive, i.e., those with whiter tails, also possess the physiology known to accompany more intense mate-acquisition behavior. This covariation between appearance, physiology, and behavior would suggest that potential competitors and mates should be able to predict the outcome of interactions with an individual by assessing his ornamentation.
Methods

Study area and species

We studied a resident population of the Carolina subspecies of the dark eyed junco (*J. h. carolinensis*) that breeds at and around Mountain Lake Biological Station in Giles County, VA (37°22'N, 80°32'W). Male juncos defend breeding territories upon which a single female nests. Both parents care for offspring, but mating often occurs outside the pair, generating opportunity for sexual selection among males (Ketterson et al. 1997; Nolan et al. 2002).

Capture

In April-August 2003-2004, males (*n* = 90) were captured using mist nets or Potter traps. Upon capture, birds were transported to a central laboratory at MLBS in a holding bag. If previously uncaptured, birds were given a numbered aluminum leg band and a unique combination of plastic color leg bands for identification. We determined age (yearling or older adult [≥ 2 years]) by examining the color of the primary wing coverts, and secondarily, the iris, which are both lighter in yearlings (Nolan et al. 2002). Mass (g) was measured using a spring balance.
Each time a bird was captured, a blood sample was obtained from the wing vein (initial sample). Handling time was recorded as the time in min from capture to collection of this blood sample, averaging 48 min (range 2 – 217 min) (Jawor et al. 2006). A solution of 1.25 μg chicken GnRH-I (Sigma L0637) in 50μl of 0.1 M phosphate-buffered saline was then injected into the pectoral muscle. The bird was returned to its holding bag, and after 30 min, a second blood sample was collected (post-challenge sample). After this sample, the bird was released at the site of capture. Plasma was separated and frozen (-20ºC) for later hormone analysis.

To control for the idiosyncrasies of capture and to obtain robust individual estimates of testosterone production, we attempted to obtain four samples each year from individual birds, collected at four sampling stages across the breeding season (Jawor et al. 2006). We attempted to obtain two samples during early breeding (21 April – 16 May) by catching birds at random in baited mist nets and traps. The first GnRH challenge was administered upon each bird’s first capture (2003: 28 April – 16 May, n = 53; 2004: 21 April – 11 May, n = 44, combined n = 97) and the second after waiting 7 – 21 days (mean 10.4; 2003: 6 May – 16 May, n = 26; 2004: 1 May – 15 May, n = 11, combined n = 37). During early breeding, many birds were beginning to nest, but the exact stage of reproduction was unknown for most of them (dates of first egg were 26 April in 2003 and 25 April in 2004). Some birds were captured and given a GnRH challenge while feeding 6-7 day old nestlings (2003: 25 May – 29 June, n = 14, 2004: 20 May – 20 July, n = 14, combined n = 28). Captures during this stage were made by placing a mist net at the nest.
A final set of birds was captured at the end of the breeding season, but prior to the onset of molt, using baited mist nets (2003: 15 July – 6 August, n = 7; 2004: 20 July – 5 August, n = 9). All sampling periods occurred after the typical early-breeding season testosterone peak (26 March – 14 April, Ketterson and Nolan, 1992). Overall, 5 individuals were challenged a total of 5 times, 5 were challenged 4 times, 12 were challenged 3 times, 29 were challenged 2 times, and 39 were challenged once. Twenty-one (21) individuals received challenges in both 2003 and 2004, 37 were challenged in 2003 only, and 32 were challenged in 2004 only.

Our GnRH-challenge method stimulates a maximal testosterone response at 30 min, and levels return to baseline within 2 hours (Jawor et al. 2006). In our population, there are significant differences among the sampling periods described above in the increase of testosterone produced, indicating a gradual seasonal decline (Jawor et al. 2006, see also Results). When seasonal variation is held constant, individuals show repeatable differences in the magnitude of testosterone increases above initial levels (repeatability = 0.36) (Jawor et al. 2006).

Tail white measurement

To assess ornament size, digital photographs of individual tail feathers were obtained upon first capture of males in 2003 and 2004 (n = 90). Photographs were taken of the 3 – 5 feathers on the right side of the tail that have some degree of white coloration by placing a piece of black paper between the tail feathers (Wolf et al. 2004; Yeh 2004). If a
feather on the right was missing, the feather on the left was photographed in its place. So that the entire feather could be seen, body feathers were held to the side with the handler’s thumb. Photographs were taken from a standardized distance, using standardized lighting, and with a ruler for scale. Using the MetaView/MetaMorph image analysis program (Universal Imaging), the proportion of white area of each feather on the right side of the tail was measured. This proportion, between 0 and 1, was the tail white value for each feather. To obtain an individual’s tail white score, the values of the feathers on the right side of the tail were added (range 1.80 – 3.26) (Chapter 2, Wolf et al. 2004). Again, we used the tail white value of the corresponding feather on the left if a feather on the right was missing.

**Simulated territorial intrusions**

In order to test the assumption that GnRH-induced testosterone levels predict those naturally produced in response to social stimuli, we performed both GnRH challenges and simulated territorial intrusions (STI) on a small sample of 10 males in May – July 2005. GnRH challenges were performed on males captured using mist nets at their nest while feeding 6-day-old offspring. STIs were conducted 2 – 4 days later, after the nestlings had been collected (as part of a separate experiment) or the nest had been naturally depredated. All STIs were started before 0900. In each STI, we placed a caged captive male in the territory near the empty nest and played a 10 min CD recording of junco songs, which were recorded in our population > 10 years before. Songs were
played at a rate of 4 songs per min. This treatment stimulated approach and directed song by the territorial male. After 10 min, the recording was stopped, at which time two mist nets near the cage were unfurled in an attempt to capture the focal male. After 10 min of silence, the playback was restarted and remained on until the male was captured. A blood sample was taken immediately and plasma was reserved for hormone analysis. Captures occurred 13 – 86 min after the playback was restarted (mean 37.1 min).

**Testosterone assays**

The level of testosterone in each plasma sample was determined using enzyme-linked immunoassays (Assay Designs #901-065). Assay methods are described in detail elsewhere (Clotfelter et al. 2004). Approximately 2000 cpm of tritiated testosterone were added to each sample in order to calculate recoveries after 2 extractions with diethyl ether. Extracts were resuspended in 50 µl ethanol and diluted to 350 µl with assay buffer from the kit. From each reconstituted sample, 100 µl were used to determine recoveries, and duplicate 100 µl quantities were used in the EIA. Testosterone concentrations were determined with a 4–parameter logistic curve-fitting program (Microplate Manager; BioRad Laboratories, Inc.) and corrected for incomplete recoveries.

Samples from different years were run in different assays. In 2003-2004, intra-plate coefficients of variation ranged from 1 – 19% (mean 9%), and intra-plate variation was 20%. In 2005, intra-plate variation was 3 – 4% (mean 3%), and inter-plate variation was 8%. Within each dataset, we corrected for inter-plate variation by multiplying each
measurement by the grand mean of assay standards across all plates within the dataset and dividing by the plate mean of these standards.

**Statistical analyses**

To test for a relationship between testosterone production and tail white, it was necessary to correct for factors that may have influenced the response to GnRH (such as season, Jawor et al. 2006) as well as the non-independence of data points caused by repeated sampling. To this end, we used restricted maximum likelihood to fit linear mixed models. Mixed models allow tests of multiple fixed effects while allowing for structured random effects (Verbeke and Molenberghs 2000). In our models, the random portion accounted for repeated measures taken from a single individual as well as the structure of our sampling regime (8 total sampling periods, i.e. 4 seasonal stages in each of 2 years). Our model estimated the error variance-covariance matrix of the data with a first-order factor analytic structure (*SPSS 14.0 Command Syntax Reference*, SPSS Inc., Chicago, IL). The diagonal elements ($\lambda_i^2 + d$) of this matrix were estimates of the error variance at each sampling period, and the off-diagonals ($\lambda_i\lambda_j$) were estimates of the within-individual covariance between sampling periods. Thus, 36 error (co)variance components were described by 9 model parameters (8 $\lambda$ and 1 $d$). This covariance matrix was similar to one generated by a model that estimated all 36 components separately, but allowed for greater power due to the reduction in the number of parameters.
In the fixed portion of the models, we included tail white as a continuous predictor. We also included other variables that might explain variation in measures of testosterone. As categorical predictors, we included year, seasonal stage (early breeding A and B, nestling feeding, or late breeding), and a year × stage interaction; we included mass (g), and handling time (min, ln-transformed) as covariates (see Jawor et al. 2006 for discussion of these variables). We used type I (sequential) sums of squares, which allowed us to control for these effects before testing for covariation with tail white, which was entered into the model last. As dependent variables, we used initial and post-challenge testosterone (both ln-transformed), as well as the GnRH-induced increase in testosterone (ln post-challenge – ln initial).

To visualize individual covariation between testosterone and tail white, we calculated average testosterone measures for each individual. This was accomplished by fitting general linear models including all the fixed effects used in the mixed model (except tail white), as well as an individual term. Least-squares means for each individual were estimated from this model and plotted against tail white (averaged if an individual was measured in both years).

Because STI data were only collected at one seasonal stage and were not repeated within individuals, we used standard parametric statistics (paired t-test and Pearson correlation) on ln-transformed testosterone values. All analyses were performed using SPSS 14 for Windows.
Results

Tail white and GnRH-challenge response

We found a significant positive correlation between tail white and the magnitude of GnRH-induced testosterone increases (Table 3.1, Figure 3.2). This relationship held even when we did not control for other fixed effects ($b = 0.26, F_{1, 137.8} = 6.82, P = 0.010$). The relationships between tail white and both initial and post-challenge testosterone levels were non-significant (Table 3.1).

Other effects on GnRH-challenge response

For each testosterone measure, there were effects of year and/or seasonal stage (Table 3.1). The strongest effect was that of stage on post-challenge testosterone and GnRH-induced increase. As discussed in Jawor et al. (2006), this effect describes a decline of responsiveness to GnRH in later seasonal stages. Handling time had significant negative effects on all measurements of testosterone (Table 3.1). Heavier birds showed lower post-challenge testosterone and GnRH-induced testosterone increase (Table 3.1). Jawor et al. (2006) suggested that this effect may have derived from an effect of dosage or differences of activity correlated to testosterone production.
Table 3.1. Linear mixed models of testosterone before and after GnRH challenges. Fixed effects with $P < 0.05$ are highlighted in bold, effects with $0.10 < P < 0.05$ are shown in italics. Random effects (not shown) described and corrected for the pattern of residual (co)variance that arose from repeated sampling from individuals. See Methods for details of the models.

1. ln Initial testosterone

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>$F$</th>
<th>df</th>
<th>$b$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>11.00</td>
<td>1, 60.7</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Stage</td>
<td>1.64</td>
<td>3, 47.5</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>Year × stage</td>
<td>0.95</td>
<td>3, 41.4</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>Mass</td>
<td>1.07</td>
<td>1, 144.5</td>
<td>0.002</td>
<td>0.30</td>
</tr>
<tr>
<td>ln Handling time</td>
<td>7.70</td>
<td>1, 136.1</td>
<td>-0.09</td>
<td>0.006</td>
</tr>
<tr>
<td>Tail white</td>
<td>2.19</td>
<td>1, 139.7</td>
<td>-0.10</td>
<td>0.14</td>
</tr>
</tbody>
</table>

2. ln Post-challenge testosterone

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>$F$</th>
<th>df</th>
<th>$b$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>7.28</td>
<td>1, 83.0</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>Stage</td>
<td>18.82</td>
<td>3, 54.2</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>Year × stage</td>
<td>2.68</td>
<td>3, 42.7</td>
<td>0.059</td>
<td></td>
</tr>
<tr>
<td>Mass</td>
<td>14.21</td>
<td>1, 116.2</td>
<td>-0.10</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>ln Handling time</td>
<td>5.68</td>
<td>1, 114.8</td>
<td>-0.10</td>
<td>0.019</td>
</tr>
<tr>
<td>Tail white</td>
<td>1.24</td>
<td>1, 119.0</td>
<td>0.14</td>
<td>0.27</td>
</tr>
</tbody>
</table>

3. GnRH-induced increase (ln post-challenge – ln initial)

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>$F$</th>
<th>df</th>
<th>$b$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>3.25</td>
<td>1, 96.5</td>
<td>0.075</td>
<td></td>
</tr>
<tr>
<td>Stage</td>
<td>22.06</td>
<td>3, 73.1</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>Year × stage</td>
<td>2.17</td>
<td>3, 65.5</td>
<td>0.101</td>
<td></td>
</tr>
<tr>
<td>Mass</td>
<td>27.81</td>
<td>1, 115.0</td>
<td>-0.10</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>ln Handling time</td>
<td>4.65</td>
<td>1, 101.3</td>
<td>-0.07</td>
<td>0.033</td>
</tr>
<tr>
<td>Tail white</td>
<td>5.17</td>
<td>1, 125.8</td>
<td>0.23</td>
<td>0.025</td>
</tr>
</tbody>
</table>
Figure 3.2. Relationship between tail white and magnitude of testosterone increase after GnRH challenge (ln post-challenge testosterone – ln initial testosterone). Data points represent individual means. Testosterone measures were adjusted for stage, year, mass, and handling time as described in Methods.

Across all measurements, post-challenge testosterone was positively correlated with initial testosterone ($r_{176} = 0.54, P < 0.0001$) and GnRH-induced increase ($r_{176} = 0.82, P < 0.0001$), but initial testosterone was not significantly correlated with GnRH-induced increase ($r_{176} = -0.04, P = 0.60$).
**Figure 3.3.** Age-specific relationship between tail white and magnitude of testosterone increase after GnRH challenge (ln post-challenge testosterone – ln initial testosterone). First-year males are shown as filled circles, and older adults are shown as open circles. As in Figure 3.2, data points represent individual means. Testosterone measures were again adjusted for stage, mass, and handling time, but not for year using a general linear model. Data were summarized separately for each age class, so there are more data points than in Figure 3.2 (4 individuals were measured as yearlings in 2003 and second-year adults in 2004).
Age effects

Jawor et al. (2006) reported no significant effect of age on initial testosterone or GnRH-induced testosterone increases; therefore, we did not include it in our statistical models here. Including age in our models (either as a linear effect, in yr, or comparing age classes, i.e. first-year breeders to older adults) did not alter any of the relationships between testosterone and tail white. However, because mating may differ between first-year breeders and older birds (Peters et al. 2006), we added age class as a main effect and nested tail white within age class to test whether the relationship between tail white and GnRH-induced increases differed among age classes. This type of model estimates age-specific slopes, and is equivalent to including an interaction term in an analysis of covariance (Engqvist 2005). In this model, there was no significant main effect of age class ($F_{1,130.5} = 1.10, P = 0.30$), but the effect of tail white significantly differed between age classes ($F_{1,101.2} = 5.50, P = 0.005$). Specifically, the slope for first-year adults was strong and positive ($b = 0.48, t_{82.4} = 3.31, P = 0.001$), whereas there was no significant relationship in older adults ($b = 0.05, t_{130.3} = 0.32, P = 0.75$; see Figure 3.3).

Simulated territorial intrusions

Territorial males presented with STIs displayed significantly elevated testosterone levels (paired $t$-test, $t_9 = 9.82, P < 0.0001$, Figure 3.4a). Further, the levels produced during STIs
Figure 3.4. Testosterone in response to a simulated territorial intrusion. (a) Testosterone shows a significant increase following a simulated territorial intrusion. The line represents the median, the shaded box represents the interquartile range, and the bars represent the range. (b) Magnitude of the hormonal response to the intrusion is predicted by the response to a GnRH challenge. Line of best fit from a least-squares regression is presented for visualization.
were highly correlated with those produced by GnRH challenges, indicating that response to a GnRH challenge is a good predictor of a male’s physiological response to a naturally occurring social stimulus ($r_s = 0.68, P = 0.03$, Figure 3.4b).

**Discussion**

Our results show that males with larger tail white ornaments produce larger short-term testosterone increases in response to an injection of GnRH, the hypothalamic hormone that leads to the secretion of testosterone. There was no significant relationship between plumage and testosterone measured before injection. Testosterone levels produced in response to GnRH were highly correlated with those produced during territorial behavior, suggesting that GnRH injections may be used to predict hormonal response to natural challenges from a conspecific. To our knowledge, our study is the first to associate short-term testosterone increases with variation in a sexual signal in a natural population. However, a recent study of captive white-throated sparrows (*Zonotrichia albicollis*) reports that alternative plumage morphs, which differ in aggressiveness and reproductive strategy, also differ in testosterone produced in response to GnRH challenges (Spinney et al. 2006). This suggests that the association demonstrated here between attractive plumage and short-term testosterone increases may also occur in other species. Combined with earlier studies in juncos, our results suggest that more attractive, socially dominant males have the physiological ability to invest more in mating effort. To the extent that individual variation in the ability to increase testosterone predicts male quality, the
correlation with testosterone production may act to enforce the signal honesty of tail white.

**Proximate mechanisms**

The mechanism responsible for the correlation between tail white and short-term testosterone elevation is unknown, but the effect is unlikely to have arisen from a direct physiological effect of adult testosterone levels on tail white. Tail white size is a relatively static trait and only changes when a new feather grows (Hill et al. 1999). This typically occurs during the annual molt. Testosterone levels are extremely low during molt in juncos, and experimentally elevating testosterone inhibits molt (Nolan et al. 1992). It is unknown whether juncos produce testosterone increases in response to social interactions during molt, and in turn, whether such short-term changes in testosterone may influence developing feathers. In house sparrows, where testosterone does not inhibit molt, such a mechanism has been suggested for socially induced changes in bib size (McGraw et al. 2003). Another possibility is that testosterone may have effects on tail white during early development. Testosterone produced by developing offspring or deposited by the mother in the yolk may affect the development of feather follicles, leading to developmental effects that persist throughout life. Early developmental effects of testosterone on plumage have also been demonstrated in house sparrows (Strasser and Schwabl 2004).
Hormones and honest signals

The correlation between tail white and testosterone production may be in part genetic, arising from genes that are pleiotropic or in linkage disequilibrium. For a genetic correlation to exist between two traits, both must be heritable. Tail white does show heritable variation (Chapter 2), but the heritability of testosterone production in juncos (as in most wild species) is unknown (Adkins-Regan 2005). However, GnRH-challenge response has been shown to be repeatable (a necessary condition for heritability) in juncos and heritable in domestic mammals (Robison et al. 1994; Jawor et al. 2006). If the correlation is genetic, one possibility is that both testosterone production and tail white are influenced by the genetic component of overall condition (Rowe and Houle 1996). Such a relationship has been suggested to account for correlations between sexually selected traits in a recent study of Drosophila bipectinata (Cooperman et al. 2007).

Another possibility is that the correlation we measured arose from the effect of a common environmental factor, such as diet quality (McGlothlin et al. 2007). Because of the different time scales on which tail white and testosterone may change, however, such an environmental effect would have to be either long-lasting or correlated across seasons (Hill et al. 1999). Quantitative genetic studies would be necessary to determine the relative importance of genetic and environmental effects.

Sexual selection

From an ultimate perspective, selection may be responsible for the observed correlation between tail white and short-term testosterone elevation. Particularly, correlational
selection, which arises when traits interact in their effects on fitness and may maintain or strengthen genetic correlations between traits, is likely to be important (Cheverud 1982; Lande and Arnold 1983; Phillips and Arnold 1989; Brodie 1992; Sinervo and Svensson 2002). Recently, correlational selection has been suggested to play a role in the evolution of signal honesty by linking ornaments to traits associated with male quality (Chapter 2, LeBas et al. 2003; Bentsen et al. 2006; Getty 2006).

Earlier studies clearly predict that tail white and testosterone are likely to interact in their effects on mating success. Males display the white patch on their outer tail feathers while courting females and during escalated aggressive encounters with other males (Balph et al. 1979; Nolan et al. 2002) (Figure 3.1). Experimentally enhancing tail white in male juncos increases their attractiveness to females, and experimentally enhanced testosterone increases both the frequency of tail white displays and the general attractiveness of males (Enstrom et al. 1997; Hill et al. 1999). Thus, to the extent that individual variation in testosterone production predicts courtship behavior, it may interact with tail white to increase the attractiveness of a male’s display; specifically, testosterone may increase display frequency while more tail white increases its size. In addition, the capability of males with more tail white to elevate testosterone to a greater degree may provide a strong advantage during male-male competition. In another study, we found that males with high maximum testosterone levels following GnRH injections also display more intense territorial aggression (Chapter 4). Because tail white may also act as a signal of male dominance (Balph et al. 1979; Holberton et al. 1989), producing a larger increase in testosterone may allow a male with a larger tail white patch to effectively reinforce its signal with aggression.
Experimentally enhanced testosterone also increases other components of mating effort (e.g. song rate, extra-territorial forays) while decreasing parental effort (e.g. feeding nestlings) and survival (Ketterson and Nolan 1992, 1999; Enstrom et al. 1997; Raouf et al. 1997; Schoech et al. 1998; Ketterson et al. 2001; Reed et al. 2006). Initial evidence in juncos suggests that the effects demonstrated in implantation studies also apply to individual variation; notably, short-term testosterone increases seem to be the best predictors of individual variation in the trade-off between mating effort and parental effort. Males that produce higher GnRH-induced increases feed their offspring less, and are presumably spending this time singing to attract additional mates (Chapter 4).

Behaviors associated with the mating effort/parental effort trade-off are likely to interact with tail white in their effects on fitness. For example, it is likely beneficial for males with whiter tails to spend more time searching for extra-pair mates, because they are more likely to be successful at attracting females they encounter. Males with less tail white would benefit from spending relatively more time caring for offspring because they are unlikely to attract extra-pair mates.

Consequently, the full potential benefits of tail white are likely to be realized only if tail white is associated with testosterone-mediated investments in mating effort and display. Put another way, males should benefit from producing behavior appropriate to their signal. This potential for interaction between tail white and behaviors mediated by testosterone in their effects on mating success suggests that correlational sexual selection may be acting to maintain the correlation observed in this study. Correlational sexual selection acting on tail white and body size has already been measured in this population (Chapter 2). Specifically, larger males with more tail white achieved the highest mating
success. However, that study did not measure natural variation in testosterone production, so further measurements of selection are necessary to test the hypothesis presented here.

**Reliable signaling**

The correlation between tail white and natural ability to elevate testosterone should allow competitors and potential mates to assess a male’s likely behavior by the size of his tail-white patch, making tail white a reliable signal of behavioral tendencies. Although the evidence presented here is not sufficient to establish tail white as an honest signal of male quality, prior knowledge about the role of testosterone suggests that this scenario is quite plausible. As described above, males are not expected to benefit equally from increasing mating effort. Increasing testosterone production and mating effort comes at a cost to parental care, immune function, and survival (Ketterson et al. 1992; Schoech et al. 1998; Casto et al. 2001; Reed et al. 2006), and therefore males cannot afford to increase mating effort unless they are likely to benefit. Optimal allocation to mating effort, parental effort, and self maintenance, and thus optimal testosterone production, is likely to depend on male quality (i.e. the amount of resources available).

Further studies are necessary to confirm whether variation in testosterone increases is related to male quality as predicted. Quality is difficult both to define and to measure, but body condition or mass is often used as an estimate (e.g. Keyser and Hill 2000; Jawor and Breitwisch 2004). Our results, however, show a negative relationship between body mass and GnRH-stimulated testosterone increases (although this pattern is likely to be a
simple effect of dosage, 2006). One complication with using body mass as a proxy for
quality is that mass itself may be affected by life-history trade-offs. In the present case,
testosterone may increase activity, and in doing so, may decrease body mass (Ketterson
et al. 1991). Perhaps the most relevant measure of male quality in the context of sexual
selection is genetic quality, which is operationally defined as the breeding value for total
fitness (Rowe and Houle 1996; Hunt et al. 2004). Measuring associations between
sexually selected traits and male genetic quality requires extremely large sample sizes
(Qvarnström et al. 2006). The variation inherent to hormone measurements may add extra
difficulty to measuring such a relationship with testosterone. However, the fact that
artificially enhanced testosterone levels lead to increased male fitness suggests that such a
relationship is plausible (Reed et al. 2006).

Age-dependent effects

Although older males tend to have more tail white (Wolf et al. 2004; Yeh 2004), average
magnitude of testosterone elevation was not higher in older adults (see Jawor et al. 2006
for further discussion). However, we found that the strength of the correlation between
the two changed with age. Specifically, the relationship was strongly positive in first-year
adults, but not significantly different from zero in older adults. We stress that this finding
does not negate the correlation between tail white and testosterone increases at the
population level. Juncos should still be able to predict the likely behavior of an unfamiliar
male, irrespective of age, based on the size of his tail white patch. That is, when the
population consists of a mixture of old and young adults, tail white should remain “honest on average” (Kokko 1997). However, the extent to which tail white is informative may depend on the age structure of the population.

The relationship between plumage and testosterone may be weaker in older males because selection pressures differ for the two age classes. One possibility is that correlational sexual selection may act more strongly on younger birds, favoring a stronger relationship between testosterone production and tail white. This may occur because the decision of whether to divert energy away from mating effort and toward parental effort is a more crucial for younger birds. Older male juncos seem to have a general mating advantage; they often retain territories and mates from previous years (Nolan et al. 2002) and are more likely than yearlings to obtain extra-pair mates, suggesting that they may be more attractive to females in general (Reed et al. 2006). Because of competition from older males, it may be beneficial for first-year males with smaller ornaments to invest more in survival in hopes of higher mating success as an older adult, whereas first-year males with large ornaments may be able to succeed in competition with older males.

Another, not mutually exclusive, explanation for why older males show a decreased correlation is that they have already experienced more episodes of survival selection. Stabilizing selection acting on tail white, testosterone production, or both could eliminate extremes from the population, weakening the correlation between the variables in older cohorts.

Interestingly, Peters et al. (2006) showed a similar pattern in blue tits. In that species, old and young males show opposite relationships between plumage and
testosterone corresponding to opposite relationships between plumage and mating success. The results of that study, combined with those presented here, suggest that future work should focus on how selection shapes the age-dependence of mating strategies and the expression of sexually selected traits.

**Conclusion**

In summary, our results indicate phenotypic integration of static, sexually selected morphology with a dynamic physiological mechanism known to mediate allocation to mating effort. This relationship suggests that correlational selection may be important for understanding the co-expression of hormones and sexual signals, and potentially, the maintenance of signal reliability (Getty 2006). Together with previous studies in this species (Enstrom et al. 1997; Raouf et al. 1997; Hill et al. 1999; Ketterson and Nolan 1999; Ketterson et al. 2001; Reed et al. 2006), our results provide links among hormone levels, ornamental traits, and behavior at the level of individual variation, the raw material on which selection can act.
Acknowledgments

We thank Jackie Gaudioso, Nicki Gerlach, Annie Lindsay, Dawn O’Neal, Katie Pavlis, Sara Schrock, Eric Snajdr, Peter Stevens, Charles Ziegenfus, and Devin Zysling for assistance in the field, Leif Engvist and Emilia Martins for statistical advice, and Butch Brodie III, Greg Ball, Greg Demas, Britt Heidinger, Mary-Louise Maas, Trevor Price, Volker Rudolf, Scott Sakaluk, and Dale Sengelaub for discussions and comments on the manuscript, and Mountain Lake Biological Station (Henry Wilbur, Director and Eric Nagy, Associate Director) and Mountain Lake Hotel for facilities and permission to work on their land. This research was supported by a National Science Foundation Doctoral Dissertation Improvement Grant (0508693) and grants from the American Ornithologists’ Union, Indiana University, Mountain Lake Biological Station, Sigma Xi and the Wilson Ornithological Society to JWM, a National Science Foundation Research Experience for Undergraduates award to JLP, and National Science Foundation grants (0216091, 0519211) to EDK. This research adhered to the Association for the Study of Animal Behaviour /Animal Behavior Society Guidelines for the Use of Animals in Research, the legal requirements of the United States of America, the State of Indiana, and the Commonwealth of Virginia, and the guidelines of the Animal Care and Use Committees at Indiana University and the University of Virginia.
CHAPTER FOUR

Natural variation in a testosterone-mediated trade-off between mating effort and parental effort

JOEL W. MCGLOTHLIN

JODIE M. JAWOR*

ELLEN D. KETTERSON

Submitted to American Naturalist

*Department of Biological Sciences, University of Southern Mississippi, Hattiesburg, MS
Male birds frequently face a trade-off between acquiring mates and caring for offspring. Hormone manipulation studies indicate that testosterone often mediates this trade-off, increasing mating effort while decreasing parental effort. Little is known, however, about individual covariation between testosterone and relevant behavior on which selection might act. Using wild male dark-eyed juncos (*Junco hyemalis*), we measured individual variation in testosterone levels before and after standardized injections of gonadotropin-releasing hormone (GnRH). GnRH challenges have been shown to produce short-term testosterone increases that are similar to those produced naturally in response to social stimuli, repeatable in magnitude, and greater in males with more attractive ornaments. We compared these testosterone increases to behavioral measures of mating and parental effort (aggressive response to a simulated territorial intrusion and nestling feeding, respectively). Males that showed higher post-challenge testosterone displayed more territorial behavior, and males that produced higher testosterone increases above initial levels displayed reduced parental behavior. Initial testosterone levels were positively but non-significantly correlated with aggression, but did not predict parental behavior. These relationships suggest that natural variation in testosterone, specifically the production of short-term increases, may underlie individual variation in the mating effort/parental effort trade-off. We discuss the implications of these results for the evolution of hormonally mediated trade-offs.
Introduction

Life-history trade-offs arise when traits that contribute to fitness are inversely linked and are thus inhibited from evolving independently (Stearns 1992; Roff 2002). When a trade-off is present, multiple beneficial traits cannot be maximized simultaneously, at least in the short term, and theory predicts that selection will favor the optimal combination of traits within the constraints imposed by the trade-off (Stearns 1992; Roff 2002; Roff and Fairbairn 2007). One of the most common trade-offs in animals involves mating effort (the amount of energy, time, or other key resources invested in competing for mates) and parental effort (the amount of resources invested in rearing offspring) (Magrath and Komdeur 2003). In many species, such as biparental birds with extra-pair fertilizations, this trade-off is particularly important because mating effort and parental care may overlap in time. Hence, investment in one activity usually requires reduced investment in the other.

Trade-offs are often mediated by physiological factors such as hormones, and understanding their physiological basis may provide insight into life history evolution (Stearns 1992; Finch and Rose 1995; Sinervo and Svensson 1998; Ketterson and Nolan 1999; Zera and Harshman 2001; Ricklefs and Wikelski 2002; Adkins-Regan 2005; Hau 2007). Hormonal manipulations provide a powerful means to dissect the mechanistic basis of life-history trade-offs and to experimentally test for effects on fitness, especially when they are used in long-term studies of natural populations (Ketterson et al. 1996; Reed et al. 2006). However, to understand the evolution of hormone-related trade-offs more fully, we must also consider naturally occurring variation among individuals, which
must be present in order for selection to act (Adkins-Regan 2005). Such variation may lie in the strength of the hormonal signal, the sensitivity to the individual to the hormones, or both. Empirically, a logical first step is to focus on hormone concentrations in the circulation, which are often much easier to measure than individual sensitivity.

Examining individual variation in the physiological basis of trade-offs involving male mating effort is of considerable interest, because theory predicts that males should invest differentially in relation to their quality (Nur and Hasson 1984; Getty 1998, 2006). The potential benefit of increasing mating effort at the expense of parental effort is often great, as total reproductive success tends to increase with mating success in males (Trivers 1972; Arnold 1994; Queller 1997; Wade and Shuster 2002; Shuster and Wade 2003). However, males must compete among themselves for a limited number of mating opportunities, and one male’s mating success necessarily comes at the expense of another male. As a consequence, the variance in male reproductive success increases as the maximum number of mates increases (Shuster and Wade 2003). Because all males cannot succeed equally at obtaining multiple mates, males may differ in their optimal investment in mating effort (Trivers 1972; Nur and Hasson 1984; Getty 1998, 2006). Males that are more likely to succeed in obtaining mates because, for example, they possess a more attractive ornament, would benefit from increased investment in mating, whereas males that are less likely to be successful at mating may benefit more from investing in parental care. Selection may thus maintain variation in the resolution of this trade-off, as well as its covariation with male attractiveness or quality (Getty 2006; Roff and Fairbairn 2007).

In birds, the trade-off between mating effort and parental effort appears to be mediated, at least in part, by the steroid hormone testosterone (Ketterson and Nolan 1992,
Many studies that have used experimental elevation of testosterone by means of subcutaneous implants have shown increases in mating behavior and decreases in parental behavior (e.g. Silverin 1980; Wingfield 1984; Hegner and Wingfield 1987; Dittami et al. 1991; Ketterson et al. 1992; Raouf et al. 1997; Van Roo 2004; but see Hunt et al. 1999, Van Duyse et al. 2000, 2002, Lynn et al. 2002, 2005). However, it has been difficult to study individual-level variation in testosterone and the behaviors it mediates, perhaps because testosterone varies so much within individuals (Adkins-Regan 2005). In many songbirds, testosterone levels show temporal variation on both relatively long-term (seasonal) and short-term scales (Wingfield et al. 1990). Short-term changes are particularly interesting, because in many species, they are induced by social stimuli and occur during the production of mate-acquisition behavior, such as territorial aggression and courtship (Harding 1981; Moore 1983; Wingfield 1985; Wingfield et al. 1990, 2001; Pinxten et al. 2003; but see Van Duyse et al. 2004; Landys et al. 2007; Lynn et al. 2007). Because of this association with behavior, variation in transient testosterone elevations may be more relevant to the mating effort/p parental effort trade-off than baseline circulating testosterone. To our knowledge, no studies have examined this relationship.

We assessed natural covariation between testosterone and behavior in a songbird, the dark-eyed junco (Junco hyemalis). A long-term implantation study has shown that testosterone generally increases mating effort and mating success, while decreasing parental behavior (Ketterson et al. 1992; Enstrom et al. 1997; Raouf et al. 1997; Cawthorn et al. 1998; Schoech et al. 1998; Reed et al. 2006). Natural testosterone levels show long-term and short-term variation, with males transiently increasing testosterone
levels during territorial interactions (Chapter 3, Ketterson and Nolan 1992; Jawor et al. 2006). The magnitude of short-term testosterone increases can be measured using gonadotropin-releasing hormone (GnRH) challenges, which are simple bioassays that are often used to measure the responsiveness of the hypothalamo-pituitary-gonadal (HPG) axis that regulates testosterone production (e.g. Millesi et al. 2002). Testosterone levels produced by male juncos in response to GnRH challenges are correlated with those produced during territorial interactions, and the magnitude of the testosterone increase varies among individuals (Chapter 3, Jawor et al. 2006). Further, males with larger plumage ornaments (a white patch on the tail) produce larger testosterone increases when injected with GnRH (Chapter 3).

Following from these results, we predicted that variation in the responsiveness of the HPG axis, and thus the ability to produce short-term testosterone increases, might account for natural variation in relative allocation to mating effort versus parental effort. To test this prediction, we measured natural covariation between behavior (as a proxy for effort) and the response to GnRH challenges. As a measure of mating effort, we assessed aggression during simulated territorial intrusions and compared this behavior to the results of GnRH challenges performed during early breeding. Territorial aggression is a major component of competition for mates in birds and is thus likely to contribute to variation in mating success. In two separate breeding seasons, we assessed nestling feeding rate as a measure of parental effort, comparing it to GnRH challenges performed within one day of the behavioral observations. Male juncos do not incubate, and thus nestling feeding represents their major contribution to parental care. If variation in the capacity of the HPG axis to generate acute increases of testosterone underlies variation in
the trade-off between mating effort and parental effort, we expected to find that response to GnRH challenges should correlate positively with aggressive behavior and negatively with parental behavior.

**Methods**

*Study species and general methods*

We studied a wild breeding population of Carolina dark-eyed juncos (*J. h. carolinensis*) near University of Virginia’s Mountain Lake Biological Station in Giles County, VA (37°22’N, 80°32’W). Nolan et al. (2002) provide a detailed account of junco breeding biology. Briefly, at the beginning of the breeding season (late March – early April) male juncos establish and defend all-purpose territories. Females build nests and begin to lay eggs in late April. Females incubate clutches of 3 – 5 eggs for an average of 12 days. Both sexes defend the nest from predators, feed nestlings after hatching for an average of 12 days, and feed fledglings after they leave the nest. Juncos in our population repeatedly renest if nests are lost, and attempt additional nests following nest success (up to 3 successful nests in a single season). Extra-pair fertilizations are common and may account up to 56% of young (Nolan et al. 2002).

The population was censused at the beginning of each breeding season by capturing adults in baited mist nests and Potter traps, the locations of which remained the same each year. At each capture, standard morphometric measurements were obtained.
from each individual. If a bird had not been captured previously, it was marked with a U.S. Fish & Wildlife Service leg band and a unique set of plastic color bands so that it could be identified from a distance. Sex was determined using cloacal protuberance (males) or brood patch (females) development when possible; otherwise, larger birds were classified as males. Such an assignment was later confirmed by sexual development and/or behavior.

When birds began nesting, we attempted to find all nests in our study area using systematic searches and by observing focal bird behavior. Once located, nests were visited every 1 – 3 days thereafter to monitor progress. We assigned each nest to a pair of adults based on their behavior at the nest.

**Hormone sampling**

To assess natural variation in circulating testosterone and the sensitivity of the HPG axis, we used intramuscular injections of GnRH. Such GnRH challenges induce the pituitary to release luteinizing hormone into circulation, which in turn stimulates the testes to release testosterone. In male juncos, GnRH challenges induce peak testosterone levels at 30 min post-challenge, which return to baseline after 2 hr (Jawor et al. 2006). Importantly, testosterone levels produced after a GnRH challenge are correlated with natural increases in testosterone produced in response to a territorial intruder (Chapter 3).

After a bird was captured, it was returned to a central laboratory where a blood sample (~100 μl) was taken from the wing vein to measure initial testosterone level. We
measured handling time (in min), which was controlled in the statistical models, as the time elapsed between capture and the collection of this blood sample. Afterward, 50 µl of a solution containing 1.25 µg of chicken GnRH-I (Sigma L0637/American Peptide 54–8–23) dissolved in 0.1 M phosphate-buffered saline solution (PBS) was injected in the left pectoralis major. The bird was immediately placed into a holding bag. Exactly 30 min after the injection, a second (~100 µl) blood sample was taken to measure post-GnRH challenge testosterone. Blood samples were centrifuged and the plasma fraction was reserved and frozen at -20°C until assayed. The sampling regime for these challenges differed among years according to the type of behavior being measured (see relevant sections below).

**Territorial aggression**

In 2006, we performed GnRH challenges on males caught at random during the early breeding season (12 April – 17 May). Because the intensity of GnRH-challenge response may decrease within males as the breeding season progresses (Jawor et al. 2006), we attempted to perform 2 GnRH challenges on each male found on our study site to obtain an average response for each male. For a given male, the second challenge was performed 6 – 29 days (mean 12.8) after the first. In total, we performed 173 GnRH challenges on 114 different males. For 36 males, we were able to locate territories and measure territorial aggression. Twenty-one of these males had previously received two GnRH challenges, separated by 7 – 22 days (mean 13.3), and 15 had received a single
challenge. For a given male, behavioral measurements were collected 9 – 40 days (mean 23.7) after a male’s first GnRH challenge and 3 – 26 days (mean 13.3) after its second challenge (if any). Handling time for these samples ranged from 8 – 96 min (mean 35).

We used observations of behavior, often stimulated by brief song playback (1 – 3 min) in order to map the territories of males. Although the period of territory mapping overlapped temporally with both hormone sampling and measurement of territorial behavior, these activities were never conducted in the same part of the study site on the same day.

Simulated territorial intrusions were conducted between 29 April and 29 May to measure territorial aggression. In each intrusion, we placed a captive lure male in a small cage in the estimated center of the focal male’s territory. The captive lures \( n = 8 \) used in this study were captured from areas that were at least 3 km away from our study site, so they were unlikely to be familiar or related to focal males. The cage was covered with a cloth until the trial began. Two nylon ropes with plastic flagging placed at distances of 5 and 10 m, which were attached to the bottom of the cage, were stretched along the ground in opposite directions and used to judge distance of the focal male from the cage. A portable compact disc player (Duraband CD-855) attached to a battery-powered speaker (Radio Shack 40–1441) was placed directly next to the cage.

At the beginning of each trial, the speaker was set to full volume, the player was started, and the cover was removed from the captive male’s cage. The two observers then retreated to a location at least 15 m away from the cage. After 5 min of silence, a 15-min recording of junco long-range song (Titus 1998) was played. The recording consisted of 5 different song types, recorded in our study population \( \geq 10 \) yr before. Each song type was
repeated 9 times at a rate of 6 songs min\(^{-1}\) (for a total of 45 songs in 7.5 min), and then this series was played a second time (90 songs in 15 min). Audacity 1.2.3 for Windows (http://audacity.sourceforge.net) was used to compile the composite recording, remove background noise, and equalize volume of different song types. The sound power level was 92 dB, measured at 1 m using a sound level meter (Radio Shack 33–2050). This sound level is comparable to a junco singing long-range song (Nolan et al. 2002).

We recorded 4 variables related to territorial aggression. First, \textit{latency} was the amount of time (in s, recorded using a stopwatch) between the beginning of the song playback and when the male was first seen or heard. Seven males that approached the lure before the song began were assigned zero latency. We used a second stopwatch to record the \textit{time spent within 5 m} of the cage (in s). We counted the number of \textit{flyovers} (a flight directly over the lure’s cage) and the number of long-range \textit{songs} produced. In each trial, a single observer (JWM) watched the bird using binoculars and noted behavior, while a second operated the stopwatches, recorded the behavior on a datasheet, and helped locate the bird if needed.

If a male did not appear in response to the stimuli, the trial was not used. Trials were abandoned or discarded if the responding male could not be identified using its color bands. We did not perform simulated territorial intrusions on males that were known to be feeding nestlings, and we excluded data from 4 males known to be feeding fledglings because of the dramatic changes in both behavior and home range at that stage of reproduction (Nolan et al. 2002). Of the remaining trials conducted on 36 males, 4 of the males had mates that had not yet produced an egg, 5 had mates that were incubating (male juncos do not incubate), 8 had recently lost nests to predators, and 19 were of
unknown nesting stage. Territorial behavior (first principal component, see Statistical Analyses) did not differ statistically among these groups ($P = 0.65$).

**Parental behavior**

In 2003 – 2004, we measured parental behavior and GnRH-challenge response in 24 males that were feeding nestlings (12 in 2003, 13 in 2004; one male was measured in both years). Each brood had 2 – 5 nestlings (mean 3.5). Family size was reduced in five of the nests observed because we collected an egg for steroid analysis as part of another study.

We temporarily removed a male’s mate when measuring its parental behavior in order to control for potential interactions between members of the pair (Clotfelter et al. 2007). These females were caught using a mist net in front of the nest. After catching the female, we counted and weighed the nestlings and placed a video camera on a tripod near the nest. Females were held in the laboratory in individual cages and provided with food and water *ad libitum*. Male behavior at the nest was recorded for 4 hr, at which time the female was returned to the site of capture and released. Recordings of parental behavior were made on either day 6 or day 7 after hatching (see below), and all recordings were begun between 0600 and 1100. Videotapes were scored for number of visits to the nest with food, and divided by total recording time to calculate feeding rate.

To measure GnRH-challenge response in relation to parental behavior, we caught males using the same mist net arrangement used to capture females. Most males were
caught the day after their behavior was recorded (n = 20). However, if we were unable to
catch the female on day 6 and instead caught the male, the male was given a GnRH
challenge on day 6 and assessed for parental behavior the following day (n = 5). We did
not include day of capture in our statistical analyses; however, controlling for this factor
capture did not affect our results.

Males were returned to the laboratory to receive GnRH challenges according to
the procedure described above. Handling time ranged from 14 – 217 min (mean 60).
Following the challenge, the male was temporarily housed in a small cage (approximately
4 hr) and provided with food (white millet and mealworms) and water while female
parental behavior was being measured as part of another study.

**Testosterone assays**

Plasma collected from GnRH challenges was analyzed for testosterone measured using
an enzyme immunoassay (EIA) kit (Assay Designs, Inc., #901–065). Assay methods are
described in detail elsewhere (Clotfelter et al. 2004). Approximately 2000 cpm of tritiated
testosterone were added to each sample in order to calculate recoveries after 2 extractions
with diethyl ether. Extracts were resuspended in 50 μl ethanol and diluted to 350 μl with
assay buffer from the kit. From each reconstituted sample, 100 μl were used to determine
recoveries, and duplicate 100 μl quantities were used in the EIA. Testosterone
concentrations were determined with a 4–parameter logistic curve-fitting program
(Microplate Manager; BioRad Laboratories, Inc.) and corrected for incomplete recoveries.

Samples from different years were run in different assays. In 2003 – 2004 assays, the intra-plate coefficient of variation (calculated from standard samples of known concentration), ranged from 1 – 19% (mean 9%), and inter-plate variation was 20%. In the 2006 assay, intra-plate variation ranged from 4 – 19% (mean 12%), and inter-plate variation was 23%. To correct for inter-plate variation, we multiplied each measurement by the grand mean of standards across all plates within a given dataset, divided by the plate mean of standards.

Within a given year, multiple plasma samples from the same individual were analyzed on the same plate. Individuals were randomly assigned to plates, and samples within a plate were randomly assigned to wells.

**Statistical analyses**

All statistical analyses were performed using SPSS 14.0 for Windows. To test for relationships between testosterone and behavior, we used restricted maximum likelihood to fit linear mixed models (Verbeke and Molenberghs 2000). Such models allow for simultaneous estimation of structured random effects (an error variance-covariance matrix) and tests of fixed effects. For the analysis of territorial aggression, we used a compound symmetrical covariance structure for the random portion of the model (*SPSS 14.0 Command Syntax Reference*, SPSS Inc., Chicago, IL). This model fits estimates of
two parameters, error variance across subjects and error covariance within subjects (which is analogous to repeatability). The latter term allowed us to account for repeated testosterone measurements from an individual. In the analysis of parental behavior, we fit a diagonal covariance structure, which estimated separate error variances for the two years of the study (SPSS 14.0 Command Syntax Reference, SPSS Inc., Chicago, IL). Because only one individual was sampled for parental behavior in both years, we could not estimate the within-subjects covariance term.

The relationships between behavior and initial (pre-GnRH challenge) testosterone, post-challenge testosterone (both natural log transformed), and GnRH-induced testosterone increase (natural log post-challenge testosterone – natural log initial testosterone) were tested in separate analyses. In order to control for variables that may have affected them (Jawor et al. 2006), testosterone measurements were used as the dependent variables in our mixed models. We used Type I (sequential) sums of squares, which allowed us to control for these variables before testing for covariation with the behavior of interest. For both datasets, we included handling time (min, natural log transformed), day of year, and mass (g) as continuous fixed effects. For the analysis of parental behavior, year was also entered as a categorical fixed effect. Behavioral measurements were the last fixed effect entered into each model.

In order to visualize the relationships between testosterone and behavior, we calculated adjusted values of all testosterone measures. For the territorial aggression dataset, we used general linear models that included handling time, day of year, and mass as well as an individual term to generate individual least squares means. For the parental care dataset, we used models that included year, handling time, day of year, and mass.
Adjusted values were calculated by adding the residual value for each individual to the overall mean for each testosterone value.

Because we were interested in generalized territorial aggression, and because territorial behaviors were intercorrelated, we extracted a single principal component to describe response to simulated territorial intrusions. The first principal component, which described 47% of variance, was loaded as in Table 1, and was used as our measurement of aggression in the statistical analyses.

**Table 1:** Loadings of the first principal component of territorial behavior measured in simulated territorial intrusions

<table>
<thead>
<tr>
<th>Behavior</th>
<th>PC1 Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>latency</td>
<td>-0.81</td>
</tr>
<tr>
<td>time spent within 5 m</td>
<td>0.74</td>
</tr>
<tr>
<td>flyovers</td>
<td>0.61</td>
</tr>
<tr>
<td>songs</td>
<td>0.53</td>
</tr>
</tbody>
</table>
Results

Territorial behavior

In birds for which territorial behavior was measured, mean initial testosterone (± 1 S. E.) was 1.85 ± 0.182 ng ml⁻¹, mean post-challenge testosterone was 7.17 ± 0.564 ng ml⁻¹, and the mean GnRH-induced increase was 5.31 ± 0.510 ng ml⁻¹ (n = 57). The natural log transformed values were correlated as follows: initial – post, \( r = 0.51 \); initial – increase, \( r = -0.41 \); post – increase, \( r = 0.58 \) (\( P \leq 0.002, n = 57 \)). On average, focal males responded to the simulated territorial intrusions in 115 ± 29.1 s, spent 542 ± 44.7 s within 5 m of the cage, performed 2.8 ± 0.46 flights over the cage, and sang 54 ± 6.5 songs (n = 36).

The first principal component of territorial behavior was positively related to post-challenge testosterone (Figure 4.1, Table 4.2). In other words, males that produced higher absolute levels of testosterone in response to the GnRH challenge tended to show shorter response latency, spent more time spent within 5 m of the lure, performed more flyovers, and produced more songs. Initial testosterone levels showed a trend toward a relationship with aggression, but there was no significant relationship with GnRH-induced increase (Figure 4.1, Table 4.2). Inspection of Figure 4.1 suggests that this pattern may have been driven by the two individuals with the highest adjusted initial testosterone levels. These individuals were only sampled once, and produced relatively low GnRH-induced increases (0.33 and 0.57), suggesting that their testosterone levels may have been elevated when they were captured.
Figure 4.1. Relationships between the first principal component of territorial aggression and initial testosterone levels (natural log transformed), post-GnRH challenge testosterone levels (natural log transformed), and the magnitude of GnRH-induced increase (In post-challenge – In initial). Aggression was measured only once for each individual. Testosterone levels were measured either once or twice. Testosterone values were adjusted for multiple measurements as well as handling time, mass, and day of year as described in the Methods.
Table 4.2. Linear mixed models of relationships between testosterone and territorial aggression (first principal component). Relationships with $P < 0.05$ shown in bold, $0.05 < P < 0.1$ shown in italics.

1. ln Initial testosterone

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>$F$</th>
<th>$df$</th>
<th>$b$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln Handling time</td>
<td>10.22</td>
<td>1, 37.6</td>
<td>-0.45</td>
<td>0.003</td>
</tr>
<tr>
<td>Day</td>
<td>0.95</td>
<td>1, 51.9</td>
<td>-0.01</td>
<td>0.33</td>
</tr>
<tr>
<td>Mass</td>
<td>1.39</td>
<td>1, 23.9</td>
<td>0.03</td>
<td>0.25</td>
</tr>
<tr>
<td>PC1 Aggression</td>
<td>3.30</td>
<td>1, 24.3</td>
<td>0.14</td>
<td>0.082</td>
</tr>
</tbody>
</table>

2. ln Post-challenge testosterone

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>$F$</th>
<th>$df$</th>
<th>$b$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln Handling time</td>
<td>4.72</td>
<td>1, 49.3</td>
<td>-0.21</td>
<td>0.035</td>
</tr>
<tr>
<td>Day</td>
<td>11.50</td>
<td>1, 42.2</td>
<td>-0.03</td>
<td>0.002</td>
</tr>
<tr>
<td>Mass</td>
<td>0.12</td>
<td>1, 34.0</td>
<td>-0.08</td>
<td>0.73</td>
</tr>
<tr>
<td>PC1 Aggression</td>
<td>29.31</td>
<td>1, 29.3</td>
<td>0.22</td>
<td>0.033</td>
</tr>
</tbody>
</table>

3. GnRH-induced increase (ln post-challenge – ln initial)

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>$F$</th>
<th>$df$</th>
<th>$b$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln Handling time</td>
<td>0.81</td>
<td>1, 52.0</td>
<td>0.28</td>
<td>0.37</td>
</tr>
<tr>
<td>Day</td>
<td>5.78</td>
<td>1, 35.8</td>
<td>-0.02</td>
<td>0.022</td>
</tr>
<tr>
<td>Mass</td>
<td>2.72</td>
<td>1, 42.3</td>
<td>-0.15</td>
<td>0.11</td>
</tr>
<tr>
<td>PC1 Aggression</td>
<td>0.69</td>
<td>1, 32.6</td>
<td>0.09</td>
<td>0.41</td>
</tr>
</tbody>
</table>
In males that were measured for parental behavior, mean initial testosterone was 2.95 ± 0.152 ng ml⁻¹, mean post-challenge testosterone was 6.88 ± 0.591 ng ml⁻¹, and the mean GnRH-induced increase was 3.93 ± 0.564 ng ml⁻¹ ($n = 25$). The natural log transformed values were correlated as follows: initial – post, $r = 0.34$ ($P = 0.10$); initial – increase, $r = -0.27$ ($P = 0.20$); post – increase, $r = 0.82$ ($P < 0.001$; $n = 25$). Mean feeding rate was 5.2 ± 0.71 visits hr⁻¹, which was comparable to the feeding rate previously shown by control males in the presence of a female (Ketterson et al. 1992).

Nestling feeding rate was negatively related to the magnitude of GnRH-induced testosterone increase (Figure 4.2, Table 4.3). There was a trend toward a relationship with post-challenge testosterone, and there was no significant relationship with initial testosterone (Figure 4.2, Table 4.3).

The average feeding rate was higher in 2003 (6.7 visits hr⁻¹) than in 2004 (3.7 visits hr⁻¹, $P = 0.04$), but there was no year difference in any of the testosterone measures (Table 4.3), suggesting that year differences did not generate our results. Indeed, our results did not differ if we used year-adjusted values for nestling feeding rate.
Figure 4.2. Relationships between nestling feeding rate and initial testosterone levels (natural log transformed), post-GnRH challenge testosterone levels (natural log transformed), and the magnitude of GnRH-induced increase (ln post-challenge – ln initial). Testosterone values were adjusted for year, handling time, mass, and day of year as described in the Methods.
Table 4.3. Linear mixed models of relationships between testosterone and nestling feeding rate.

Relationships with $P < 0.05$ shown in bold, $0.05 < P < 0.1$ shown in italics.

1. ln Initial testosterone

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>$F$</th>
<th>df</th>
<th>$b$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>0.11</td>
<td>1, 14.9</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>ln Handling time</td>
<td>2.26</td>
<td>1, 14.4</td>
<td>-0.12</td>
<td>0.15</td>
</tr>
<tr>
<td><strong>Day</strong></td>
<td><strong>6.34</strong></td>
<td>1, <strong>15.7</strong></td>
<td><strong>-0.01</strong></td>
<td><strong>0.023</strong></td>
</tr>
<tr>
<td>Mass</td>
<td>0.16</td>
<td>1, 18.9</td>
<td>-0.03</td>
<td>0.70</td>
</tr>
<tr>
<td>Feeding rate</td>
<td>1.57</td>
<td>1, 9.9</td>
<td>0.02</td>
<td>0.24</td>
</tr>
</tbody>
</table>

2. ln Post-challenge testosterone

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>$F$</th>
<th>df</th>
<th>$b$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>0.41</td>
<td>1, 16.9</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>ln Handling time</td>
<td>0.02</td>
<td>1, 14.7</td>
<td>-0.04</td>
<td>0.89</td>
</tr>
<tr>
<td><strong>Day</strong></td>
<td><strong>0.30</strong></td>
<td>1, <strong>14.5</strong></td>
<td><strong>-0.002</strong></td>
<td><strong>0.59</strong></td>
</tr>
<tr>
<td>Mass</td>
<td>3.23</td>
<td>1, 11.2</td>
<td>-0.11</td>
<td>0.099</td>
</tr>
<tr>
<td><strong>Feeding rate</strong></td>
<td><strong>3.43</strong></td>
<td>1, <strong>19.0</strong></td>
<td><strong>-0.05</strong></td>
<td><strong>0.079</strong></td>
</tr>
</tbody>
</table>

3. GnRH-induced increase (ln post-challenge – ln initial)

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>$F$</th>
<th>df</th>
<th>$b$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>0.87</td>
<td>1, 18.2</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>ln Handling time</td>
<td><strong>3.59</strong></td>
<td>1, <strong>16.7</strong></td>
<td><strong>0.17</strong></td>
<td><strong>0.076</strong></td>
</tr>
<tr>
<td>Day</td>
<td>0.001</td>
<td>1, 16.5</td>
<td>-0.001</td>
<td>0.98</td>
</tr>
<tr>
<td>Mass</td>
<td><strong>4.16</strong></td>
<td>1, <strong>13.0</strong></td>
<td><strong>-0.12</strong></td>
<td><strong>0.062</strong></td>
</tr>
<tr>
<td><strong>Feeding rate</strong></td>
<td><strong>4.51</strong></td>
<td>1, <strong>18.7</strong></td>
<td><strong>-0.05</strong></td>
<td><strong>0.047</strong></td>
</tr>
</tbody>
</table>
Discussion

To address natural variation in the hormonal resolution of the trade-off between mating effort and parental effort, we assessed individual variation in the responsiveness of the HPG axis and its covariation with aggressive and parental behavior. Implantation studies of free-living dark-eyed juncos have shown that experimentally enhanced testosterone can alter the resolution of this trade-off (Ketterson and Nolan 1992, 1994, 1999). Our results based on testosterone response to a GnRH challenge indicate that this conclusion can be generalized to natural differences among individuals upon which selection may act. In general, the behavior of males that produced higher testosterone levels suggested higher mating effort and lower parental effort. Both territorial aggression and parental behavior were predicted by aspects of the hormonal response to stimulation of the HPG axis. Specifically, males that produced higher maximum testosterone levels in response to GnRH were more aggressive when responding to a simulated territorial intrusion, and males that showed a greater increase above initial testosterone levels after a GnRH challenge fed their offspring less often. Initial testosterone showed a trend toward a positive relationship with aggression but was not significantly related to parental behavior. These results suggest that individuals vary along a testosterone-mediated continuum between individuals that invest heavily in the survival of their offspring and those that avoid parental care to seek additional mating opportunities.

Although these relationships were measured in different individuals in different years, the common physiological link suggests that short-term testosterone elevations may underlie individual variation in the resolution of the mating effort parental effort
trade-off in this species. To the extent that this variation is heritable, the mating effort/parental effort trade-off should be able to respond to selection. Below, we discuss the implications of our findings for understanding both the mechanistic basis and the evolution of hormonally mediated trade-offs.

**Short-term testosterone elevation and behavior**

The “challenge hypothesis” states that testosterone should be most closely associated with behavior during periods of social instability (Wingfield et al. 1987, 1990). Such a relationship is expected to arise because testosterone levels increase in response to social stimuli, likely via stimulation of the HPG axis (Harding 1981; Wingfield 1985). One common pattern is that testosterone levels and aggression show a concomitant increase in response to simulated territorial intrusions (Wingfield 1985).

In this study, we found that the testosterone levels produced in response to a GnRH challenge (which predicts testosterone levels produced in response to a male social stimulus, Chapter 3) were positively correlated with the aggressive response to a territorial intruder. Because GnRH challenges and measurement of territorial aggression were separated by as much as a month, this result suggests that the relationship represents a property of an individual; in other words, males likely vary consistently in both ability to produce testosterone and aggressiveness. This suggestion is supported by the repeatability of GnRH challenge response across the breeding season previously demonstrated in male juncos (Jawor et al. 2006). Although we did not measure
aggressive response repeatedly in this study, response to a territorial intrusion has been shown to be repeatable in a closely related species (song sparrows, *Melospiza melodia*, Nowicki et al. [2002]).

We also found a trend toward a relationship between initial testosterone levels before GnRH challenge and aggressive response. Initial and post-challenge levels showed a strong positive correlation during early breeding, suggesting that the correlations between aggression and these testosterone measures may be indicating the same relationship. Alternatively, some males may have been captured after engaging in a territorial dispute or courtship, and thus had elevated initial levels of testosterone. An examination of the data shows that this may be the case. The positive trend appears to be driven by the two males with the highest initial levels, which also displayed weak increases in response to the GnRH challenge, suggesting that their HPG axis may have already been maximally stimulated. A third possibility is that baseline testosterone levels and short-term elevations may both be important for producing aggressive behavior. Although aggressive response is clearly related to individual variation in the HPG system, further study is necessary to disentangle the influences of baseline testosterone levels and short-term increases.

Male parental behavior was negatively related to the testosterone response to a GnRH challenge, and showed a non-significant trend toward a negative relationship with absolute post-challenge levels. There was no relationship between parental behavior and initial testosterone levels. Despite maintaining low initial plasma levels during nestling feeding, males retain the physiological ability to produce short-term testosterone elevations that in some cases approach breeding-season peak levels (Ketterson and Nolan
Our results suggest that variation in the magnitude of these elevations, rather than initial testosterone levels, may underlie natural variation in parental behavior. Short-term testosterone elevation may act as a mechanism allowing a male to alternate between feeding nestlings and other behaviors such as territory defense and mate search.

If a causal relationship between short-term testosterone elevations and behavior exists, it may arise by two mechanisms that are not mutually exclusive. First, previous socially induced testosterone elevations may have had a “priming” effect on the behavior. It is well established that testosterone elevations contribute to the persistence of territorial aggression, particularly in winners of contests (e.g. Wingfield 1994; Trainor et al. 2004). Males may respond more aggressively to an intruder because high levels of testosterone produced in response to previous interactions with competitors have had persistent effects on brain regions related to aggression. Such an effect is likely to act through regulation of gene expression, which is the classical mechanism of steroid action (Nelson 2005). The extent to which individuals vary in the physiology related to this effect is unknown. Males that produced higher levels of testosterone in response to GnRH challenges are likely to have produced larger natural testosterone increases during past encounters with conspecifics, and may thus have up-regulated expression of certain genes necessary for producing territorial aggression.

Second, testosterone may have rapid activational effects on behavior, likely occurring by way of mechanisms that do not involve gene expression. In an elegant experiment using toadfish, Remage-Healey and Bass (2006) demonstrated a rapid increase in calling behavior when males were fed 11–ketotesosterone, the primary
androgen in fishes. Neurophysiological studies of a closely related species suggest that these behavioral changes occur due to hormonal effects on the activity of the vocal control region of the brain (Remage-Healey and Bass 2004).

Such rapid effects of testosterone on behavior may be mediated by conversion to estrogens by the enzyme aromatase at the target location. Testosterone often exerts its effects on the brain via this mechanism (Nelson 2005). There is strong evidence that estrogens may have rapid neuromodulatory effects, which may in turn cause behavioral shifts (Maggi et al. 2004; Cornil et al. 2006). For example, in Japanese quail (Coturnix japonica), rapid changes in sexual behavior have been linked to changes in the bioavailability of estrogen (Balthazart et al. 2006). Cornil et al. (2006) suggest that aromatization may commonly underlie rapid behavioral effects of testosterone, such as responses to territorial intruders or the production of sexual behavior. This may occur via rapid changes in circulating plasma testosterone, as seen in many songbirds, or by rapid modulation of brain aromatase activity, as demonstrated in quail (Cornil et al. 2006).

Although it is clear how short-term testosterone increases might directly mediate the expression of territorial behavior, it is less obvious how such changes might affect parental behavior. One possibility is that males respond to stimuli, such as a neighbor’s song or the presence of a female, by temporarily shifting their activity from parental behavior to song or courtship. The magnitude of testosterone increases produced in response to these stimuli may affect the likelihood or duration of such a shift. In support of this hypothesis, testosterone-implanted males show increased song rate coincident with decreased parental care (Ketterson et al. 1992). The temporal pattern of nestling feeding we observed is consistent with intermittent shifts in behavior. Visits to the nest were not
evenly distributed in time (see also Clotfelter et al. 2007). Males with the lowest feeding rates (less than 4 visits hr\(^{-1}\)) often left nestlings unattended for over an hour at a time (J. W. McGlothlin, personal observation). These long gaps suggest that males that feed infrequently are not poor foragers but rather are allocating effort to other activities. Although social stimulation may cause testosterone elevations, leading to behavioral shifts, an alternative explanation is that males produce such elevations spontaneously. Further study is necessary to examine whether male or female stimuli have the capacity to alter male parental behavior.

_Evolution of testosterone-mediated trade-offs_

A long-term implantation study conducted in this population found that on average, males treated with testosterone had higher fitness than controls (Reed et al. 2006). This effect occurred because testosterone-treated males had higher mating success (as measured by extra-pair fertilizations), which more than compensated for their decreased survival (Raouf et al. 1997; Reed et al. 2006). Such results suggest that selection should favor males with constitutively elevated testosterone.

However, our data suggest an alternative option. Levels of testosterone produced in response to GnRH challenge are similar to those produced both at the early breeding season peak (late March – early April) and by treatment with testosterone implants (Ketterson et al. 1992; Jawor et al. 2006). Further, response to GnRH covaries with the magnitude of natural testosterone increases produced in response to social stimuli.
(Chapter 3). This shows that males can produce short-term increases in testosterone as needed to support territorial (and perhaps sexual) behavior, without maintaining circulating testosterone at a constitutively high level. Males may thus avoid or moderate some of the costs of high testosterone without losing the ability to produce testosterone-mediated behavior (Wingfield et al. 2001). In the long term, males with flexible HPG axes would likely be favored over males with inflexible, but high, testosterone levels.

The potential costs of producing short-term testosterone increases have not been explored. To the extent that the costs of testosterone are related to the production of testosterone-mediated behavior (rather than systemic effects such as immunosuppression), short-term elevation may indeed be costly. One of the common results of implantation studies is that experimentally enhanced testosterone leads to increased activity (e.g. Lynn et al. 2000). This effect is likely to be associated with short-term testosterone elevation as well, as evidenced by the association with territorial aggression shown here. Increased activity may be beneficial in terms of mating success, leading to more vigorous territorial defense or the ability to encounter more females (Chandler et al. 1994; Raouf et al. 1997), while at the same time imposing survival costs such as depletion of energy stores or increasing visibility to predators (Ketterson et al. 1991; Reed et al. 2006). Future studies should examine the relationship of natural levels of testosterone with trade-offs involving survival in more detail.

Although selection pressures on baseline testosterone levels may indeed differ, it is not clear whether these aspects of the HPG system are likely to evolve separately. In this study, initial and post-challenge testosterone levels were positively correlated, which reflects that they are two manifestations of a common hormonal system. The maintenance
of circulating levels of testosterone as well as the production of short-term increases depends on the stimulation of the HPG axis, suggesting that common genes are likely to be associated with variation in both. The evolutionary independence of different aspects of the HPG system may depend on whether behavioral effects depend more upon absolute or relative levels of testosterone. The results presented here, although not conclusive, suggest that both may be important.

Although our results indicate that natural variation in testosterone levels is associated with behavioral variation, we do not wish to diminish the importance of other aspects of the hormonal system. Our GnRH challenge protocol was designed to assess variation in pituitary and gonadal response, but important variation is likely to exist both upstream and downstream of the HPG axis. In order for GnRH to be released, individuals must integrate environmental and social stimuli in the neural pathways that stimulate the hypothalamus. Individuals are likely to vary in sensory as well as neural mechanisms. Downstream, testosterone is often converted to another hormone, and regardless of whether conversion occurs, hormones must interact with receptors to have an effect. Variation likely exists in enzyme activity, receptor expression, and the pathways activated by the hormone-receptor complex. Important evolutionary changes may occur at any of the steps along this complex pathway (Hau 2007).
Conclusion

We have demonstrated that individual variation in parental and territorial behavior is related to individual variation in testosterone production. Furthermore, our results suggest that the ability to produce short-term testosterone increases may be more important for the mediation of this trade-off than circulating baseline levels, at least in our species. This is not likely to be true in all species, however. For example, some species do not decrease parental care when testosterone is elevated (Hunt et al. 1999; Van Duyse et al. 2000, 2002; Lynn et al. 2002, 2005), and some do not increase testosterone in response to social stimuli (reviewed in Landys et al. 2007). Interestingly, these species differences seem to be related to changes in the mating effort/parental effort tradeoff. Behavioral insensitivity to testosterone seems to occur in species where male parental care is critical to offspring survival (Lynn et al. 2005). Further, species that rear only a single brood (and thus may have less conflict between periods of mating effort and parental effort) are less likely to show socially modulated testosterone increases (Landys et al. 2007). Across species, testosterone-mediated trade-offs seem to be evolutionary labile, responding to changes in social and environmental selection pressures (Hau 2007). Our study provides initial evidence that the testosterone-mediated trade-off between mating effort and parental effort varies among individuals, and thus, the raw material necessary for selection to generate among-species patterns seems to exist within populations.
Acknowledgments

We thank Amanda Bessler, Jackie Gaudioso, Nicki Gerlach, Tim Greives, Carrington Knox, Dawn O’Neal, Sara Schrock, Eric Snajdr, Charles Ziegenfus, and Devin Zysling for assistance in the field, Butch Brodie III, Greg Demas, Britt Heidinger, Emília Martins, Trevor Price, and Dale Sengelaub for helpful discussions and advice, Henry Wilbur and Eric Nagy for assistance at Mountain Lake Biological Station, and Mountain Lake Hotel and Wilderness Conservancy for permission to work on their property. This study was funded by grants from the National Science Foundation (DEB 0508692), the American Ornithologists’ Union, Indiana University, MLBS, Sigma Xi, and the Wilson Ornithological Society to JWM, and National Science Foundation grants (IBN 9701334, IBN 0216091, IOS 0519211) to EDK. This research adhered to the Association for the Study of Animal Behaviour/Animal Behavior Society Guidelines for the Use of Animals in Research, the legal requirements of the United States of America, the Commonwealth of Virginia, and the State of Indiana, and the guidelines of the Animal Care and Use Committees at the University of Virginia and Indiana University.
Chapter Five

Stabilizing selection on natural variation in androgen responsiveness

Joel W. McGlothlin
Jodie M. Jawor*
Ellen D. Ketterson

*Department of Biological Sciences, University of Southern Mississippi, Hattiesburg, MS
Summary

Because of their roles in mediating life-history trade-offs and the integrated expression of functionally related traits, hormones may be important to understanding both adaptation and evolutionary constraint. However, few studies have examined how selection acts on hormones and hormonally mediated traits in natural populations, perhaps because hormone levels may show substantial variation within individuals due to seasonal and social effects. Using breeding males in a population of dark-eyed juncos (*Junco hyemalis*), we measured selection acting on natural variation in testosterone via differences in annual survival. We measured testosterone levels at multiple times during the breeding season to account for seasonal variation, and used a standardized injection of gonadotropin-releasing hormone (GnRH) to assess variation in males’ ability to produce short-term testosterone increases. The magnitude of short-term testosterone increases produced in response to GnRH, or androgen responsiveness, has previously been shown to vary repeatably among individuals and to be correlated with attractive plumage, aggressive behavior, and parental behavior. We found no selection acting on testosterone levels measured before GnRH injection, but there was strong stabilizing selection acting on the magnitude of response to GnRH challenges. In other words, males with either relatively high or relatively low androgen responsiveness were less likely to return the following year. This result indicates that selection may act to preserve an intermediate multivariate optimum for the suite of traits mediated by testosterone. We suggest that differences in fitness may arise due to the survival costs imposed by both mating effort and parental effort.
Introduction

Natural selection acts directly on phenotypes, causing evolutionary change when phenotypic variation is associated with genetic variation. To understand how populations evolve, it is important to understand the relationships between genotype and phenotype (the “phenotype landscape,” Rice 1998, 2002; Wolf et al. 2001) as well as between phenotype and fitness (the “adaptive landscape,” Simpson 1944; Lande 1976, 1979; Arnold et al. 2001). The study of endocrinology in an evolutionary/ecological context may provide important insight into both of these relationships. Hormones are often crucial to 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; Adkins-Regan 2005). Hormones are also expected to be intimately related to fitness and often 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).

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 (e.g. Ketterson and Nolan 1999; Adkins-Regan 2005; Reed et al. 2006). Despite their power to probe mechanistic relationships and demonstrate causality, such experiments can provide only limited information about evolutionary processes. This is because the act of hormonal manipulation inherently obscures the phenotypic variation upon which natural
Selection on testosterone

selection acts. For this reason, manipulative experiments should be combined with studies that focus on the consequences of individual hormonal variation (Chapter 1).

Measuring phenotypic variation and fitness in natural populations allows the estimation of the current strength and form of natural selection (Lande and Arnold 1983; Endler 1986; Brodie et al. 1995; Kingsolver et al. 2001). If information about inheritance is also available, measurements of selection in the wild can be used to predict the evolutionary trajectory of the population (Lande and Arnold 1983; Grant and Grant 1995). Despite the expected relationship between hormones and fitness, we have little information about how selection acts on hormonally mediated traits in the wild (Adkins-Regan 2005). Most of the selection studies reviewed by Kingsolver et al. (2001) involve morphology; there are no measurements of selection on hormone levels or traits that are explicitly hormonally mediated. Since this meta-analysis was conducted, a small number of studies have reported relationships between hormone levels and fitness components. In a large study of cliff swallows (Petrochelidon pyrrhonota), Brown et al. (2005) found evidence for stabilizing selection on circulating levels of corticosterone (a glucocorticoid stress hormone) via annual survival, but no significant selection on testosterone levels. Bonier et al. (2007) showed a negative effect of corticosterone on female reproductive success in white-crowned sparrows (Zonotrichia leucophrys). In a study that manipulated operational sex ratio, Mills et al. (2007) found evidence of directional selection on male testosterone levels in male bank voles (Clethrionomys glareolus).

One of the reasons measuring selection on hormones and hormonally-mediated traits has proved difficult is the commonness of within-individual variation. 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). Short-term variation 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; Goymann et al. 2007; Landys et al. 2007). Socially modulated testosterone change (also known as androgen responsiveness) is likely to be particularly important when considering the evolution of testosterone-mediated traits. Comparative studies have associated species differences in androgen responsiveness to differences in mating system and life history (Wingfield et al. 1990; Goymann et al. 2007; Landys et al. 2007). Short-term testosterone changes have long been associated with territoriality (Wingfield 1985; Wingfield et al. 1987), and have recently been associated with parental behavior as well (Chapter 4).

In this study, we examined the effects of natural testosterone levels on annual survival of adult males in a population of dark-eyed juncos (*Junco hyemalis*). A long-term study in this population found that experimentally elevated testosterone levels decreased survival (Reed et al. 2006). However, these males more than compensated for reduced survival by siring more offspring via extra-pair fertilizations, and as a result, had higher lifetime fitness (Raouf et al. 1997; Reed et al. 2006). Despite this negative effect of enhanced testosterone, we do not necessarily expect to find that survival selection favors lower testosterone levels in nature. Males may differ in their optimal testosterone levels due to differences in quality. In this case, survival selection may favor intermediate or even higher testosterone levels. In addition, selection may act differently on constitutively expressed (baseline) testosterone levels and androgen responsiveness.
In order to test these hypotheses, we measured survival selection on circulating testosterone and short-term testosterone increases using the regression method of Lande and Arnold (1983). Testosterone levels were measured at multiple points in the breeding season, and short-term increases were measured using a standardized injection of gonadotropin-releasing hormone (a “GnRH challenge”). In nature, 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 maximum androgen responsiveness, and stimulates testosterone increases that are repeatable (Jawor et al. 2006) and that predict those produced in response to social stimuli in the wild (Chapter 3). In addition, these increases are correlated with attractive plumage as well as variation in behavior (Chapters 3 & 4). These considerations suggest that measuring selection on short-term testosterone increases should provide insight into the evolution of the HPG axis and the traits it mediates.

Methods

Study area and species

We studied a resident population of the Carolina subspecies of the dark eyed junco (J. h. carolinensis) that breeds at and around Mountain Lake Biological Station (MLBS) in Giles County, VA (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 and incubate clutches, usually of 4 eggs, alone. Both parents care for offspring after hatching. Mating often occurs outside the pair, and experimentally elevated testosterone has been shown to increase extra-pair mating success (Ketterson et al. 1997; Raouf et al. 1997; Nolan et al. 2002; Reed et al. 2006).

**Capture**

In April-August 2003-2004, males ($n = 91$) were captured using mist nets or Potter traps. Upon capture, birds were transported to a central laboratory at MLBS in a holding bag. If previously uncaptured, birds were given a numbered aluminum leg band and a unique combination of plastic color leg bands for identification. We determined age (yearling or older adult [$\geq 2$ years]) by examining the color of the primary wing coverts, and secondarily, the iris, which are both lighter in yearlings (Nolan et al. 2002). Mass (g) was measured using a spring balance.

**GnRH challenges**

Each time a bird was captured, a blood sample was obtained from the wing vein (initial sample). Handling time was recorded as the time in min from capture to collection of this blood sample, averaging 48 min (range 2 – 217 min) (Jawor et al. 2006). A solution of
1.25 μg chicken GnRH-I (Sigma L0637) in 50μl of 0.1 M phosphate-buffered saline was then injected into the pectoral muscle. The bird was returned to its holding bag, and after 30 min, a second blood sample was collected (post-challenge sample). After this sample, the bird was released at the site of capture. Plasma was separated and frozen (-20ºC) for later hormone analysis.

To control for the idiosyncrasies of capture and to obtain robust individual estimates of testosterone production, we attempted to obtain four samples each year from individual birds, collected at four sampling stages across the breeding season (Jawor et al. 2006). We attempted to obtain two samples during early breeding (21 April – 16 May) by catching birds at random in baited mist nets and traps. The first GnRH challenge was administered upon each bird’s first capture (2003: 28 April – 16 May, n = 54; 2004: 21 April – 11 May, n = 46, combined n = 100) and the second after waiting 7 – 21 days (mean 10.4; 2003: 6 May – 16 May, n = 26; 2004: 1 May – 15 May, n = 11, combined n = 37). During early breeding, many birds were beginning to nest, but the exact stage of reproduction was unknown for most of them (dates of first egg were 26 April in 2003 and 25 April in 2004). Some birds were captured and given a GnRH challenge while feeding 6-7 day old nestlings (2003: 25 May – 29 June, n = 14, 2004: 20 May – 20 July, n = 14, combined n = 28). Captures during this stage were made by placing a mist net at the nest. A final set of birds was captured at the end of the breeding season, but prior to the onset of molt, using baited mist nets (2003: 15 July – 6 August, n = 7; 2004: 20 July – 5 August, n = 9). All sampling periods occurred after the typical early-breeding season testosterone peak (26 March – 14 April, Ketterson and Nolan, 1992; Ketterson et al. 2005). Overall, 5 individuals were challenged a total of 5 times, 6 were challenged 4
times, 12 were challenged 3 times, 28 were challenged 2 times, and 40 were challenged once. Twenty-three (23) individuals received challenges in both 2003 and 2004, 36 were challenged in 2003 only, and 32 were challenged in 2004 only.

Our GnRH-challenge method stimulates a maximal testosterone response at 30 min, and levels return to baseline within 2 hours (Jawor et al. 2006). In our population, there are significant differences among the sampling periods described above in the increase of testosterone produced, indicating a gradual seasonal decline (Jawor et al. 2006). When seasonal variation is held constant, individuals show repeatable differences in the magnitude of testosterone increases above initial levels (repeatability = 0.36) (Jawor et al. 2006).

**Testosterone assays**

We determined testosterone concentrations using an EIA kit (Assay Designs, Inc., #901-065) (described in Clotfelter et al. 2004). For the analysis of samples, approximately 2000 cpm of H$^3$-T were added to allow calculation of recoveries after 2 extractions with diethyl ether. Extracts were resuspended in 50 μl of ethanol and diluted to 350 μl with assay buffer from the kit. From each reconstituted sample, 100 μl were used to determine recoveries, and duplicate 100 μl quantities were used in the EIA. T concentrations were determined with a 4-parameter logistic curve-fitting program (Microplate Manager; BioRad Laboratories, Inc.) and corrected for incomplete recoveries.
Intra-plate coefficients of variation ranged from 1 – 19% (mean 9%), and inter-plate variation was 20%. We corrected for inter-plate variation by multiplying each measurement by the grand mean of assay standards across all plates within the dataset and dividing by the plate mean of these standards.

Selection analysis

Before selection analysis, it was necessary to summarize repeated testosterone measurements into a single measurement for each individual, as well as to correct for other variables that may have affected our measurements (Jawor et al. 2006). To accomplish this, we fit general linear models that included individual and sampling stage as categorical predictors and ln handling time and mass as continuous predictors. From these models, we estimated the least-squares mean for each individual. Mean initial testosterone (ng ml⁻¹, ln transformed) and GnRH-induced testosterone increase (ln post-challenge testosterone – ln initial testosterone) were estimated in separate models. We used GnRH-induced increase rather than absolute post-challenge testosterone to reduce collinearity of variables in the selection analysis. Means were calculated separately for each year (2003, n = 59; 2004, n = 55).

Estimates of annual survival were based on recapturing or sighting a male in the following breeding season (April – August). Census methods remained consistent from year to year (see Chapter 2, Reed et al. 2006 for details). If a male was caught or seen at any time in 2004, he was assigned a survival of 1 for 2003; otherwise, he was assigned a
survival of 0. The population was censused in 2005 to estimate 2004 survival, and again in 2006 to check these estimates. Juncos are highly philopatric between breeding seasons, suggesting that our estimates of survival were accurate (Nolan et al. 2002).

Selection gradients were estimated using linear regression (Lande and Arnold 1983; Brodie et al. 1995). Selection was measured both for each year separately and for the two years combined. Before analysis, individual survival was divided by average survival so that the dependent variable would be relative fitness. In 2003, 27 of 59 males (45.7%) survived, and in 2004, 32 of 55 males (58.2%) survived. For the combined-year analysis, we used the overall average survival (51.8%). Linear (directional) selection gradients were estimated in a model that did not include the non-linear terms to reduce collinearity. Non-linear (quadratic and correlational) selection gradients were estimated using a full model. The quadratic terms from the regression (and their standard errors) were doubled to generate the quadratic selection gradients (Brodie et al. 1995). Hormone measurements were standardized to zero mean and unit variance before analysis; thus, we report standardized selection gradients (Lande and Arnold 1983; Brodie et al. 1995).

Statistical significance of selection gradients was tested using binomial logistic regression. To visualize the form of selection, we fit univariate cubic splines using glms v. 4.0/glmsWIN v. 1.0 (Schluter 1988). If necessary, we used “R” to fit thin-plate splines to visualize multivariate effects (http://www.r-project.org). For both types of spline, the smoothing parameter ($\lambda$) was chosen by minimizing generalized cross-validation scores.
Results

When both years of our study were combined, there was significant negative quadratic selection acting on the magnitude of GnRH-induced testosterone increase (Table 5.1). Because there was no significant directional selection, this may be interpreted as stabilizing selection. Males with very high or very low GnRH challenge response were less likely to return the following breeding season (Figure 5.1). Stated another way, survivors had lower variance in GnRH-induced testosterone increases than did non-survivors (Figure 5.2). Stabilizing selection was stronger in 2003 than in 2004, when it was not statistically significant (Table 5.1).

In 2003, positive correlational selection acted on initial testosterone and GnRH-induced testosterone increases (Table 5.1). Thin-plate spline visualization indicates that this effect was due to low survival of males with high initial testosterone and low GnRH-induced testosterone increases (Figure 5.3).

We detected no significant directional selection, and there were no other significant non-linear gradients (Table 5.1).
Table 5.1. Linear (β) and non-linear (γ) selection gradients (± 1 S.E.) for measures of testosterone, estimated using linear regression. *P* values were derived from binomial logistic regressions. See Methods for details. Effects with *P* < 0.05 are shown in bold.

<table>
<thead>
<tr>
<th>Trait</th>
<th>β/γ</th>
<th>S.E.</th>
<th><em>P</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Both years combined</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln Initial testosterone</td>
<td>-0.04</td>
<td>0.09</td>
<td>0.65</td>
</tr>
<tr>
<td>GnRH-induced testosterone increase</td>
<td>0.11</td>
<td>0.09</td>
<td>0.24</td>
</tr>
<tr>
<td>Initial&lt;sup&gt;2&lt;/sup&gt;</td>
<td>-0.02</td>
<td>0.12</td>
<td>0.93</td>
</tr>
<tr>
<td>Increase&lt;sup&gt;2&lt;/sup&gt;</td>
<td>-0.37</td>
<td><strong>0.14</strong></td>
<td><strong>0.010</strong></td>
</tr>
<tr>
<td>Initial × increase</td>
<td>0.06</td>
<td>0.11</td>
<td>0.48</td>
</tr>
<tr>
<td><em>n</em> = 114</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Trait</th>
<th>β/γ</th>
<th>S.E.</th>
<th><em>P</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln Initial testosterone</td>
<td>-0.27</td>
<td>0.19</td>
<td>0.19</td>
</tr>
<tr>
<td>GnRH-induced testosterone increase</td>
<td>0.06</td>
<td>0.14</td>
<td>0.96</td>
</tr>
<tr>
<td>Initial&lt;sup&gt;2&lt;/sup&gt;</td>
<td>0.39</td>
<td>0.25</td>
<td>0.60</td>
</tr>
<tr>
<td>Increase&lt;sup&gt;2&lt;/sup&gt;</td>
<td><strong>-0.56</strong></td>
<td><strong>0.21</strong></td>
<td><strong>0.014</strong></td>
</tr>
<tr>
<td>Initial × increase</td>
<td><strong>0.55</strong></td>
<td><strong>0.27</strong></td>
<td><strong>0.048</strong></td>
</tr>
<tr>
<td><em>n</em> = 59</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Trait</th>
<th>β/γ</th>
<th>S.E.</th>
<th><em>P</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln Initial testosterone</td>
<td>0.11</td>
<td>0.15</td>
<td>0.59</td>
</tr>
<tr>
<td>GnRH-induced testosterone increase</td>
<td>0.21</td>
<td>0.11</td>
<td>0.15</td>
</tr>
<tr>
<td>Initial&lt;sup&gt;2&lt;/sup&gt;</td>
<td>-0.08</td>
<td>0.17</td>
<td>0.71</td>
</tr>
<tr>
<td>Increase&lt;sup&gt;2&lt;/sup&gt;</td>
<td>-0.29</td>
<td>0.23</td>
<td>0.20</td>
</tr>
<tr>
<td>Initial × increase</td>
<td>0.02</td>
<td>0.15</td>
<td>0.86</td>
</tr>
<tr>
<td><em>n</em> = 55</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Selection on testosterone

Figure 5.1. Cubic spline visualization of survival selection acting on (a) initial (pre-GnRH-challenge) testosterone levels and (b) GnRH-induced testosterone increases in 2003 and 2004. Rectangles represent individual data points. Dotted lines represent ± 1 S.E. based on 50 bootstrap replicates.
Figure 5.2. Frequency distributions of GnRH-induced testosterone increases of individuals that survived (top) or died (bottom).

Figure 5.3. Thin-plate spline visualization of selection acting on initial testosterone levels and GnRH-induced testosterone increases in 2003. Relative fitness ($w$) is plotted on the vertical axis.
Our results indicate that stabilizing selection acts on the magnitude of short-term testosterone increases produced in response to injections of GnRH via differences in annual survival. Stabilizing selection was relatively strong; the magnitude of selection was well above the median quadratic selection ($|γ| = 0.10$) reported by Kingsolver et al. (2001). Although the strength varied, selection was consistently stabilizing across years. Because testosterone increases produced in response to GnRH injections predict those produced in response to territorial intruders (Chapter 3), our results suggest that natural selection acts to maintain socially modulated androgen responsiveness at its survival optimum. We found no significant evidence of selection on initial testosterone levels, providing evidence that breeding baseline testosterone is a less important determinant of survival than androgen responsiveness. Although our measurements of initial testosterone are not necessarily identical to breeding baseline because of the handling time involved in obtaining the blood sample, our results are consistent with those of Brown et al. (2005), who found no evidence of selection acting on baseline testosterone.

Although we measured selection by relating fitness directly to hormonal measurements, we do not mean to suggest that selection acts directly on testosterone levels. Just as genetic evolution arises from selection on expressed phenotypic variation, evolution of hormone levels is likely to occur when selection acts on hormonally mediated traits (Chapter 1). Hormone levels are “invisible” to selection unless they have effects on an organism’s phenotype. Hormones often mediate the expression of integrated suites of traits (Finch and Rose 1995; Ketterson and Nolan 1999). Because hormones
Selection on testosterone

often mediate suites of functionally related traits, traits mediated by hormones are expected to experience correlational selection (Chapter 1). Correlational selection occurs when traits interact in their effects on fitness, and is predicted to maintain trait integration over time (Chapter 2, Lande 1980a, 1984; Cheverud 1982; Lande and Arnold 1983; Phillips and Arnold 1989; Brodie 1992; Sinervo and Svensson 2002). Quadratic (stabilizing and disruptive) selection is the multivariate analog of correlational selection, and these selection modes may be transformed from one to the other by rotation of the γ matrix (Cheverud 1982; Schluter and Nychka 1994; Blows and Brooks 2003; Blows 2007). Thus, apparent stabilizing selection on hormone levels may arise from multivariate stabilizing selection acting on the traits mediated by that hormone.

In juncos, experiments that manipulate testosterone levels have shown that the hormone has multiple phenotypic effects. For example, males with increased testosterone show increased mating effort, including more frequent song and courtship display and larger home range size, which likely contributes to their success at obtaining extra-pair fertilizations (Ketterson et al. 1992; Chandler et al. 1994; Enstrom et al. 1997; Raouf et al. 1997; Reed et al. 2006). However, males with increased testosterone show reduced parental care, as measured by nestling feeding rate and nest defense, and self-maintenance, as measured by body mass and immune function (Ketterson et al. 1991, 1992; Cawthorn et al. 1998; Schoech et al. 1998; Casto et al. 2001). Our initial results indicate that these phenotypic effects of testosterone implants translate to individual variation in testosterone production (Chapter 4). In particular, GnRH-induced increases predicted reduced parental behavior and absolute testosterone levels following GnRH challenges predicted increased territorial aggression, suggesting that androgen
responsiveness underlies the resolution of the trade-off between mating effort and parental effort (Chapter 4). Androgen responsiveness has not yet been linked to immunocompetence. Circulating testosterone has been shown to negatively predict innate immunity (Greives et al. 2006), but this relationship did not hold for GnRH-challenge response (Greives et al. unpublished data). We have not examined the relationship between natural testosterone levels and acquired immunity (which was shown to be suppressed by experimentally enhanced testosterone, Casto et al. 2001).

The curvilinear form of selection on androgen responsiveness is likely to have arisen from survival costs associated with both mating effort and parental effort. Mating effort is energetically expensive, as it involves active territorial defense, song, and display. Investment in mating effort may increase a male’s predation risk (by increasing his conspicuousness) or sensitivity to harsh environmental conditions (by depleting energy stores). Thus, males with high androgen responsiveness may have reduced survival because of their relatively high investment in mating. Such males may be similar to the testosterone-treated males in Reed et al. (2006). Note, however, that the average survival of testosterone-treated males in that study, 38%, is somewhat lower than that predicted for males with very high androgen responsiveness in this study. This suggests that the ability to modulate testosterone levels when needed may ameliorate some of the costs of testosterone, and may be preferred to maintaining constitutively high testosterone levels (Wingfield et al. 2001).

Parental effort is also energetically expensive, and may imposes survival costs of its own (Clutton-Brock 1991). The reduced survival of males with low androgen responsiveness may derive from an increased investment in parental care relative to
males with high and intermediate androgen responsiveness (Chapter 4). Another, not mutually exclusive, possibility is that reduced survival and lower androgen responsiveness may arise from a common source. For example, these males may be unable to produce high testosterone increases in response to GnRH injections because they are in poor condition, and as a result, they are also unable to survive the winter. Survival was not significantly correlated with body mass or age when considering either the entire dataset ($P > 0.5$) or only males with below-average androgen responsiveness ($P > 0.6$). However, the reduced survival of these males may have derived from an aspect of condition or quality that we did not measure.

To our knowledge, our results provide the first evidence for selection on androgen responsiveness in a natural population. We have suggested here that the curvilinear nature of the survival function arises from trade-offs between mating effort and other components of fitness. A complete understanding of how selection shapes natural variation in testosterone would require measuring selection acting at multiple fitness components other than survival, and ideally, measuring selection via total lifetime fitness as has been possible for experimentally manipulated males (Reed et al. 2006). Perhaps the most important component of fitness to examine is mating success. To that end, we are currently assessing mating success of the males in this study using DNA paternity analysis.
Acknowledgments

We thank Joe Casto, Jackie Gaudioso, Annie Lindsay, Tim Greives, Dawn O’Neal, Sara Schrock, Eric Snajdr, Peter Stevens, Charles Ziegenfus, and Devin Zysling for assistance in the field, Butch Brodie III, Britt Heidinger, Mary-Louise Maas, Emília Martins, and Dale Sengelaub for discussions and comments on the manuscript, and Mountain Lake Biological Station (Henry Wilbur, Director and Eric Nagy, Associate Director) and Mountain Lake Hotel for facilities and permission to work on their land. This research was supported by a National Science Foundation Doctoral Dissertation Improvement Grant (0508693) and grants from the American Ornithologists’ Union, Indiana University, Mountain Lake Biological Station, Sigma Xi and the Wilson Ornithological Society to JWM, and National Science Foundation grants (0216091, 0519211) to EDK. This research adhered to the Association for the Study of Animal Behaviour /Animal Behavior Society Guidelines for the Use of Animals in Research, the legal requirements of the United States of America, the State of Indiana, and the Commonwealth of Virginia, and the guidelines of the Animal Care and Use Committees at Indiana University and the University of Virginia.
Conclusions and future directions

Research summary

In this dissertation, I have examined the integration of the dark-eyed junco mating phenotype from both proximate and ultimate perspectives. In Chapter 1, I proposed that correlational selection may be an important factor affecting the evolution of integrated, hormonally mediated suites. In Chapter 2, I demonstrated how correlational selection may act to generate integration in a natural population of dark-eyed juncos. Correlational sexual selection favored an increase in the correlation between tail white (a plumage ornament) and body size (as measured by wing length). There was also a genetic correlation between these traits, suggesting that the action of correlational selection over many generations favored genetic integration of the two traits.

Chapters 3-5 focused on testosterone as a physiological integrator of the mating phenotype. This work was based on findings from the long-term field of juncos by Ketterson, Nolan, and colleagues (Ketterson and Nolan 1992, 1999; Reed et al. 2006) that revealed that experimentally enhanced testosterone increased investment in mating effort at the expense of parental effort and survival. Because males are predicted to differ in the optimal resolution of such trade-offs based on their quality, I expected to find an association between natural variation in testosterone and other components of the mating phenotype. Surprisingly, I found that a male’s ability to produce short-term testosterone increases (his androgen responsiveness) was more important than his constitutively
Conclusions

Circulating “baseline” levels. In Chapter 3, I found that males with more tail white were able to produce higher testosterone increases in response to a challenge with gonadotropin-releasing hormone (GnRH). These increases predicted a male’s hormonal response to a territorial intruder. In Chapter 4, I found evidence suggesting that androgen responsiveness was correlated with variation in the trade-off between mating effort and parental effort. Males that produced higher testosterone levels in response to GnRH were more aggressive, and those that produced higher GnRH-induced increases were less parental. Taken together, the results from Chapters 3 and 4 suggest that the integration of plumage and androgen responsiveness may act to enforce the signal honesty of tail white. I proposed that correlational sexual selection acting on tail white and androgen responsiveness (or the traits mediated by testosterone) may be responsible for generating this integration, because attractiveness and the behaviors controlled by testosterone are likely to have interactive effects on mating success.

Although I have been unable to measure sexual selection as of yet, Chapter 5 shows that selection may act on androgen responsiveness. I found that males with very high or very low androgen responsiveness were less likely to survive. Tying this to the results of Chapter 4, I suggested that the energetic costs of investing in mating effort or parental effort may lead to the reduced survival of males at each end of the distribution of androgen responsiveness. This study presents one of the few examples of selection acting on hormone levels in a natural population.

Taken together, and combined with previous results from studies of this population, the results presented here implicate correlational selection and testosterone as evolutionary and physiological integrators of the mating phenotype in dark-eyed juncos.
The results of Chapter 5 show how ultimate and proximate factors may interact and suggest potentially rewarding avenues of future research.

Evolution of phenotypic integration

In Chapter 1, I argued that correlational selection should be able to assemble and disassemble suites of hormonally mediated traits, and in Chapter 2, I presented evidence that correlational selection helped maintain a genetic correlation between plumage and body size. There are still few measurements of correlational selection in nature, despite Kingsolver et al.’s (2001) call for more such measurements. (One could argue that six years is not enough time to design, carry out, and publish a proper field study of complex non-linear selection, and that more measurements should be on the way soon.) Although more measurements of correlational selection gradients are important, alone they can tell us only about the predominance of this form of selection, rather than its effects. Brodie and McGlothlin (2007) outline some of the important issues to address in the future in order to understand the multivariate evolution of the phenotype; I will briefly summarize these views here.

Selection gradients are more useful if they are accompanied by estimates of inheritance. Along with the study presented in Chapter 2, there are only a limited number of studies that measure both correlational selection and the $G$ matrix associated with the traits under selection (Brodie 1989, 1992; Conner and Via 1993; Morgan and Conner 2001; Blows et al. 2004). With so few studies available, we have no idea whether
Conclusions

Concordance between $G$ and $\gamma$ is a general phenomenon (as would be predicted by theoretical considerations) or an uncommon occurrence seen only in a few special cases. Further, although quantitative genetic models (Lande 1980a, 1984; Phillips and Arnold 1989; Phillips and McGuigan 2006) and computer simulations (Jones et al. 2003) predict that $G$ should evolve in response to correlational selection, there are no direct experimental demonstrations of this occurrence. In contrast, the predictions of the multivariate breeder’s equation for the simultaneous evolution of trait means have been tested in a fairly large number of artificial selection experiments (reviewed by Roff 2007). Similar experiments should apply artificial correlational selection and examine the effects on $G$. Such experiments are more difficult than those applying artificial directional selection because they require measuring $G$ at both the beginning and the end of the experiment, but could be accomplished in a number of insect or plant species.

Much recent work has focused on comparing $G$ across populations and species, with the preliminary conclusion that $G$ seems to differ among species more so than among populations of the same species (Steppan et al. 2002). Future studies should focus on closely related species where multivariate phenotypic evolution has obviously occurred, preferably multiple times in multiple directions. Such adaptive radiations may act as natural experiments, providing a testing ground for many evolutionary hypotheses (Schluter 2000). Adaptive radiations tend to occur relatively rapidly when there are multiple open niches, as may occur following colonization of recently formed islands. Quantitative genetic theory that attempts to explain adaptive radiation relies on a static $G$ matrix (Schluter 1996), although there is no reason to expect this to be the case. Rather, correlational selection is highly likely to occur, because colonization of a new niche is
likely to involve simultaneous selection for multiple functionally related traits. My postdoctoral work, in collaboration with E. D. Brodie III, will focus on how G changes across multiple origins of ecomorphs in an adaptive radiation of Caribbean Anolis lizards (Losos et al. 1998).

Beyond studies of the G matrix, there is a great need for studies of the physiological and developmental bases of phenotypic integration and complex adaptation. As I argued in Chapter 1 and throughout this dissertation, understanding how hormonal mechanisms evolve may provide insight into such problems. Studies that explore how the hormonal basis of integrated suites of traits may evolve using experimental methods such as artificial selection are rare (See Zijlstra et al. 2004; Davidowitz et al. 2005; Zera 2005 for recent examples). Studying genetic differences in physiological mechanisms by examining targeted genes may provide insight into physiological evolution in species that are not amenable to artificial selection experiments (e.g. Geffeney et al. 2002, 2005). Quantifying sequence variation in genes related to growth hormone pathways, such as insulin-like growth factor, may uncover the physiological basis of adaptive changes involving size and shape differences in vertebrates (e.g. Sutter et al. 2007).

**Mechanistic links between testosterone and plumage**

Although I argued in Chapter 3 that the link between testosterone and tail white is not likely to be a mechanistic one, I have not been able to rule out this possibility.
Testosterone is not generally responsible for plumage dimorphism in songbirds, and few studies have shown a mechanistic link between testosterone and continuous variation in male plumage (Owens and Short 1995; Kimball 2006). In house sparrows, testosterone implants have been shown to increase the size of the bib, a melanin-based plumage patch used as a badge of status (Evans et al. 2000), while testosterone implants decreased the expression of bright, carotenoid-based plumage in house finches by delaying molt (Stoehr and Hill 2001). Testosterone implants that maintain the breeding-season peak completely suppress molt in juncos (Nolan et al. 1992), but it was previously unknown whether low circulating levels of testosterone would also suppress molt, or simply delay it as in house finches. In collaboration with Jodie Jawor, I attempted to test for a mechanistic link between testosterone and tail white using very small testosterone implants (1 mm of crystalline testosterone). We found that these implants either completely suppressed molt or caused a very abnormal molt. As I suggested in Chapter 3, studies that elevate testosterone on a short-term basis during molt, using hormonal injections, might be used to test for a mechanistic link.

I also attempted to test for a maternal effect of testosterone on tail white. The genetic analysis in Chapter 2 showed that 13.8% of the variation in tail white was due to a maternal variance, although the source of this variation could not be determined. Yolk androgens have been shown to affect variation in many offspring traits, including bib size in house sparrows (Strasser and Schwabl 2004). To test for an effect of maternally derived testosterone on tail white, I hand-reared nestlings, beginning at day 6 after hatching, from nests of control and testosterone-implanted females. Testosterone-implanted females had previously been shown to deposit more androgens in their yolk.
Conclusions

In 2005, I hand-reared 59 nestlings from 22 broods to independence. However, because testosterone implants seemed to suppress nesting in females (perhaps because females were implanted too early in the breeding season), only 2 of these broods were from testosterone-treated females. Due to feather breakage, I was able to measure tail white from only one testosterone-treated offspring. Although I do not recommend repeating this study in the future, another possible way to test this hypothesis is injecting yolks directly with testosterone. This method was not employed originally due to the small size of junco eggs. Recent work suggests a third possibility, though it lacks direct experimental control. Yolk testosterone has been shown to affect digit ratio in some avian species (Burley and Foster 2004; Romano et al. 2005). Digit ratio could be compared to tail white in a large number of males caught in the wild to determine whether a yolk androgen effect is plausible. Kristal Cain, in collaboration with Ellen Ketterson, has begun a study of digit ratios in a natural population of juncos.

Natural variation in testosterone and behavior

In Chapter 4, I demonstrated relationships between natural variation in testosterone, specifically short-term changes or androgen responsiveness, and behavior. However, this study did not explore all behaviors relevant to the mating phenotype. Perhaps the most important behavior to focus on next is courtship of females. Considering courtship together with male-male aggression is likely to be provide a more thorough estimation of mating effort, because both types of behavior are necessary to obtain mates. Evidence
suggests that male songbirds may produce short-term testosterone increases in response to females as well as males (Moore 1983; Pinxten et al. 2003). I attempted to measure such an increase in a pilot study of captive juncos in Spring 2007. However, most males did not show courtship behavior in response to the female. One potential reason was that my experimental design did not allow males time to establish ownership of his “territory” (a room in an indoor aviary). Rather, males were placed into a new room minutes before the female was introduced (compare Pinxten et al. 2003). Nevertheless, 3 of 22 males showed evidence of an increase in testosterone levels. All males had been given GnRH challenges a few days prior to behavioral testing. These 3 males showed post-female levels that were closer to GnRH-induced levels while the rest were similar to or below baseline levels. These data suggest that a properly designed study of the relationships between androgen responsiveness and courtship may be worthwhile.

To this end, Elizabeth Schultz is measuring courtship, aggression, and androgen responsiveness in the field this summer as her Research Experience for Undergraduates project at Mountain Lake Biological Station. In spring 2006, I developed a method, based on simulated territorial intrusions, to elicit courtship behavior in territorial male juncos. A caged female was placed in a male’s territory and a recording of a female’s “trill,” which is associated with precopulatory display (Nolan et al. 2002), was broadcast. In 3 different territories, this protocol elicited approach by the male, and in 2 of these, the male courted vigorously. The results of the study this summer will determine whether courtship and aggression are correlated across males, and perhaps, whether they are both related to a common hormonal mechanism.
Sexual selection on androgen responsiveness

Sexual selection is likely to be the major evolutionary force shaping the integration of the mating phenotype. Male juncos vary in mating success, creating an opportunity for sexual selection, primarily due to variation in extra-pair fertilization success (Ketterson et al. 1997; Raouf et al. 1997; Reed et al. 2006). Testosterone implants increase extra-pair mating success, but we know little about how natural variation in circulating testosterone and androgen responsiveness contribute to extra-pair success. Measuring this relationship is a major goal of ongoing work.

In collaboration with Sara Schrock and Danielle Whittaker, I am continuing to assign offspring to their genetic fathers using microsatellite genotyping. Because we did not find nests for all the males sampled in Chapter 5, our sample size for measuring sexual selection is much smaller (by approximately half) than that for measuring survival selection. We have nearly completed genotyping all individuals from 2003. Using the mating success of 28 males, there is no significant evidence of directional ($P > 0.17$) or non-linear ($P > 0.12$) selection. Because of poor amplification, it was necessary to re-extract many of the DNA samples from 2004; this work is currently underway.

Understanding how sexual selection acts on natural variation in testosterone (or more properly, the traits it mediates) is important for understanding how the trade-offs demonstrated in testosterone implant studies evolve. Another motivation for this work is to test our hypothesis from Chapter 3 that correlational sexual selection maintained the relationship between tail white and androgen responsiveness. Using the small dataset from 2003, I have been unable to detect such selection as of yet ($P > 0.6$). How dynamic
testosterone-mediated behaviors interact with plumage ornaments to determine mating
success is an important question for understanding evolutionary integration of the mating
phenotype. Discovering the answer, however, may await a much larger dataset.

**Evolution of the junco mating phenotype**

The work presented here indicates that the mating phenotype of male juncos is an
integrated suite of characters that varies among individuals. This variation, along with the
estimates of selection, suggests that the mating phenotype should be able to evolve, and
may do so as an integrated whole (although, as I argued in Chapter 1, selection may be
able to dissociate one or more of its component traits). Comparing the mating phenotype
across populations that have diverged may allow us to test this hypothesis. Juncos have
recently become established as an “island” population on the campus of University of
California, San Diego (Yeh 2004; Yeh and Price 2004). Interestingly, this population has
experienced rapid evolution of decreased tail white and body size (Rasner et al. 2004;
Yeh 2004). Preliminary evidence suggests that males on the campus are also less
aggressive (Newman et al. 2006). These changes may have occurred because the
opportunity for sexual selection is lower in the new habitat, and the traits associated with
the mating phenotype have consequently been reduced in response. Jonathan Atwell, a
graduate student with Ellen Ketterson, in collaboration with Trevor Price, is currently
studying morphological and behavioral differences between the campus population and
the ancestral, mountain population. This study promises to reveal how the junco mating phenotype may evolve in response to a novel environment.
References


Heidinger, B. J., I. C. T. Nisbet, and E. D. Ketterson. 2006. Older parents are less responsive to a stressor in a long-lived seabird: a mechanism for increased


Ketterson, E. D., V. Nolan, Jr., J. M. Casto, C. A. Buerkle, E. Clotfelter, J. L. Grindstaff, K. J. Jones et al. 2001. Testosterone, phenotype, and fitness: a research program


References


References


CURRICULUM VITAE

Joel W. McGlothlin

Indiana University
Department of Biology
1001 E 3rd St., Jordan Hall 142
Bloomington, IN 47405
jmcgloth@indiana.edu
(812) 855-1096

Education

2001-07  Ph. D. in Biology, Indiana University, Bloomington, IN
Advisor: Ellen D. Ketterson
1997-2001 B. A. in Biology (Honors), Vanderbilt University, Nashville, TN
Advisor: David E. McCauley

Research Grants

2005-07 NSF Doctoral Dissertation Improvement Grant (DEB 0508692, $11,783)
2005 Indiana University Graduate School Grant in Aid ($600)
Wilson Ornithological Society Louis Agassiz Fuertes Award ($2,500)
2004 American Ornithologists Union Research Award ($1,800)
Mountain Lake Biological Station Research Fellowship ($1,000)
2003-04 Sigma Xi Grant in Aid of Research ($750, $500)

Fellowships and Awards

2005-07 NIH Trainee, Research Training Grant: “Common Themes in Reproductive Diversity”
2005 Center for the Integrative Study of Animal Behavior (CISAB) Fellowship (Summer)
2005 Indiana Univ. College of Arts & Sciences McCormick Science Grant
2001-05 NSF Graduate Research Fellowship
2001 Founder’s Medal for First Honors, Vanderbilt University
2000 Research Experience for Undergraduates, Mountain Lake Biological Station
Publications


Manuscripts in Review


Conference Presentations

**Talks**

2006 Animal Behavior Society, Snowbird, UT (Symposium talk)
Evolution, Stony Brook, NY
NSERC/NSF/ESF Workshop: Individual variation, Vancouver, BC

2005 Animal Behavior Society, Snowbird, UT
Society for Integrative and Comparative Biology, San Diego, CA

2004 ESF/NSF/NSERC Workshop: Trade-offs and constraints, Wageningen, Netherlands

2002 Animal Behavior Society, Bloomington, IN

**Posters**

2006 Animal Behavior Society, Snowbird, UT
2005 NSF/ESF/NSERC Workshop: Maternal effects. Seattle, WA
2002 Evolution, Urbana, IL

**Invited Seminars**

2006 Animal Behavior Brown Bag, University of Chicago, Chicago, IL
2005 Mountain Lake Biological Station, Pembroke, VA

**Press Coverage**

2006 Paul Hess, “News and notes” *Birding* 38:30
**Professional Society Affiliations**

American Society of Naturalists  
Animal Behavior Society  
Society for the Study of Evolution

**Teaching Experience**

Department of Biology, Indiana University, associate instructor  
2003  BIOL L111, Introductory Biology, (discussion)  
2002  BIOL L376, Biology of Birds (lab)

Department of Biological Sciences, Vanderbilt University, teaching assistant  
2001  BIOL 119, General Zoology (lab)  
2000  BIOL 100, Introductory Biology (lab)  
2000  BSCI 110a, Introduction to Biological Sciences (lab)

**Undergraduate Mentoring**

2006-07  Elizabeth Schultz, Indiana University STARS Program  
2005  Jennifer Phillips, Mountain Lake Biological Station REU Program  
2003  Jacqueline Gaudioso, Mountain Lake Biological Station REU Program  
2002  Jessica Henry, CISAB REU Program

**Departmental Service**

2005-06  CISAB Steering Committee  
2005-06  Biology Faculty Search Committee

**Referee for Professional Publications**

American Naturalist (3), Animal Behaviour (1), Auk (1), Behavioral Ecology (1),  
Behavioral Ecology and Sociobiology (2), Behaviour (1), Evolution (3), Functional  
Ecology (1), Heredity (1), Hormones and Behavior (2), Proceedings of the Royal Society  
B (6)