Testosterone reduces promiscuity of female Blue Tits (Cyanistes caeruleus)

de Jong, Berber; Lens, Luc; van der Velde, Marco; Korsten, Peter; Groothuis, Ton; Komdeur, Jan

Published in:
Ethology

DOI:
10.1111/eth.12574

ethology international journal of behavioural biology

Testosterone Reduces Promiscuity of Female Blue Tits (*Cyanistes caeruleus*): An Experimental Study

Berber de Jong*†‡, Luc Lens‡, Marco van der Velde*, Peter Korsten§, Ton Groothuis† & Jan Komdeur*

* Behavioural and Physiological Ecology, Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, The Netherlands
† Behavioural Biology, Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, The Netherlands
‡ Terrestrial Ecology Unit, Department Biology, Ghent University, Ghent, Belgium
§ Department of Animal Behaviour, Bielefeld University, Bielefeld, Germany

Correspondence
Berber de Jong, Behavioural and Physiological Ecology, Groningen Institute for Evolutionary Life Sciences
University of Groningen, P.O. Box 11103, 9700 CC, Groningen, The Netherlands.
E-mail: berber.dejong@gmail.com

Received: January 8, 2016
Initial acceptance: April 19, 2016
Final acceptance: September 29, 2016
(L. Fusani)

doi: 10.1111/eth.12574

Keywords: extra-pair paternity, female testosterone, female attractiveness, mate choice, blue tit, *Cyanistes caeruleus*

Abstract

In many animal species, extra-pair copulations (EPCs) are common and can increase fitness in both sexes. In males, EPCs can increase total reproductive output, whereas in females benefits of EPCs can be indirect through improving the genetic quality of their offspring. Males and females of many vertebrates show an increase in levels of the hormone testosterone (T) during the mating period. In males, T plays an important role in regulating mating behaviour including increasing their EPC rate. While much is known about the role of T in male mating behaviour, the role of T in female reproduction remains unclear. To study the influence of T on extra-pair paternity rates in females in a field setting, we created three experimental groups of female blue tits (*Cyanistes caeruleus*): treated with either T, flutamide (Flu; an androgen receptor blocker) or empty implants before egg laying. Subsequently, we scored the number of extra-pair offspring (EPO) in their broods. We also assessed the attractiveness of females treated with either T or Flu to males in mate choice trials in the laboratory. The overall proportion of EPO was lower for the T-implanted group compared with the control group, whereas Flu had no effect. Given that males did not show a preference for Flu- vs. T-treated females in the mate choice trials, it appears less likely that the reduction in EPO in the T-implanted females was due to a reduction in their attractiveness. T levels may have negatively influenced EPO rate by affecting female within-pair and/or extra-pair mating behaviour. Future behavioural studies should investigate how elevated T levels reduce the number of EPO.

Introduction

In many animal species, extra-pair copulations (EPCs) resulting in extra-pair offspring (EPO) are common (Griffith et al. 2002). In more than 85% of socially monogamous passerine bird species, some offspring are not sired by the social partner, but by another male (Griffith et al. 2002; Griffith & Immler 2009). For males, EPCs can be beneficial because they have the potential to increase their reproductive success by increasing the number of offspring males sire (Møller & Ninni 1998). The role of EPCs is less clear in females, but there is evidence that EPCs can yield direct or indirect benefits (Jennions & Petrie 2000; Griffith et al. 2002), which can increase female fitness (Gerlach et al. 2012). Direct benefits may be gained through extra-pair males providing access to breeding resources (Birkhead & Møller 1992), paternal care to the offspring (Nakamura 1998) or fertility insurance (Sheldon 1994). Indirect benefits may concern the increase in the additive genetic quality ('good genes'; Griffith et al. 2002) or heterozygosity of the offspring ('compatible genes'; Griffith & Immler 2009; Jennions & Petrie 2000).

The ability of males to obtain EPCs has been shown to be influenced directly or indirectly by the steroid

hormone testosterone (T). T plays an important role in male mating behaviour and reproduction (Wingfield et al. 1990; Hau 2007). In many male vertebrate species, T levels increase at the beginning of the breeding season when females are sexually receptive and territorial aggression occurs frequently, and decrease at the end of the breeding season or during the period of paternal care (Moore 1982, 1984; Wingfield et al. 1990; Adkins-Regan 2005). Several studies have shown that natural plasma T concentrations in males are positively correlated with mating success (Alatalo et al. 1996), mate guarding and extra-pair paternity (Saino & Møller 1995; Garamszegi et al. 2005). Moreover, experimental studies in which T levels were elevated have shown that T enhances mating and courtship-related traits, such as singing rate (De Ridder et al. 2000), attractiveness (Enstrom et al. 1997), courtship display (Enstrom et al. 1997; De Ridder et al. 2000; Alonso-Alvarez 2001; Edler et al. 2011), as well as the probability of siring EPO (Raouf et al. 1997; but see Foerster & Kempenaers 2004). Overall, high T levels may increase male fitness by positively affecting (extra-pair) mating behaviour and thereby increasing reproductive success.

Females also have measurable T plasma levels (Staub & DeBeer 1997), which show the same seasonal variation as in males (with the highest peak coinciding with the mating period, Ketterson et al. 2005), although in general T levels in females are considerably lower than in males (Ketterson et al. 2005; Møller et al. 2005; Goyman & Wingfield, 2014). In contrast to the role of elevated T in males, its role in female mating behaviour is not clear. Previous studies have shown that experimentally elevated T levels in females (induced with hormone implants) can reduce attractiveness (Ketterson et al. 2005), choosiness in selecting a mate (McGlothlin et al. 2004) and induce male-like mating behaviour, such as establishing a territory (Lank et al. 1999), courtship feeding and mounting attempts directed towards other females (Nespor et al. 1996; Lahaye et al. 2012). Recently, it has been shown that experimentally elevated T in females resulted in a decrease in the number of EPO in spotless starlings (*Sturnus unicolor*; García-Vigón et al. 2008), but not in dark-eyed juncos (*Junco hyemalis*, Gerlach & Ketterson 2013).

So far only few studies have investigated experimentally whether T influences the rate of EPO in free-living females. Such studies are rare, because experimental manipulations of hormone levels are challenging to carry out in the wild. Here, we report on an experimental test of the effect of female T levels on EPO rates in blue tits (*Cyanistes caeruleus*), which are a key model for studying extra-pair mating. The blue tit was one of the first species for which EPO were reported (Kempenaers et al. 1992). Extra-pair offspring occur frequently in blue tits; for example, in our study population, on average 12.8% of the young is extra-pair and on average 46.7% of the broods contain at least one EPO (Magrath et al. 2009; see also Brommer et al. 2007; Vedder et al. 2011). Furthermore, female blue tits have been reported to actively search for EPCs (Kempenaers et al. 1992), and paternity is probably largely under female control (Kempenaers 1995). Therefore, if T plays a role in regulating mating behaviour in female blue tits, it is likely to also affect female extra-pair mating behaviour and consequently the rate of extra-pair paternity.

To investigate the effect of female T on extra-pair paternity rates, we treated females with T-releasing implants (Data S1). We attempted to mimic the naturally observed brief peaks in female T during the mating period by removing the implants again just after the onset of egg laying. Additionally, in contrast to the two previous studies, we also used a treatment with flutamide (Flu), an androgen receptor blocker. Flutamide competes with T for binding to androgen receptors and thereby reduces the effectiveness of T. Hence, we created three experimental groups: one group received T implants, one group received Flu implants, and one group received empty implants as a control (C). To investigate the potential long-term effects of short-term exposure to elevated T, we also examined female EPO rates in the breeding season 1 yr after the manipulation.

As it was previously found that increased T levels in females negatively affected the number of EPO (García-Vigón et al. 2008; but see Gerlach & Ketterson 2013), an experimental increase in T levels may also reduce the number of EPO in blue tits. Conversely, blocking T receptors with Flu may then increase the number of EPO. Alternatively, if T in females decreases choosiness (McGlothlin et al. 2004), or increases sexual activity or male-like behaviour (Lank et al. 1999; Ketterson et al. 2005; Lahaye et al. 2012), then experimentally increased T levels may lead to an increase in the number of EPO. A possible effect of our hormone treatments on female EPO rate may also be mediated by mate preferences of the male extra-pair partners (Ketterson et al. 2005). Therefore, we also aimed to test whether possible differences in the number of EPO are potentially related to differences in female attractiveness due to the hormone treatment by performing male mate choice tests in outdoor cages.

Materials and Methods

Study Area and Species

Data were collected during the breeding seasons (Mar.–Jun.) of 2010 and 2011 at 'De Vosbergen' estate in the north-east of The Netherlands (53°08′N, 06°35′E). The study area contains 188 blue tit nest boxes that are distributed over a 54-ha forest. The forest is a mix of deciduous and coniferous trees (Aminasab et al. 2016). Blue tits have biparental care with both sexes defending the nest and feeding the nestlings, but only the female builds the nest and incubates the eggs (Cramp & Perrins 1993). The blue tit population has been part of a long-term study since 2001 (for further details, see Korsten 2006).

Hormone Implants and Monitoring: A Field Experiment

In 2010 and 2011 at the beginning of the breeding season (25 Mar.–13 Apr. in 2010, 23 Mar.–13 Apr. in 2011), all nest boxes were checked daily for nest building activity. Once nest material was present, females were caught and randomly assigned to either a T (2010: n = 21; 2011: n = 20), C (2010: n = 18; 2011: n = 16) or Flu treatment (2010: n = 20; 2011: n = 21). Each of the females was only implanted once, either as a control or as a T- or Flu-treated individual. Thus, females included in the experiment in 2010 were not implanted in 2011 (neither as a control nor as a T/Flu individual). On average, females were caught 14 d (range 7–23 d) before the onset of egg laying in 2010 and 13 d (7–26 d) in 2011. Females were caught inside their nest box during the day (2010: n = 45 (T = 17, C = 13, Flu = 15); 2011: n = 57) or at night when roosting in their nest box (2010: n = 14 (T = 4, C = 5, Flu = 5). We think it is unlikely that the different catching methods had an effect on our results. Of the 14 females that were captured at night, two had EPO in their nest (one T and one Flu female). Two of these 14 females were not resighted in the population after implantation. Thus, 17% (2/12) of broods of females that were captured at night contained at least one EPO. This percentage hardly differs from the 19% (8/42) of broods that contained at least one EPO in 2010.

In order to give the females hormone-releasing implants, they were transported to and arrived at the centrally located field station within 10 min. after capture. Before implantation, a blood sample was taken from the wing vein for hormone and molecular parentage analyses. The blood sample was always taken immediately after the bird arrived at the field station, within 15 min after capture. To take a blood sample, the wing vein was punctured with a sterile needle (Terumo, 27 g × 0.75; 0.4 × 20 mm) and approximately 100–120 μl of blood was collected with heparinized micro-haematocrit capillaries. For the hormone analyses, the blood was transferred into an Eppendorf tube. Within 2 h of sampling, the blood was centrifuged for 10 min at 2173 g. Plasma was collected and stored in a −20°C freezer until analysis. For the molecular parentage analyses, a minimum of 5 μl from the collected blood was transferred to 90% ethanol for storage until DNA extraction. After blood sampling, each female received one implant that was placed subcutaneously along the left flank under local anaesthesia (xylocaine 10% spray). The small wound was sutured with tissue glue (1 × 0.5 ml Histoacryl, Braun, Germany). The implant consisted of a silicone tube that was sealed on both ends with 1-mm silicon glue. The T and control implants were 4.4 mm long (id. 0.5 mm; od. 1.0 mm) and filled with crystalline T (Sigma©) or left empty, respectively. The Flu implants were 7.2 mm long (id. 1.47 mm; od. 1.96 mm) and filled with crystalline flutamide (Sigma©). We followed the procedure of Van Duyse et al. (2005) when choosing the doses to be administered to the Flu females. Following the implantation, the females were weighed. There was no significant difference in body mass among the treatment groups, although there was a difference in mean body mass between years (GLM: treatment: $F_{2,116} = 0.17$, p = 0.84; year: $F_{1,116} = 4.38$, p = 0.04; mean body mass 2010 = 10.94 g ± 0.07 SE; 2011 = 10.72 g ± 0.07 SE). After implantation, 12 birds were not resighted in the population in the same year (2010: n = 6 (T = 2, C = 2, Flu = 2); 2011: n = 6 (T = 1, C = 2, Flu = 3)).

Following the implantation, nest boxes were checked daily until the onset of egg laying. Females were recaptured after the second egg was laid to remove the implant. On average, females were implanted for 17.4 ± 5.4 (SD) days in 2010 (T = 17.5 ± 4.9, C = 17.8 ± 7.0, Flu = 16.9 ± 4.7) and 17.8 ± 4.4 d in 2011 (T = 19.2 ± 5.1, C = 18.7 ± 4.3, Flu = 15.8 ± 2.7). In 2010, the second egg was collected and replaced with a dummy egg for a different experiment. The females were either recaptured with a mist net (2010: n = 30; 2011: n = 11), at night (2010: n = 2) or with a hand net (following De Heij et al. 2008) when leaving their nest box in the morning (2010: n = 15; 2011: n = 37). During recapture, a second blood sample was taken to measure the effects of the implants on

plasma T levels. This blood sample was always taken within 10 min after capture. After implant removal, females were returned to their territory or nest box (when captured at night) within 30 min after capture.

Nest checks were resumed near the end of the incubation period, around 3 d before the anticipated hatch date (9 d after laying the last egg), to determine the date of hatching of the first egg for each nest. When the nestlings were 3 d old (day of hatching of the first egg was taken as day 0), a blood sample (5–25 μl) was taken for molecular parentage analysis. The nestlings were individually marked by clipping the tip of one or two toenails. Of the 104 females that were implanted and resighted, 91 successfully raised nestlings until the age of 3 d old (2010: n = (T = 16, C = 14, Flu = 15); 2011: n = (T = 16, C = 13, Flu = 17)). Thirteen of the resighted birds were unsuccessful in raising nestlings until the age of 3 d old (2010: n = (T = 1, C = 3, Flu = 3); 2011: n = (T = 3, C = 1, Flu = 2)). On day 8 after the day of hatching, all nestlings received a metal ring with a unique number. Social partners of the females were caught in the nest box with flap traps for blood sampling when broods were 9–14 d old. Near the end of the nestling phase, daily nest checks were recommenced to determine total numbers of fledged young. Twenty-five females of 2010 (T = 9, C = 9 and Flu = 7) were breeding in the population in 2011, and 16 females of 2011 were breeding in the population in 2012 (T = 8, C = 5 and Flu = 3). Of these 31 females that were breeding in the population 1 yr after they were implanted, we included 28 females to test whether there were any long-term effects of short-term hormone treatments on the number of EPO.

Molecular Parentage

DNA was extracted from the blood samples of both parents and nestlings using a Chelex extraction method (for blood samples; Walsh et al. 1991). To assign parentage, parents and nestlings were genotyped for eight microsatellite loci, Ase18 (Richardson et al. 2001), Pca3, Pca7 and Pca8 (Dawson et al. 2000), PmaGAn27 and PmaTGAn45 (Saladin et al. 2003), Pocc6 (Bensch et al. 1997) and Pdo5 (Griffith et al. 1999). PCRs were carried out in 10 μl volume using 20–50 ng of template DNA, a QIAGEN Multiplex PCR Kit and the manufacturer's protocol. The microsatellites were amplified in two separate multiplex panels (panel 1: Ase18, Pca8, PmaGAn27, Pdo5; panel 2: Pca3, Pca7, PmaTGAn45, Pocc6) using the following PCR programme: 15 min. 95°C, 35 cycles of 94°C for 30 s, 55°C for 90 s and 72°C for 60 s, followed by 60°C for 30 min. Fluorescently labelled PCR products were separated on an AB3730 DNA analyser. Subsequently allele lengths were determined using GENEMAPPER 4.0 software. For the blood samples for which markers were not visible, analyses were carried out again. In the case of too small blood samples, markers were still not visible during the second analysis. For these blood samples, we could not identify the microsatellite genotypes. Using Cervus 3.0 (Kalinowski et al. 2007), mean exclusion probability of the eight markers was calculated to be 0.999951 for the first (female) parent and 0.999999 for the second (male) parent (given the genotype of the first parent). Maternity of the social female was confirmed by the microsatellite data for all nestlings. Paternity of the social male was excluded if there were at least two mismatches between the social father's and offspring's genotype, and those nestlings were regarded as EPO (following Magrath et al. 2009). We only determined whether nestlings were extra-pair or not and did assign extra-pair sires.

Mate Choice Experiment using Captive Birds

For the mate choice experiment conducted in 2011, eight female and eight male captive blue tits were used. The birds were hand-raised in 2007 and 2009 (for details, see Vedder et al. 2010) and males and females were always separately housed in aviaries. One month before the experiment, all birds were housed in individual cages (80 × 40 × 40 cm) placed within a large outdoor aviary with *ad libitum* food and water. Early Apr. 2011, four females received a T implant and four females a Flu implant (for details on implants, see above). Given the limited sample size, a control group was not included in the experiment and we predicted the T and Flu treatment to have potentially opposing effects on female attractiveness. After a first round of mate choice trials (which started 6 d after implantation), we reversed the treatments of the females for a second round of mate choice trials.

Upon capture just prior to the implantation, a blood sample was taken to determine basal testosterone levels. Seven days after implantation, a second sample was taken to measure the effect of the implants on plasma T concentrations. After 18 d, a third blood sample was taken and the implants were removed. Subsequently, the birds were allowed to recover for 16 d after implant removal. The half-life of T is of the order of minutes to an hour (Adkins-Regan 2005). Thus, the period of 16 d would be sufficient time for circulating T levels to return to baseline levels before the next round of implantations and subsequent mate

choice trials. In the next round of implantations, the treatments were reversed such that the females that were T-treated first subsequently received a Flu implant and vice versa. The blood sampling scheme of the second experiment was equal to the first experiment. The method of blood sampling was equal to the field experiment. Within 2 h after sampling, the blood was centrifuged for 10 min at 2173 g. After centrifugation, plasma was removed and stored in a −20°C freezer until analyses.

The mate choice experiment started 6 d after implantation (11th and 12th of Apr. and the 21st and 22nd of May). For each test, two females and one male were placed into the mate choice set-up in the afternoon preceding the test the following morning to let them habituate to the set-up (Fig. 1). During habituation, the compartments were separated by opaque partitions so that the birds could not see each other. The females were unfamiliar to the males. Both females were put in small cages (25 × 25 × 40 cm) with wire mesh sides that were placed along both sides of a larger cage (80 × 40 × 40 cm), where the male was housed. The male had a long perch over the whole length of his cage. The perch was divided into five equal parts of 16 cm each to enable scoring of the position of the male. The females also had a perch that was placed at the same height as the male perch. At both sides of the male cage, there was a small window through which the male could see the female.

The preference tests were all conducted between 9:00 and 10:00 am. The tests were conducted in the morning because in free-living blue tits most EPC take place early in the morning (Kempenaers et al. 1992). The test started by removing the opaque partition between the male and the females and lasted for two periods of 15 min. To account for a potential side bias in the males, females were swapped between the two sides of the male cage after the first 15-min. period (by swapping their cages around), after which the preference test was continued for another 15 min. At the end of the test, all birds were returned to their original cages. All tests were recorded with a digital video camera (JVC Everio) and analysed in the laboratory by one observer who was blind to the treatment of the females. Eight days after the reversal of the implantation treatments, the females and males underwent a second preference test. For this test, a new set of four males was used. Thus, in total, eight male preference scores were obtained. The same female dyads were used during the second round of preference tests.

When analysing the video recordings, the amount of time a male spent in each of the five perch areas was scored. To present the preferences of males for either the T- or Flu-treated females, we calculated the percentage of time a male spent in close proximity to the Flu female of the total time spent in close proximity to either the Flu or T female (i.e. the time spent at either one of the extremes of the perch). Thus, a percentage above 50% represented a preference for the Flu female, whereas a percentage below 50% represented a preference for the T female.

Hormone Analyses

Plasma samples were thawed and hormone concentrations were measured using radioimmunoassays (RIAs). After measuring plasma volume of all samples, 50 µl radioactively labelled testosterone (Perkin Elmer Life and Analytical Science BV) was added to calculate the recovery of testosterone during the extraction process. After an incubation time of 1 h, 2.5 ml diethyl ether/petroleum benzine (70:30) was added and samples were vortexed and centrifuged. Samples were snap-frozen by a mixture of ethanol and dry ice and decanted. The supernatant was dried under streaming nitrogen, the remaining pellet was again dissolved in 1.5 ml 70% methanol, and samples were stored overnight at −20°C. In the morning, samples were centrifuged, the methanol phase was decanted and the samples were dried again under streaming nitrogen. The pellet was resuspended in 200 µl PBS buffer; 30 µl of this mixture was used for measuring recoveries. Average recovery rate for testosterone was

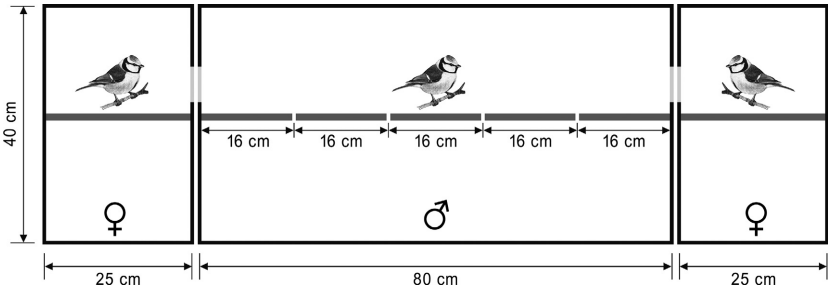

Fig. 1: The experimental set-up of the mate choice test. The male could see the females through a small window covered with wire mesh.

82% ± 4.67 (SD). Hormone measures for all samples were corrected for extraction efficiency. Plasma samples were analysed in one assay per experiment (field or captive female plasma samples) using a commercial kit (Orion Diagnostica, Spectria Testosterone RIA kit, Espoo, Finland) with a sensitivity of 0.04 ng/ml testosterone and cross-reactivities of 4.5% with 5α-dihydrotestosterone (DHT) and 0.01% with androstenedione (A4). The dilution curve ran parallel to the standard curve. Based on the standard curve, values smaller than 0.10 ng/ml were considered to be below the detection limit. Due to a freezer defect, our samples were lost in 2010. Therefore, we could only analyse plasma samples from 2011 for the field experiment.

In total, 56 (T = 20, C = 15 and Flu = 21) baseline plasma T samples could be analysed as well as 37 samples (T = 14, C = 8 and Flu = 15), which were collected at recapture. Before implantation, there was no significant difference in plasma T levels among the treatments (GLM; $F_{2,52} = 0.52$, p = 0.59, see Figure S1). The implants significantly increased plasma T concentrations in T females, but did not affect T concentrations of Flu and C females (GLM treatment x sample: $F_{2,32} = 17.29$, $p < 0.0001$, see Figure S1).

In the captive females, all the baseline T levels were below the detection limit. The low T levels may have occurred, because T levels in captive birds are generally lower compared with free-living individuals (Wingfield et al. 1990). Average recovery rate for testosterone in this analysis was 80% ± 9.8 (SD). The intra-assay variation was 4.7%. To analyse the effects of the hormone implants on plasma T levels, we only included the T plasma concentration after implantation, because all of the baseline T levels were below the detection limit.

The order of treatment (first T and then Flu or vice versa) did not have an effect on plasma T levels (paired *t*-test: $t = 1.93$, df = 6, p = 0.10). One female of the first implantation round, who was implanted with Flu, had a T level of 2.74 ng/ml. All the other females that were implanted with Flu had T levels lower than the detection limit. For the statistical analyses, their T levels were set to the minimum of 0.10 ng/ml (=the minimum detection limit). Because there was no order effect of treatment on plasma T levels, order of treatment was not included in the final analysis. The mean plasma T levels (2.66 ± 0.18 ng/ml) after implantation in captive females treated with testosterone were significantly higher than in females treated with Flu (0.30 ± 0.16 ng/ml, Mann–Whitney *U*-test: $U = 98$, p = 0.001 see Data S1).

Data Selection and Sample Sizes

For the analyses of the short-term effects (i.e. within the same breeding season) of the hormone treatments on the EPO rates, only females that were implanted in that particular year were included (including females where the implant could not be removed). Seven of the 91 broods were excluded (2010: T = 3; C = 1; 2011: Flu = 3) because the DNA was not collected from both parents (2011: Flu = 1), or females were disturbed (stopped egg laying for a couple of days) which may have affected their (extra-pair) copulation behaviour (2010: T = 2), or the second egg of the female was not removed in 2010 (T = 1; C = 1), or the female lost her implant (2011: Flu = 2). This left us with 15 T, 12 C and 15 Flu broods in 2010 and 16 T, 13 C and 13 Flu broods in 2011. For testing the effects of our hormone treatment on general breeding parameters (onset of egg laying, clutch size, proportion of offspring alive at day 3 and proportion of offspring fledged), we used the same set of females.

For analyses of the long-term effects of the hormone manipulations on the number of broods with EPO, we only included females in which the implant was successfully removed in the previous year. This is because we wanted to test whether a short-term elevation in T levels at the beginning of the breeding season would have carry-over effects to the following breeding season. For these analyses, 13 of the 41 broods of the females that were treated in 2010 or 2011 and were breeding in the population the following year were excluded (2011:T = 5, C = 3; 2012: T = 1, C = 2, Flu = 2) because either their implant was not removed (n = 8) or was lost (n = 1), or parentage could not be determined (n = 3), or eggs were collected for another experiment (n = 1). In total, 11 T, 9 C and 8 Flu females were used to test whether there was a carry-over effect of the hormone treatment on the occurrence of EPO to the following year (see Table 1 for an overview of the sample sizes).

Statistical Analyses

For testing the effect of treatment on the onset of egg laying and clutch size, general linear models (with a normal error distribution) were used. To test whether our treatment had an effect on the hatching success (the number of offspring alive on day 3 after hatching divided by the clutch size) and fledging success (the number of fledglings divided by the number of young on day 3), we fitted generalized linear models (GLM) with a binomial error distribution and a logit link

function. Treatment and year, as well as their interaction, were included as factors.

To test whether our treatment had an effect on the overall proportion of EPO (of the total number of offspring genotyped per brood), we fitted GLMs with a binomial error distribution and a logit link function. Treatment and year as well as their interaction were included as factors. A similar model was fitted to examine whether the effects of the short-term hormone implants had carried over to the following year. The differences in the proportions of broods with at least one EPO among treatments were analysed with a generalized linear model with a binary error distribution and a logit link function. A brood was scored as 1 when it contained at least one EPO and 0 when it contained no EPO. We included the same predictors as above, as well as the number of offspring genotyped per brood (because the more offspring are genotyped, the greater the likelihood may be to encounter at least one EPO in a brood). To test whether males spent more time in close proximity to Flu compared with T females during the mate choice trials, we used a nonparametric Wilcoxon matched-pairs signed-rank test for paired data comparing the amounts of time spent by each male in close proximity to either the T- or Flu-treated female. All statistical tests are two-tailed and the level for significance was $p < 0.05$. GLMs were implemented in IBM SPSS Statistics (version 20) software.

Table 1: Sample sizes of the statistical analyses

	2010	2011	Total
Implanted			
T	21	20	41
C	18	16	34
Flu	20	21	41
EPO analysis			
T	15	16	31
C	12	13	25
Flu	15	13	28
Long-term effect on EPO rate			
T	4	7	11
C	6	3	9
Flu	7	1	8

Results

Breeding Parameters and Reproductive Success

There was no effect of the hormone treatment on the onset of egg laying (general linear model; treatment: $F_{2,84} = 0.42$, $p = 0.66$; see Table 2). Females started laying later in 2011 than in 2010 (year: $F_{1,84} = 5.72$, $p = 0.019$), but there was no significant interaction with treatment (treatment × year interaction, $F_{2,84} = 0.95$, $p = 0.39$).

There was also no effect of hormone treatment on mean clutch size (treatment: $F_{2,84} = 0.89$, $p = 0.41$; Table 2). Females laid fewer eggs in 2011 than in 2010 (year: $F_{1,84} = 13.36$, $p = 0.0005$), but there was no significant interaction with treatment (treatment × year interaction, $F_{2,84} = 1.55$, $p = 0.22$).

There was no significant effect of hormone treatment on hatching success (treatment: Wald $\chi^2 = 0.54$, $df = 1$, $p = 0.46$; Table 2). The hatching success was higher in 2011 than in 2010 (year: Wald $\chi^2 = 45.47$, $df = 1$, $p < 0.0001$), but there was no significant interaction with treatment (treatment × year: Wald $\chi^2 = 0.73$, $df = 1$, $p = 0.39$).

There was no significant effect of hormone treatment on fledging success (treatment: Wald $\chi^2 = 0.01$, $df = 1$, $p = 0.94$; Table 2). Fledging success was higher in 2011 than in 2010 (year: Wald $\chi^2 = 7.0$, $df = 1$, $p = 0.008$), but there was no significant interaction with treatment (treatment × year: Wald $\chi^2 = 2.01$, $df = 1$, $p = 0.16$). In summary, neither increased levels of testosterone (T treatment) nor a decrease in T effectiveness (Flu treatment) significantly affected any of the investigated breeding parameters or reproductive success.

Table 2: Breeding parameters for nests of testosterone (T)- and flutamide (Flu)-treated females compared with control (C) females. Means ± SE are given. Sample sizes are in brackets

	2010			2011		
	T nests	C nests	Flu nests	T nests	C nests	Flu nests
Onset of egg laying (March date)	47.20 ± 0.81 (15)	47.17 ± 0.90 (12)	47.87 ± 0.65 (15)	51.81 ± 2.27 (16)	50.0 ± 1.45 (13)	48.62 ± 1.47 (13)
Clutch size	11.27 ± 0.60 (15)	11.58 ± 0.40 (12)	11.33 ± 0.44 (15)	9.75 ± 0.41 (16)	9.31 ± 0.52 (13)	10.77 ± 0.41 (13)
Brood size (d3)	7.07 ± 0.88 (15)	7.25 ± 0.96 (12)	8.4 ± 0.49 (15)	7.94 ± 0.58 (16)	7.92 ± 0.76 (13)	10.08 ± 0.50 (13)
Fledgling number	6.33 ± 0.95 (15)	7.92 ± 1.02 (12)	7.0 ± 0.58 (15)	7.25 ± 0.78 (16)	7.62 ± 0.76 (13)	9.15 ± 0.61 (13)

Hormone Implants and Extra-Pair Offspring

In total, 27% (23/84) of broods of the females treated with hormone implants contained at least one EPO (2010: 19%, 8/42; 2011: 36%, 15/42). There were significant main effects of hormone treatment and year on the overall proportion of EPO in broods (Table 3). T-treated females had a lower overall proportion of EPO in their broods compared with C and Flu-treated females (Fig. 2a; T vs. C: Wald $\chi^2 = 10.820$, df = 1, p = 0.001; T vs. Flu: Wald $\chi^2 = 4.028$, df = 1, P = 0.045). The proportion of EPO in broods was higher in 2011 than in 2010 (Fig. 2b). The significant effect of treatment on the proportion of EPO in broods appeared to be largely driven by an effect on the proportion of broods that contained at least one EPO (Table 4; Fig. 3; T vs. C: Wald $\chi^2 = 5.067$, df = 1, p = 0.024; T vs. Flu: Wald $\chi^2 = 6.233$, df = 1, p = 0.013). Remarkably, for broods that had at least one EPO (n = 23), the proportion of EPO was significantly higher for the three T-treated females (Table 5; Fig. 4; T vs. C: Wald $\chi^2 = 4.226$, df = 1, p = 0.040; T vs. Flu: Wald $\chi^2 = 8.080$, df = 1, p = 0.004). The proportion of EPO in EPO-containing broods was significantly higher in 2011 than in 2010 (Table 4).

Table 3: Generalized linear models results for the effect of the three hormone-implant treatments (testosterone, control and flutamide) on the proportion of extra-pair offspring (n = 84 broods). The numbers in bold are statistically significant

Final Model	Estimates	SE	Wald (χ^2)	df	p
Included					
Intercept	−2.562	0.318			
Treatment (ref: C)			**10.937**	**2**	**0.004**
Testosterone	−1.281	0.390	**10.820**	**1**	**0.001**
Flu	−0.487	0.311	2.447	1	0.12
Year (ref: 2010)	1.125	0.320	**12.383**	**1**	**<0.001**
Excluded					
Treatment × year			4.504	2	0.11

Long-Term Effects of Hormone Implants

Neither increased levels of testosterone (T) nor a decrease in T effectiveness (Flu) significantly affected EPO rate in the following breeding season compared with the control treatment (C) in the following breeding season. For the females that were implanted in 2010 and 2011 and were breeding in the population in 2011 and 2012, respectively, 46% (13/28) of the

Table 4: Generalized linear models results for the effect of the three hormone-implant treatments (testosterone, control and flutamide) on the proportion of broods containing extra-pair offspring (n = 84 broods). The numbers in bold are statistically significant

Final Model	Estimates	SE	Wald (χ^2)	df	p
Included					
Intercept	0.575	0.417			
Treatment (*ref: C*)			**6.746**	**2**	**0.034**
Testosterone	1.658	0.737	**5.067**	**1**	**0.024**
Flu	−0.140	0.569	0.061	1	0.80
Excluded					
Year (ref: 2010)			3.448	1	0.063
Treatment × year			1.044	2	0.59
N offspring genotyped			0.679	1	0.41

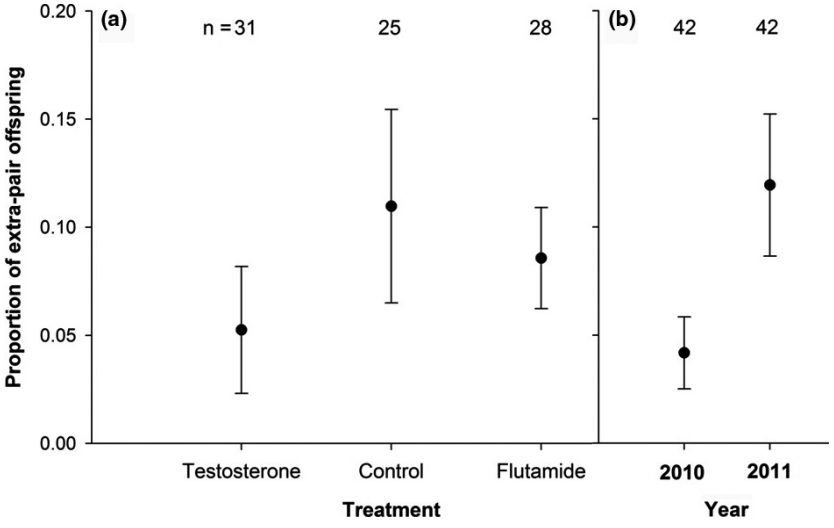

Fig. 2: Proportions of extra-pair offspring (EPO) in broods in relation to treatment (a) and study year (b). Means ± SE are given; n indicates the number of broods.

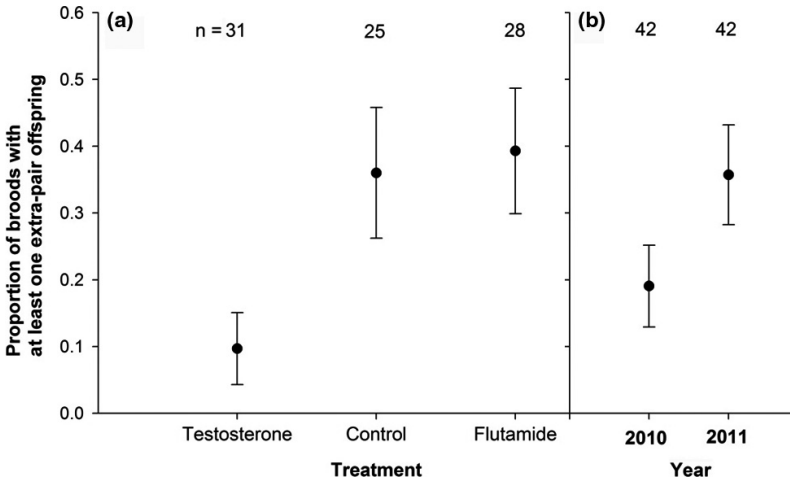

Fig. 3: Proportion of broods with at least one extra-pair offspring (EPO) in relation to treatment (a) and study year (b). Means ± SE are given; n indicates the number of broods.

Table 5: Generalized linear models results for the effect of the three hormone-implant treatments (testosterone, control and flutamide) on the number of extra-pair offspring per brood for broods containing at least one EPO (n = 23 broods). The numbers in bold are statistically significant

Final Model	Estimates	SE	Wald (χ^2)	df	p
Included					
Intercept	−1.356	0.388			
Treatment (ref: C)			**8.200**	**2**	**0.017**
Testosterone	1.123	0.546	**4.226**	**1**	**0.040**
Flu	−0.417	0.347	1.447	1	0.23
Year (ref: 2010)	0.755	0.374	**4.085**	**1**	**0.043**
Excluded					
Treatment × year			2.711	2	0.26

nests contained at least one EPO. There were no long-term (carry-over) effects of the treatments on either the overall proportion of EPO in broods (Wald $\chi^2 = 3.107$, df = 2, p = 0.21) or the proportion of broods with at least one EPO (Wald $\chi^2 = 1.411$, df = 2, p = 0.49).

Male Preference Tests

During the mate choice tests, the average percentage of time the males spent with the Flu females compared to T females was 47%, indicating no clear preference (Wilcoxon test: p = 0.26, Fig. 5). The total amount of time the males spent in close proximity to either one of the females was very short ($\bar{x}$ ± SE: 130 ± 20 s; 7% of total test trial time of 1800 s).

Discussion

Female blue tits that were treated with T had a lower overall proportion of EPO in their nest than control females and Flu-treated females. The lower overall proportion of EPO in nests of T-treated females

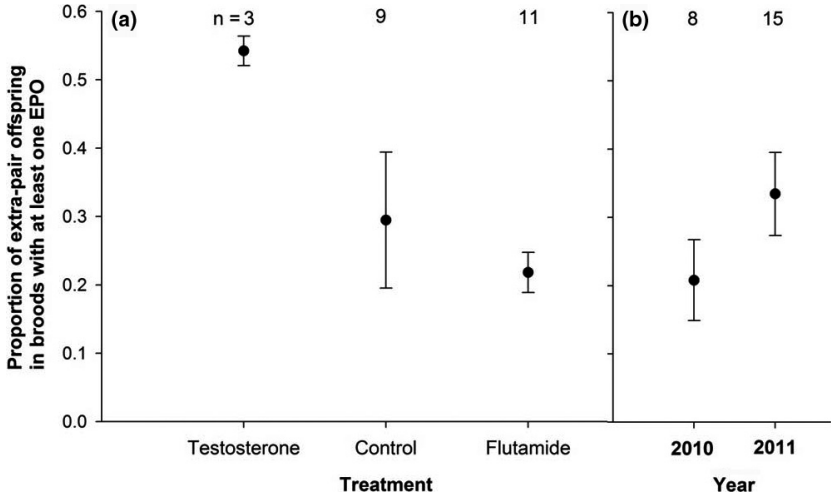

Fig. 4: Proportion of extra-offspring (EPO) in broods with at least one EPO in relation to treatment (a) and study year (b). Means ± SE are given; n indicates the number of broods.

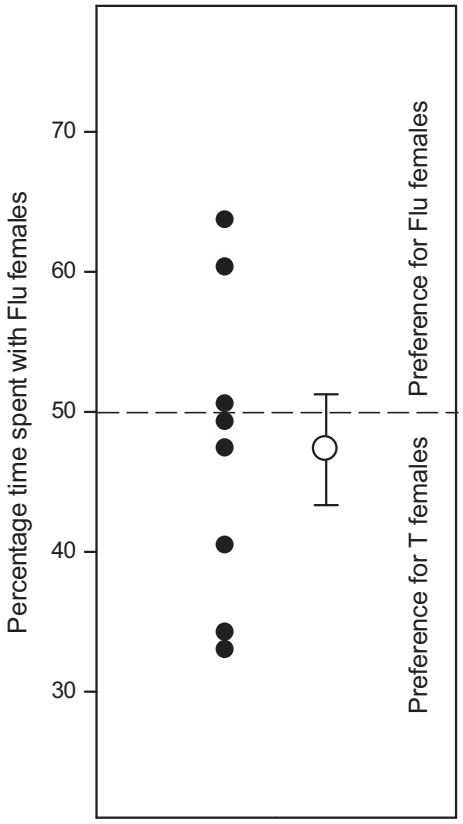

Fig. 5: Male preference measured as the percentage of time (in seconds) a male spent in close proximity to a Flu female relative to a T female (n = 8). A percentage of 50% means an equal amount of time spent with either one of the females. Closed circles indicate the preferences of individual males; the open circle with error bars indicates the mean ± SE.

appeared to be driven by a reduction in the number of nests that contained at least one EPO in T-treated females compared with control and Flu-treated females. The number of nests with EPO was similar for Flu-treated females and control females. Remarkably, the proportion of EPO was higher in nests of T-treated females compared with nests of C and Flu-treated females for the subset of nests that contained at least one EPO. However, sample size for T-treated females was very low for this particular analysis (n = 3), and we prefer not to do too much *post hoc* speculation on why this unexpected pattern occurred. There were no effects of the hormone treatments on general breeding parameters, like lay date and clutch size, or components of reproductive success, like the proportion of offspring successfully fledging.

The proportion of EPO was lower in 2010 compared with 2011. This may have been caused by the collection of the second-laid egg in 2010. Previously, it has been shown for our blue tit study population that most of the EPO occur in the first half of the clutch (Magrath et al. 2009). The observed year difference in the proportion of EPO is therefore consistent with this previous finding. Furthermore, T treatment only affected the probability of having EPO during the period when T levels were experimentally elevated. The effect of T had disappeared in the breeding season 1 yr later. Thus, it appears that there were no carry-over effects of the treatments to the following year. Also, the Flu treatment had no effect on the number of nests with EPO in the breeding season 1 yr after.

To our knowledge, so far only two other studies have investigated the effect of T in females on the number of EPO produced, but with different results. Elevated T had no effect on the number of EPO in the dark-eyed junco (Gerlach & Ketterson 2013), whereas the proportion of EPO in T-treated females was reduced in the spotless starling (García-Vigón et al. 2008). The results of the latter study are congruent with our findings. However, in contrast to our study, T-treated female starlings also had fewer nests with EPO in the subsequent years (García-Vigón et al. 2008). In our study, we did not find carry-over effects of the hormone treatments to the following year.

This difference between our study and the study of García-Vigón et al. (2008) could be due to the fact that in the latter study the implants were not removed. Therefore, in the spotless starlings, T levels probably remained elevated for a longer period compared with our study where hormone implants were removed just after the onset of egg laying, which may more closely mimic naturally occurring peaks in female T levels. The observed long-term effect of elevated T levels on the number of EPO in the spotless starlings might have been the result of long-lasting phenotypic changes (Staub & DeBeer 1997; Abitbol et al. 1999; Roberts et al. 2009). For example, the voice of women injected with testosterone develops masculine characteristics and these changes are irreversible (Abitnol et al. 1999). In our study, the experimental elevation of T was likely too short to cause any long-lasting physiological changes.

There are at least two possible hypotheses to explain why females treated with testosterone had fewer EPO. First, T may have reduced female attractiveness to males and therefore females acquired fewer EPCs. Male dark-eyed Juncos were less attracted to females treated with T than to C females (Ketterson et al. 2005). It has been suggested that this reduction in attractiveness might be the result of females with elevated T levels showing male-like mating behaviour (García-Vigón et al. 2008). However, a more recent study on budgerigars (*Melopsittacus undulatus*) did not

find that T-induced male-like behaviour in females negatively affected their attractiveness to males (Lahaye et al. 2013). These findings are in agreement with the results of our mate choice experiment where we also did not find a difference in male preference between Flu- and T-treated females. However, some caution should be taken when interpreting our results: (1) the sample size used in our mate choice experiment was small. Therefore, it was difficult to control for individual variation in mate preference among males; (2) the experiment was conducted in an unnatural setting with hand-reared birds kept in captivity, thus making it difficult to directly link these results to the findings of our field experiment; (3) the total amount of time the males spent with the females, regardless of female treatment, was rather short. This may indicate that males may have had relatively low motivation to express mate choice-related behaviours.

Although our results should be interpreted with caution (see limitations discussed above), the fact that two studies found no evidence for experimentally elevated T levels to reduce female attractiveness might indicate that the reduction of EPO is, most likely, not explained by a reduction in female attractiveness. A more parsimonious hypothesis explaining why females treated with T had fewer EPO is that T inhibits EPC-seeking behaviour in females. In blue tits, females actively go out to seek EPCs (Kempenaers et al. 1992). Therefore, it is more likely that the effect of experimentally elevated T on the number of EPO was the result of a reduction in the females' extra-pair mate-seeking behaviour rather than a change in female attractiveness. This potential change in mate-seeking behaviour might be an indirect effect of a testosterone-induced shift in behaviour from mate seeking towards behaviours like aggression (Sandell 2007; De Jong 2013) and nest building (De Jong 2013). It is also possible that T females mated more frequently with their own partner and therefore had fewer EPO. To improve our understanding of the behavioural mechanisms involved, future studies should particularly focus on the effect of female T on mating frequencies of females with both their social male and extra-pair males.

In contrast to elevated T, we did not find an effect of Flu on the proportion of EPO, with similar proportions of EPO in the Flu and control groups. Thus, reducing the effectiveness of T does not appear to influence the probability of extra-pair paternity in a female's nest. This may also be due to the fact that T may be aromatized into oestradiol (Nelson 2011). In females, oestrogens and progesterone were shown to exert stronger effects on female mating behaviour than testosterone (Adkins-Regan 2005). For example, experimentally elevated oestrogen induces mating behaviour in many female vertebrate species (Moore 1982; Takahashi 1990; Tokarz & Crews 1980; but see Hunt 1997). Thus, the aromatization of T into oestradiol may explain why the Flu treatment did not affect female (extra-pair) mating behaviour. However, a previous study has shown that excess in T levels, from T implants, was not aromatized into oestradiol (Clotfelter et al. 2004). Therefore, the lack of an effect of flutamide on female (extra-pair) mating behaviour is probably not the result of the conversion of T into oestradiol. There is also evidence that a prolonged antiandrogen treatment can lead to a strong upregulation of androgen receptors, which leads to a higher sensitivity of tissues to low concentrations of androgens (Chen et al. 2004). Thus, flutamide can change from an antiandrogen to androgen agonist. This may even result in a reversed effect on the behaviour measured (Fusani et al. 2007). This potential change in the action of flutamide might also explain why we did not find any difference between the Flu and C females.

Finally, one needs to consider the possibility that females may both express extra-pair mating behaviour and peaks in circulating T at the start of breeding as a non-adaptive side effect of selection on males (Ketterson et al. 2005; Forstmeier et al. 2011). As the male and female sex largely share the same genome, many traits may be genetically correlated between the sexes, leading to correlated evolutionary responses to selection (e.g. Cox & Calsbeek 2009; Poissant et al. 2010). In case past selection has strongly favoured certain reproductive traits in the male sex, the same traits may be expressed in the female sex as a result of a correlated evolutionary response. Indeed, in support of this idea, it has recently been found that male and female extra-pair behaviours are strongly genetically correlated (Forstmeier et al. 2011).

Concluding Remarks

We found that an experimental increase in T levels in females reduces the proportion of EPO in their broods. This effect was mainly driven by a lower proportion of broods with at least one EPO in nests of T-treated females. The effect of the T treatment did not carry over to the breeding season in the following year and appears less likely to be caused by a T-induced reduction in female attractiveness. Future studies should investigate the underlying behavioural mechanisms

causing T to reduce the proportion of EPO. These studies should, for example, investigate whether the reduction in EPO in the broods of females with elevated T is driven by a decrease in the frequency of extra-pair mating and/or increase in the frequency of within-pair mating.

Acknowledgements

We thank P. Wolfs, C. Kuipers, O. Vedder and T. Egberts for field assistance, C. Kujpers, B. de Vries and I. Weites for laboratory assistance and E. Schut and V. Hulst for assisting with the mate choice experiment. The study was conducted in full compliance with Dutch law and regulations. Funding from the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT Vlaanderen) and from the Dr. J.L. Dobberke Foundation was provided to BdJ. PK was financially supported by grants from the Netherlands Organisation for Scientific Research (NWO, 821.01.008) and the European Union's Seventh Framework Programme (PCIG10-GA-2011-304280).

Literature Cited

Abitbol, J., Abitbol, P. & Abitbol, B. 1999: Sex hormones and the female voice. J. Voice **13**, 424—446.

Adkins-Regan, E. 2005: Hormones and animal social behavior. Princeton Univ. Press, Princeton.

Alatalo, R. V., Hoglund, J., Lundberg, A., Rintamaki, P. T. & Silverin, B. 1996: Testosterone and male mating success on the black grouse leks. Proc. R. Soc. Lond. B Biol. Sci. **263**, 1697—1702.

Alonso-Alvarez, C. 2001: Effects of testosterone implants on pair behaviour during incubation in the Yellow-legged Gull *Larus cachinnans*. J. Avian Biol. **32**, 326—332.

Amininasab, S. M., Vedder, O., Schut, E., de Jong, B., Magrath, M. J. L., Korsten, P. & Komdeur, J. 2016: Influence of fine-scale habitat structure on nest-site occupancy, laying date and clutch size in Blue Tits *Cyanistes caeruleus*. Acta Oecologica **70**, 37—44.

Bensch, S., Price, T. & Kohn, J. 1997: Isolation and characterization of microsatellite loci in a Phylloscopus warbler. Mol. Ecol. **6**, 91—92.

Birkhead, T. R. & Møller, A. P. 1992: Sperm Competition in Birds: Evolutionary Causes and Consequences. Academic Press, London.

Brommer, J. E., Korsten, P., Bouwman, K. A., Berg, M. L. & Komdeur, J. 2007: Is extrapair mating random? On the probability distribution of extrapair young in avian broods. Behav. Ecol. **18**, 895—904.

Chen, C. D., Welsbie, D. S., Tran, C., Baek, S. H., Chen, R., Vessella, R., Rosenfeld, M. G. & Sawyers, C. L. 2004: Molecular determinants of resistance to antiandrogen therapy. Nature Med. **10**, 33—39.

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

Cox, R. M. & Calsbeek, R. 2009: Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. Am. Nat. **173**, 176—187.

Cramp, S. & Perrins, C. M. 1993: Birds of the Western Palearctic, Vol. **7**. Oxford Univ Press, Oxford.

Dawson, D., Hanotte, O., Greig, C., Stewart, I. & Burke, T. 2000: Polymorphic microsatellites in the blue tit *Parus caeruleus* and their cross-species utility in 20 songbird families. Mol. Ecol. **9**, 1941—1944.

De Heij, M. E., Ubels, R., Visser, G. H. & Tinbergen, J. M. 2008: Female great tits *Parus major* do not increase their daily energy expenditure when incubating enlarged clutches. J. Avian Biol. **39**, 121—126.

De Jong, B. 2013: Testosterone a female hormone: testing the function and evolution of testosterone in female birds. Univ. of Groningen, Groningen.

De Ridder, E., Pinxten, R. & Eens, M. 2000: Experimental evidence of a testosterone-induced shift from paternal to mating behaviour in a facultatively polygynous songbird. Behav. Ecol. Sociobiol. **49**, 24—30.

van Duyse, E., Pinxten, R., Snoeijs, T. & Eens, M. 2005: Simultaneous treatment with an aromatase inhibitor and an anti-androgen decreases the likelihood of dawn song in free-living male great tits, *Parus major*. Horm and Behav. **48**, 243—251.

Edler, R., Goymann, W., Schwabl, I. & Friedl, T. W. P. 2011: Experimentally elevated testosterone levels enhance courtship behaviour and territoriality but depress acquired immune response in Red Bishops *Euplectes orix*. Ibis **153**, 46—58.

Enstrom, D. A., Ketterson, E. D. & Nolan, V. Jr 1997: Testosterone and mate choice in the dark-eyed junco. Anim. Behav. **54**, 1135—1146.

Foerster, K. & Kempenaers, B. 2004: Experimentally elevated plasma levels of testosterone do not increase male reproductive success in blue tits. Behav. Ecol. Sociobiol. **56**, 482—490.

Forstmeier, W., Martin, K., Bolund, E., Schielzeth, H. & Kempenaers, B. 2011: female extra-pair mating behavior can evolve via indirect selection on males. Proc. Natl Acad. Sci. USA **108**, 10608—10613.

Fusani, L., Day, L. B., Canoine, V., Reinemann, D., Hernandez, E. & Schlinger, B. A. 2007: Androgen and the elaborate courtship behavior of a tropical lekking bird. Horm. Behav. **51**, 62—68.

Garamszegi, L. Z., Eens, M., Hurtrez-Bousses, S. & Moller, A. P. 2005: Testosterone, testes size and mating success

in birds: a comparative study. Horm. Behav. **46**, 111—112.

García-Vigón, E., Cordero, P. J. & Veiga, J. P. 2008: Exogenous testosterone in female spotless starlings reduces their rate of extrapair offspring. Anim. Behav. **76**, 345—353.

Gerlach, N. M. & Ketterson, E. D. 2013: Experimental elevation of testosterone lowers fitness in female dark-eyed juncos. Horm. Behav. **63**, 782—790.

Gerlach, N. M., McGlothlin, J. W., Parker, P. G. & Ketterson, E. D. 2012: Promiscuous mating produces offspring with higher lifetime fitness. Proc. R. Soc. **279**, 860—866.

Griffith, S. C. & Immler, S. 2009: Female infidelity and genetic compatibility in birds: the role of the genetically loaded raffle in understanding the function of extrapair paternity. J. Avian Biol. **40**, 97—101.

Griffith, S., Stewart, I., Dawson, D., Owens, I. & Burke, T. 1999: Contrasting levels of extra-pair paternity in mainland and island populations of the house sparrow (*Passer domesticus*): is there an "island effect"? Biol. J. Linn. Soc. **68**, 303—316.

Griffith, S. C., Owens, I. P. F. & Thuman, K. A. 2002: Extra pair paternity in birds: a review of interspecific variation and adaptive function. Mol. Ecol. **11**, 2195—2212.

Goyman, W. & Wingfield, J. C. 2014: Male-to-female testosterone ratios, dimporphism, and life history - what does it really tell us? Behav. Ecol. **25**, 685—699.

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

Hunt, K.E. 1997: Testosterone, Estrogen, and Breeding Behaviour in an Arctic Bird, The Lapland Longspur. Univ. of Washington, Seattle

Jennions, M. D. & Petrie, M. 2000: Why do females mate multiply? A review of the genetic benefits. Biol. Rev. Camb. Philos. Soc. **75**, 21—64.

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

Kempenaers, B. 1995: Polygyny in the blue tit – intra-sexual and inter-sexual conflicts. Anim. Behav. **49**, 1047—1064.

Kempenaers, B., Verheyen, G. R., Vandenbroeck, M., Burke, T., Vanbroeckhoven, C. & Dhondt, A. A. 1992: Extra-pair paternity results from female preference for high-quality males in the blue tit. Nature **357**, 494—496.

Ketterson, E. D., Nolan, V. & Sandell, M. 2005: Testosterone in females: mediator of adaptive traits, constraint on sexual dimorphism, or both? Am. Nat. **166**, S85—S98.

Korsten, P. 2006: Avian sex allocation and ornamental coloration: a study on blue tits. Univ. of Groningen, Groningen.

Lahaye, S. E. P., Eens, M., Darras, V. M. & Pinxten, R. 2012: Testosterone stimulates the expression of male-typical socio-sexual and song behaviors in female budgerigars (*Melopsittacus undulatus*): an experimental study. Gen. Comp. Endocrinol. **178**, 82—88.

Lahaye, S. E. P., Eens, M., Darras, V. M. & Pinxten, R. 2013: Hot or not: the effects of exogenous testosterone on female attractiveness to male conspecifics in the budgerigar. PLoS ONE **8**, e74005.

Lank, D. B., Coupe, M. & Wynne-Edwards, K. E. 1999: Testosterone-induced male traits in female ruffs (*Philomachus pugnax*): autosomal inheritance and gender differentiation. Proc. Biol. Sci. **266**, 2323—2330.

Magrath, M. J. L., Vedder, O., van der Velde, M. & Komdeur, J. 2009: Maternal effects contribute to the superior performance of extra-pair offspring. Curr. Biol. **19**, 792—797.

McGlothlin, J. W., Neudorf, D. L. H., Casto, J. M., Nolan, V. & Ketterson, E. D. 2004: Elevated testosterone reduces choosiness in female dark-eyed juncos (*Junco hyemalis*): evidence for a hormonal constraint on sexual selection? Proc. Biol. Sci. **271**, 1377—1384.

Møller, A. & Ninni, P. 1998: Sperm competition and sexual selection: a meta-analysis of paternity studies of birds. Behav. Ecol. Sociobiol. **43**, 345—358.

Møller, A. P., Garamszegi, L. Z., Gil, D., Hurtrez-Bousses, S. & Eens, M. 2005: Correlated evolution of male and female testosterone profiles in birds and its consequences. Behav. Ecol. Sociobiol. **58**, 534—544.

Moore, M. C. 1982: Hormonal response of free-living male white-crowned sparrows to experimental manipulation of female sexual behavior. Horm. Behav. **16**, 323—329.

Moore, M. C. 1984: Changes in territorial defense produced by changes in circulating levels of testosterone: a possible hormonal basis for mate-guarding behavior in white-crowned sparrows. Behaviour **88**, 215—226.

Nakamura, M. 1998: Multiple mating and cooperative breeding in polygynandrous alpine accentors. II. Male mating tactics. Anim. Behav. **55**, 277—289.

Nelson, R. J. 2011: An introduction to behavioral endocrinology. Sinauer Associates, Sunderland, MA.

Nespor, A. A., Lukazewicz, M. J., Dooling, R. J. & Ball, G. F. 1996: Testosterone induction of male-like vocalizations in female budgerigars (*Melopsittacus undulatus*). Horm. Behav. **30**, 162—169.

Poissant, J., Wilson, A. J. & Coltman, D. W. 2010: Sex-specific genetic variance and the evolution of sexual dimorphism: a systematic review of cross-sex genetic correlations. Evolution **64**, 97—107.

Raouf, S. A., Parker, P. G., Ketterson, E. D., Nolan, V. & Ziegenfus, C. 1997: Testosterone affects reproductive success by influencing extra-pair fertilizations in male dark-eyed juncos (Aves: *Junco hyemalis*). Proc. Biol. Sci. **264**, 1599—1603.

Richardson, D., Jury, F., Blaakmeer, K., Komdeur, J. & Burke, T. 2001: Parentage assignment and extra-group paternity in a cooperative breeder: the Seychelles warbler (*Acrocephalus sechellensis*). Mol. Ecol. **10**, 2263—2273.

Roberts, M. L., Ras, E. & Peters, A. 2009: Testosterone increases UV reflectance of sexually selected crown plumage in male blue tits. Behav. Ecol. **20**, 535—541.

Saino, N. & Møller, A. P. 1995: Testosterone correlates of mate guarding, singing and aggressive behavior in male barn swallows (*Hirundo rustica*). Anim. Behav. **49**, 465—472.

Saladin, V., Bonfils, D., Binz, T. & Richner, H. 2003: Isolation and characterization of 16 microsatellite loci in the Great Tit *Parus major*. Mol. Ecol. Notes **3**, 520—522.

Sandell, M. I. 2007: Exogenous testosterone increases female aggression in the European starling (*Sturnus vulgaris*). Behav. Ecol. Sociobiol. **62**, 255—262.

Sheldon, B. C. 1994: Male phenotype, fertility, and the pursuit of extra-pair copulations by female birds. Proc. R. Soc. B Biol. Sci. **257**, 25—30.

Staub, N. L. & DeBeer, M. 1997: The role of androgens in female vertebrates. Gen. Comp. Endocrinol. **108**, 1—24.

Takahashi, L. K. 1990: Hormonal regulation of sociosexual behavior in female mammals. Neurosci. Biobehav. Rev. **14**, 403—413.

Tokarz, R. & Crews, D. 1980: Induction of sexual receptivity in the female lizard, *Anolis carolinensis*: effects of estrogen and the antiestrogen CI-628. Horm. Behav. **14**, 33—45.

Vedder, O., Schut, E., Magrath, M. J. L. & Komdeur, J. 2010: Ultraviolet crown colouration affects contest outcomes among male blue tits, but only in the absence of prior encounters. Funct. Ecol. **24**, 417—425.

Vedder, O., Komdeur, J., van der Velde, M., Schut, E. & Magrath, M. J. L. 2011: Polygyny and extra-pair paternity enhance the opportunity for sexual selection in blue tits. Behav. Ecol. Sociobiol. **65**, 741—752.

Walsh, P., Metzger, D. & Higucji, R. 1991: Chelex-100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques **10**, 506—513.

Wingfield, J. C., Hegner, R. E., Dufty, A. M. & Ball, G. F. 1990: The challenge hypothesis – theoretical implications for patterns of testosterone secretion, mating Systems, and breeding strategies. Am. Nat. **136**, 829—846.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web-site:

Figure S1: Levels of plasma T concentrations (ng/ml) before (baseline, plotted left of the dashed line) and after implantation (right from the dashed line).

Figure S2: Testosterone levels before implantation and after implantation. The black bars are the testosterone females, the gray bars are the control females and the white bars the flutamide females. Bars represent the means ± SEM per treatment for all the females that were tested. The number of females is presented above the bars.

Table S1: Median baseline testosterone levels and its range in ng/ml of blue tit females during the nest building phase.

Data S1: Effect of hormone implants on plasma testosterone (T) level.