Title

Fine-tuning of seasonal timing of breeding is regulated downstream in the underlying neuro-endocrine system in a small songbird

Authors

Irene Verhagen[1], Veronika N. Laine[1], A. Christa Mateman[1], Agata Pijl[1], Ruben de Wit[1], Bart van Lith[1], Willem Kamphuis[2], Heidi M. Viitaniemi[3], Tony D. Williams[4], Samuel P. Caro[5], Simone L. Meddle[6], Phillip Gienapp[1], Kees van Oers[1] & Marcel E. Visser[1]

[1] Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW), The Netherlands

[2] Netherlands Institute for Neuroscience (NIN-KNAW), The Netherlands

[3] Department of Biosciences, University of Helsinki

[4] Department of Biological Sciences, Simon Fraser University, Canada

[5] Departement d'Ecologie Evolutive, Centre d'Ecologie Fonctionnelle & Evolutive, France

[6] Department of Behavioural Neuroendocrinology, University of Edinburgh, Scotland

Corresponding author: Irene Verhagen, i.verhagen@nioo.knaw.nl

Running title: Downstream regulation of avian breeding time

Summary statement: This unique experiment shows that variation in candidate gene expression in ovary and liver explains variation in egg-laying dates in a songbird indicative of downstream regulation of timing of breeding.

Abstract

Timing of breeding is under selection in wild populations due to climate change, and understanding the underlying physiological processes mediating timing provides insight in the potential rate of adaptation. Current knowledge on this variation in physiology is, however, mostly limited to males. We assessed whether individual differences in timing of breeding in females are reflected in differences in candidate gene expression and if so, whether these differences occur in the upstream (hypothalamus), or downstream (ovary and liver) parts of the neuroendocrine system. We used 72 female great tits from two generations of lines artificially selected for early and late egg-laying, which were housed in climate controlled aviaries and went through two breeding cycles within one year. In the first breeding season we obtained individual egg-laying dates, while in the second breeding season, using the same individuals, we sampled several tissues at three time points based on timing of the first breeding attempt. For each tissue, mRNA expression levels were measured using qPCR for a set of candidate genes associated with timing of reproduction and subsequently analysed for differences between generations, time points and individual timing of breeding. We found differences in gene expression between generations in all tissues with most pronounced differences in the hypothalamus. Differences between time points, and early and late laying females, were found exclusively in ovary and liver. Altogether, we show that fine-tuning of seasonal timing of breeding, and thereby the opportunity for adaptation in the neuroendocrine system, is regulated mostly downstream in the neuro-endocrine system.

Key words: reproductive timing, individual variation, gene expression, qPCR, *Parus major*

Introduction

Variation in avian seasonal timing of breeding is ultimately rooted in its underlying physiology, as, after transduction and integration of cues, reproductive timing is the outcome of a neuro-endocrine cascade along the so-called hypothalamic-pituitary-gonadal-liver-axis (HPGL axis). The hypothalamus, considered as the final integration point of environmental cues, the pituitary gland, and the neural centres are widely assumed to primarily guide top-down hormonal regulation and in this way direct ovarian function to time breeding (Dawson 2008; Tsutsui et al. 2012). Many studies have therefore focused on these upstream levels of the HPGL axis (Nakane & Yoshimura 2014 and references therein). Though photoperiod, perceived by three types of photoreceptors (Underwood, Steele & Zivkovic 2001), is a proximate cue for birds to time breeding (Sharp 1996; Silverin, Massa & Stokkhan 1993; Wingfield 1993), it cannot solely explain individual year to year variation in timing of breeding, as the change in day length over the season is invariable among years (Bradshaw & Holzapfel 2007; Visser, Both & Lambrechts 2004). A potential explanation for the variation in timing of breeding is an "alternative, female-specific hypothesis" where females use (changes in) supplementary cues to fine-tune downstream mechanisms at the level of the ovary and/or liver and so may regulate vitellogenesis, follicle development and timing of egg-laying (Caro et al. 2009; Lambrechts & Visser 1999; Williams 2012). In general, little work has integrated downstream levels in females, let alone multiple levels of the neuro-endocrine cascade in relation to cues and/or reproductive traits (Cánovas et al., 2014; Laine et al., *subm.*; MacManes et al., 2017; Maruska & Fernald, 2011; Maruska, Levavi-Sivan, Biran, & Fernald, 2011; Perfito, Guardado, Williams, & Bentley, 2015).

Evidence of possible downstream mechanisms regulating timing of breeding has been found in a few occasions. A study in two wild populations of blue tits (*Cyanistes caeruleus*) breeding at different times, suggested that females have similar photoperiodic sensitivities but that the population differences in seasonal timing could be explained by differences in the response of the ovary to

gonadotropins, or the liver to oestrogens (Caro et al. 2009). Work on great tits (*Parus major*) (Schaper, Dawson, Sharp, et al. 2012) and European blackbirds (*Turdus merula*) (Partecke, Van't Hof & Gwinner 2005) showed significant differences in egg laying dates between females from different temperature treatments and populations respectively, but similar plasma luteinizing hormone (LH) levels. Individual variation in luteinizing hormone receptor (LHR) transcript in the testes and developing follicles was found in dark-eyed juncos (*Junco hyemalis*) respectively, but no differences in, again, LH (Bergeon Burns et al. 2014; Needham et al. 2019). A study in male European starlings (*Sturnus vulgaris*) found that the inhibition of gonadal sex steroid secretion is seasonally regulated within the testes by mechanisms involving melatonin receptors and the gonadotropin-inhibiting hormone (GnIH) system present in the gonads (McGuire, Kangas & Bentley 2011). Direct evidence for downstream regulation of timing of breeding was, however, found in female European starlings housed with or without males (Perfito et al. 2015). Female starlings housed with males showed elevated levels of LHR, follicle stimulating hormone receptor (FSHR) and vitellogenin (VTG) mRNA only immediately before, or coincident with, rapid yolk development (RYD), together with increased plasma yolk precursor levels (Perfito et al. 2015). This is consistent with a "lack of ovarian competence" to respond to elevated circulating gonadotropins until just before egg-laying. In addition, when female starlings housed without males were briefly subjected to males, mRNA levels and yolk precursor levels elevated, indicating that the ovary depends on the "supplemental cue" of male presence (Perfito et al. 2015). Multiple, if not all, levels of the HPGL axis need close and simultaneous examination to gain knowledge on, or identify where species differ in executing physiological mechanisms resulting in variation in timing of breeding. This then, would set the stage for understanding where selection could act and how animals could respond via genetic adaptation to changing environments.

A wealth of studies measuring hormone concentrations in circulation, using endocrine and receptor agonists and antagonists to study physiological and behavioural effects, and assessment of protein levels by immunochemistry, have resulted in the extensive knowledge on HPGL axis functioning so

far. However, despite this knowledge and the understanding of which cues (i.e. photoperiod, temperature, food, social cues) influence timing of breeding, understanding of the mechanisms regulating a female's "decision" to initiate egg laying is far behind. Recent and current developments in genomic technologies have started to provide new options to explore and identify the links between genetic and phenotypic variation (Cheviron, Whitehead & Brumfield 2008; Fidler et al. 2007). For the great tit, a model species in ecology and evolution, such tools, including a well annotated reference genome, have recently become available (Derks et al. 2016; Kim et al. 2018; Laine et al. 2016).

Here, we use female great tits from selection lines where birds were genomically selected for either early or late timing of breeding (Gienapp et al. 2019; Verhagen et al. 2019). Birds were subjected to two contrasting temperature environments in climate-controlled aviaries. A recent study in these great tits reports that genes show differential expression under the influence of temperature in the hypothalamus, and, when females were expected to initiate egg laying, genes highly differentially expressed in liver, but especially ovary (Laine et al. *subm.*). However, because pooled samples (three females per sample) were used in that study gene expression levels could not be related to *individual* egg-laying dates. Using samples from the same great tits as in Laine et al (*subm.*), we assess (1) whether individual differences in gene expression levels could explain differences in individual egg-laying dates, and if so, (2) where (upstream or downstream in the HPGL axis) these differences in gene expression occur. By making use of the great tit genome (Laine et al. 2016), we take a candidate gene approach and measure individual expression levels using qPCR. Key genes known to be important mediators in reproductive endocrine pathways upstream (i.e. the hypothalamus) and downstream (i.e. the ovary and liver) in the HPGL axis in female great tits were targeted. In addition, we selected genes, potentially important in reproductive biology, from the abovementioned genome-wide study (Laine et al., *subm.*).

Materials and methods

Selection lines in timing of breeding

Selection lines for early and late timing of breeding in great tits (*Parus major*) were created using bi-directional genomic selection (see Gienapp et al. 2019; Verhagen et al. 2019 for details). To summarize, from wild broods of our long-term study population in the Hoge Veluwe, nestlings (F_1 generation) of which the mother had initiated egg laying either extremely early ('early line') or extremely late ('late line') in the wild, were brought into the aviary-facilities at the NIOO-KNAW (Wageningen, the Netherlands) 10 days post-hatching for further hand raising. Subsequently, chicks were genotyped using a 650 SNP chip (Kim et al. 2018) to predict their 'genomic breeding values' (GEBVs, i.e. the value estimating the relationship between genotype and phenotype based on genetic markers). F_1 generation individuals were, based on their GEBVs, selected for early and late line breeding pairs to produce the F_2 generation in captivity. The F_2 generation eggs were transferred to wild 'foster-nests' for incubation and hatching. F_2 generation chicks were also collected and hand-raised in the laboratory. In their turn, the F_2 offspring were genotyped and selected to produce the F_3 generation, which was then genotyped and selected.

The selection line study results are reported elsewhere (Verhagen et al. 2019). Briefly, we found that on average early line birds laid about six days earlier than late line birds. Further, the difference in average laying date increased (from about 2 to 10 days) from F_1 to F_3 generation, with non-significant line effects for the F_1 and F_2 generation, but highly significant line differences for the F_3 (Verhagen et al. 2019).

We like to point out here that these results were from birds housed in outdoor aviaries. For the present study, we housed the F_1 and F_2 generation birds, in their first year of age, in climate-controlled aviaries for two consecutive breeding seasons (see "Experimental setup" below). As opposed to outdoor aviaries (Verhagen et al. 2019), neither selection line, temperature environment, nor their interaction, explained females' reproductive phenotypes (i.e. laying dates and follicle widths) in these climate-controlled aviaries (Appendix 1). Those variables will thus be left out of further analyses in the present study, meaning that birds originating from both generations of selection line birds, and exposed to both temperature treatment are indiscriminately used to increase the sample size.

151

152 **Experimental setup**

153 F_1 generation (n = 36) and F_2 generation (n = 36) selection line pairs of great tits (Gienapp et al. 2019;
154 Verhagen et al. 2019) were housed in 36 climate controlled aviaries (2m × 2m × 2.25m) at the
155 Netherlands Institute of Ecology (NIOO-KNAW) in 2015 and 2016, respectively. Birds were
156 subjected to an artificial photoperiod mimicking the change in natural photoperiod. Per aviary light
157 was provided by one full spectrum daylight fluorescent lamp (58W, 5500K, True-light, The
158 Netherlands) and two fluorescent lamps (58W, Philips, The Netherlands). A roof shaft (SolaTube)
159 provided additional natural light (total average daily light intensity ~500 lux per aviary, Table S1). A
160 light bulb (7W, Philips, The Netherlands) mimicked dawn and dusk, which turned on half an hour
161 before lights went on and stayed on half an hour after lights went off respectively (Caro & Visser
162 2009). In addition, the pairs were subjected to two contrasting temperature environments mimicking
163 an extreme cold spring (2013) and an extreme warm spring (2014) in the Netherlands (Figure S1):
164 average laying dates were May 5 ± 5.18 days (n = 112) and April 11.8 ± 5.46 days (n = 124) for 2013
165 and 2014, respectively in the wild long-term study population at the Hoge Veluwe. Temperatures
166 changed every hour to follow as closely as possible the observed hourly temperatures in these years
167 (note that the minimum temperature in the aviaries was 2°C so any temperature below 2°C in the
168 temperature time series from outside were set to 2°C). The combination of selection line and
169 temperature environment resulted in four groups of nine pairs: 'early-warm', 'early-cold', 'late-warm'
170 and 'late-cold'. We like to state here again, that the variables selection line and temperature
171 environment are left out of further analyses (see above), but are mentioned here to explain the
172 experiment. Birds were fed *ad libitum* with food sources reported elsewhere (Visser et al. 2011) and
173 had water available for drinking and bathing. All F_1 and F_2 generation pairs went through two
174 experimental breeding cycles; a 'first breeding season' and a 'second breeding season' (see below and
175 Figure S2). This study was performed under the approval by the Animal Experimentation Committee
176 (DEC), Amsterdam, The Netherlands, protocol NIOO 14.10 addendum 1.

177

First breeding season

Pairs of all four groups were put in the climate controlled aviaries in the beginning of January 2015 and 2016, where birds followed the natural photoperiod. We provided nesting material (moss and hair) from the second week of March onwards. Birds went through their breeding season in which reproductive behaviours (e.g. nest-building and date of the first egg i.e. laying date) were recorded. Laying dates were recorded as January dates (i.e. 1 January = 1, 1 April = 91, etc.). Birds were blood sampled bi-weekly as part of another study (Mäkinen et al. 2019). Females could choose between three nest boxes of which two were accessible to the researcher from the outside to minimalize disturbance of the birds.

Second breeding season

After this first breeding season, when birds were photorefractory and well on their way into moult (~mid-July), days were shortened to 9L:15D and temperatures decreased to 10°C for seven weeks to make the birds photosensitive and temperature sensitive again (Dawson 2015). From September onwards, birds were again subjected to the same contrasting environments as in spring, to bring the birds into a second breeding season within the same calendar year. Because of this, and two subsequent years (2015 and 2016) with two breeding seasons to fit in one year, the second breeding season (of both 2015 and 2016) started with the photoperiod and temperatures corresponding with February 1 instead of January 1. As such, one month of photoperiodic and temperature input is missing, but it is likely that the most important period for temperatures to affect timing of breeding is from March onwards (Visser, Holleman & Gienapp 2006). SolaTubes that bring natural light from outside to the aviaries (see above) were closed, because of the mismatching photoperiods.

Females that did not initiate egg laying (n = 4 in 2015, n = 4 in 2016) in the first breeding season were replaced with their sisters for the second breeding season. However, the latter were not further used in this study (see “Statistical analysis – explaining variation in mRNA expression”). Pairs were divided

in three groups and sacrificed at three time points (see "Tissue collection and preparation") for tissue collection.

Tissue collection and preparation

For both generations, pairs were categorized in three groups (n = 12 pairs per group) based on their laying dates from the first breeding season, resulting in groups with a roughly similar average laying date and distribution (Figure S3). Three time points were chosen based on the laying dates of 2015 (F_1 generation); (1) October 7 (which corresponds to March 7 of the first breeding season) when gonadal maturation is initiated, i.e. photoperiod exceeded 11hrs (Silverin, Massa & Stokkhan 1993), (2) October 28 (i.e. March 30) when nest building occurred in the first breeding season, but prior to laying and (3) November 18 (i.e. April 20) when about 25% of the females had initiated egg laying in the first breeding season. The same time points were used in 2016 (F_2 generation) to be able to compare the experiments of 2015 and 2016, and increase sample size. Per time point one group was sacrificed (both males and females, but we focus on the females in this study). Pairs of birds were caught from the aviaries, deeply anaesthetized with Isoflurane (IsoFlo, Zoetis, Kalamazoo, Michigan) and a blood sample of 300 µl was taken from the jugular vein for possible future use. Brain, ovary and liver were dissected out. Brains were flash-frozen on dry ice and stored in 5ml RNA-free tubes at - 80°C (Qiagen, The Netherlands), whereas the other dissected tissues were placed in Eppendorf tubes and temporarily stored in liquid nitrogen. The width of the largest follicle was taken to an accuracy of 0.1 mm before freezing. All tissues were stored at - 80°C until further processing. From the frozen brains sagital cryo-sections (40 µm) were cut (Leica CM3050 S). The hypothalamus and hippocampus were located by use of online zebra finch brain atlases (Karten et al. 2013), such as ZEBrA (Oregon Health & Science University, Portland, OR 97239; http://ww.zebrafinchatlas.org) and directly isolated from the frozen brain sections using surgical punches (Harris Uni-Core, 2.0 mm). Isolated tissue was collected into 1ml TRIzol (Invitrogen, Thermo Fisher Scientific) immediately, homogenized by vigorous vortexing, and stored at -80°C until RNA isolation.

Real-Time quantitative polymerase chain reaction (RT qPCR)

Isolation of total RNA and cDNA synthesis

For RNA extraction from the hypothalamus, samples were defrosted and 0.2ml chloroform add to the 1 ml TRIzol. From the liver and ovary samples, a small piece was taken, and RNA extracted using 1 ml of TRIzol. Note that for the ovary samples we avoided using the largest follicles in order to compare between time points. RNA yield was measured on a Nanodrop 2000 (ThermoFisher Scientific, The Netherlands) and used to adjust the concentration for cDNA synthesis.

For cDNA synthesis from the isolated RNA samples we used the QuantiTect Reverse Transcription Kit (Qiagen, The Netherlands). A fixed amount of total RNA (150 ng in 6 µl RNase-free water, for hypothalamus 50 ng RNA) was incubated in gDNA Wipeout Buffer (1 µl) for the removal of genomic DNA. cDNA was generated (final volume 10 µl) following the manufacturer's instructions (Quantitect-Qiagen). A dilution of 1:5 for hypothalamus and 1:20 for liver and ovary was used for RT-qPCR analysis. Until analysis, all cDNA samples were stored at -20°C.

Primer design

We made a list of genes (1) known to be important or potentially important mediators of reproductive biology from the literature and (2) based on RNAseq data from the same F_2 generation females used in this study (Laine et al., *subm.*) (Table S2). In addition, we made a list of reference genes to allow for normalization of the gene expression levels (see "Reference genes and normalization of candidate gene expression" below, Table S2). Primers were then built based on the great tit reference genome build 1.1 (https://www.ncbi.nlm.nih.gov/assembly/GCF_001522545.2) (Laine et al. 2016) and annotation release 101 https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Parus_major/101/) with Geneious version 10.0.2 (Kearse et al. 2012) and tested (see "Real-Time quantitative PCR,

amplification efficiency" below). Primers were checked against the great tit reference genome using a BLAST search to confirm that primers were specific for the intended target genes.

Real-Time quantitative PCR, amplification efficiency

Amplification efficiency of each primer pair was determined through RT- qPCR by a 5-point standard curve based on a 5 dilution series (1:10, 1:20, 1:40, 1:80 and 1:160) of cDNA samples. Most assays for the candidate genes studied showed an efficiency (E) within the desired optimal range of 90 – 110%. Some fell outside this range, but were nevertheless included in the analysis based on a linear relation between the inverse^{10}log dilution value and the cycle threshold (C_t) ($R^2 > 0.90$) and a melt curve showing a single amplicon being formed. Selected primer pairs for the final candidate gene list are listed in Table S3. Relative transcript levels were measured by real-time quantitative PCR using the SYBR Green method; PowerUp SYBR Green Master Mix (Thermo Fisher Scientific). Fluorescence was measured with the CFX Connect Real-Time PCR Detection System (Bio-Rad Laboratories, The Netherlands) and fluorescent data analysed with the CFX Manager Software (Bio-Rad Laboratories, The Netherlands) from which C_t were obtained for subsequent analyses. Amplifications were always run in duplicate (in a different analysis and a different random sample order).

Reference genes and normalization of candidate gene expression

Although cDNA was generated from identical amounts of RNA, variations between samples may arise due to different RT efficiencies and RNA quality. Such variations were corrected for by normalizing the expression level of the target gene to a normalization factor (NF) based on the expression level of a set of reference genes determined for each cDNA sample (Vandesompele et al. 2002). We started out by selecting three candidate reference genes per tissue. Reference gene expression stability was calculated using the application geNorm (Vandesompele et al. 2002) based on which was decided whether or not to add additional candidate reference genes for accurate normalization of the mRNA

expression levels (Appendix 2). This resulted in the selection of the following reference genes: protein kinase C alpha (*PRKCA*), ribosomal protein L19 (*RPL19*) and succinate dehydrogenase complex flavoprotein subunit A (*SDHA*) for hypothalamus, beta-2-microglobulin (*B2M*), *PRKCA*, *RPL19* and *SDHA* for liver and hypoxanthine phosphoribosyltransferase 1 (*HPRT*), *PRKCA*, ribosomal protein L13 (*RPL13*), *RPL19* and tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation (*YWHAZ*) for ovary. Absolute amounts of cDNA were calculated by converting the C_t values ($C \times E^{-Ct}$, with $C = 10^{10}$ and $E = 2$) (Dijk, Kraal-Muller & Kamphuis 2004). The absolute amounts of the candidate genes were normalized against the normalization factor (NF) calculated by taking the geometric mean from the absolute amounts of the reference genes, resulting in relative mRNA expression levels of the candidate genes (arbitrary amounts).

Statistical analysis

Correlating phenotypes from the first and second breeding season. We used the laying dates in breeding season one as a measure for whether females are early or late breeders in season two. Follicle widths were log10 transformed before performing simple linear regression to investigate the relationship between laying date in the first breeding season and follicle width of the largest follicle in the second breeding season. This relationship was subsequently tested per time point.

Explaining variation in mRNA expression. Removing females from the data due to death or not initiating egg laying in the first breeding season or having unreliable mRNA level measurements, resulted in n = 59, n = 58 and n = 59 individual females for hypothalamus, ovary and liver, respectively. Individual mRNA expression data were subjected to both principal component analysis (PCA) and univariate statistical analyses, which were all performed in R (version 3.3.1). Prior to subjecting the data to PCA, we log10 transformed the individual gene expression data. Using the function *prcomp*, PCA was performed, which consolidates the individual mRNA expression level data into new variables known as principal components (PCs) and so reducing the number of dimensions of

the data. These PCs allowed for simultaneous assessment of expression values of the genes measured for hypothalamus, liver and ovary and give an indication of the variables that best explain variation in gene expression levels per tissue. Horn's analysis was performed to determine which PCs to retain (i.e. eigenvalue > 1). Assessment of the association between PCs and explanatory variables (i.e. time point, laying date and generation) was determined by performing ANOVAs, with the following model $PC_x \sim$ time point × laying date + generation. P-values were adjusted for multiple comparisons using Benjamini and Hochberg's False Discovery Rate (FDR) (Benjamini and Hochberg, 1995), accepting an FDR of 0.05. Females did not differ in either laying dates or largest follicle widths between selection lines, temperature treatments or their interaction within a generation (see above, Appendix 1). To exclude these variables from the study, we performed an initial analysis to test whether these variables would influence individual gene expression levels. As opposed to the study of Laine et al (*subm.*), in which genome-wide gene expression patterns were tested compared to our limited number of candidate genes, selection line, temperature environment or their interaction did not influence gene expression levels (Table S4) and were therefore left out for further analyses. Subsequently, the same procedure as applied to the PCs was used to analyze the expression level of an individual candidate gene, with $expression_{gene} \sim$ time point × laying date (from the first breeding season) + generation.

Pairwise correlations between gene pairs. Pearson's correlation coefficients between every gene pair possible were calculated and visualized with the *rcorr* and *corrplot* functions in R, respectively, in order to determine which gene pairs tend to change significantly (accepting a $p < 0.05$) together within and across the tissues examined.

Results

Relationship between laying dates and follicle widths

There is a weak but significant negative linear relationship (r = -0.32; $F_{1,59}$ = 6.88, p = 0.01) between laying date and largest follicle (Figure S4A). When analyzing per time point (Figure S4B), the relationship between laying date and largest follicle went from no relationship at time point 1 (r = -0.14, $F_{1,17}$ = 0.33, p = 0.57), to a moderate negative relationship at time point 2 (r = -0.57, $F_{1,18}$ = 8.81, p = 0.01) and a strong negative relationship at time point 3 (r = -0.66, $F_{1,20}$ = 15.18, p < 0.001). Also given the significant difference in follicle widths (Appendix 1), we are confident that the mRNA expression levels from the second breeding season are representative of the laying dates recorded (Appendix 3).

Gene expression assessment through Principal Component Analysis (PCA)

PC1 and PC2, the dimensions with eigenvalues >1 according to Horn's analysis, explain together 86.8%, 48.1% and 73.7% of the variance in gene expression among females in hypothalamus (n = 59), ovary (n = 58) and liver (n = 59) respectively (Table S5-S7).

Hypothalamus

Based on the loadings, mRNA expression of iodothyronine deiodinase type 2 (*DIO2*), opsin 5 (*OPN5*), thyrotropin releasing hormone (*TRH*) and nuclear factor interleukin-3-regulated protein (*NFIL3*) are accounting for the variance in PC1, whereas mRNA expression of vasoactive intestinal peptide (*VIP*) in the hypothalamus explains a large part of the variance in PC2 (Table S5). In addition, the similar loadings (Table S5) and the small angle between the vectors of *OPN5* and *DIO2* (Figure 1A) suggest a correlation between these genes. Females showed different candidate gene expression profiles between generations ($F_{2,\ 54}$ = 143, FDR corrected p < 0.0001, Table S8), as shown by two distinct, but overlapping clusters along PC1 in hypothalamus (Figure 1A). No distinction in expression profile was found when clustering females per 'time point' for PC1 or PC2 (Figure 1B), nor did time point explain variance in any of the PCs (Table S8) and also no association between expression of these genes and the interaction between laying date and time point was found.

Ovary

In ovary, the variance in PC1 is mainly explained by mRNA expression of the androgen receptor (*AR*), luteinizing hormone receptor (*LHR*), matrix metallopeptidase 15 (*MMP15*) and interferon related developmental regulator 1 (*IFRD1*) (Figure 1C, Table S6). Whereas the variance in PC2 is mainly explained by mRNA expression of heat shock protein family B member 1 (*HSPB1*), cytochrome P450 17A1 (*CYP17A1*) and very low-density lipoprotein receptor (*VLDLR*) (Figure 1C, Table S6). Although not shown in Figure 1, but based on similar loadings in PC1 and 2 (Table S6), expression levels of *CYP17A1*, *ER*, and *VLDLR* are correlated. Females show distinct differences in candidate gene expression profile between generations ($F_{1, 54}$ = 269.57, FDR corrected $p < 0.0001$) along PC1 (Figure 1C, Table S9) and a gradual change in expression profiles when clustering for time point ($F_{2, 55}$ = 22.01, FDR corrected $p < 0.0001$) along PC2 (Figure 1D, Table S9). PC1 and PC2, together accounting for ~48% of the total variance, are highly significantly associated with both generation and time point (Table S9).

Liver

PC1, accounting for ~46% of the variance among females, is associated with a laying date × time point interaction ($F_{2, 53}$ = 11.019, FDR corrected $p < 0.001$, Table S10) and is mainly explained by mRNA expression of apovitellenin 1 (*APOV1;* LOC107200088), bestrophin 3 (*BEST3*), CathepsinE-A-like protein (*CTSEAL;* LOC10720510) and vitellogenin 2 (*VTG2*) (Table S7). Generation explains the variation in gene expression (~28%) along PC2 ($F_{2, 53}$ = 38.09, FDR corrected $p < 0.0001$, Table S10). Although not shown in Figure 1, but based on similar loadings in PC1 and 2 (Table S7), *BEST3*, *CTSEAL* and *VTG2* are correlated in terms of expression among these females, as are *MR* and *HSPB1*. As in hypothalamus, but along PC2 instead of PC1, females show overlapping but different candidate gene expression profiles between generations in liver (Figure 1E). Similar to ovary, but again along opposite PCs, females show a gradual change in expression profile over time points (Figure 1F).

Variation in hypothalamic, ovarian and liver candidate gene expression

Hypothalamus

We found no differences in candidate gene expression in hypothalamus between time points, laying dates or their interaction (Table 1). The F_1-generation females had significantly higher expression levels in each time point for *DIO2*, *NFIL3*, *OPN5* and *TRH* compared to F_2-generation females (*DIO2*: $F_{1,57} = 82.52$, FDR corrected $p < 0.0001$; *NFIL3*: $F_{1,57} = 58.03$, FDR corrected $p < 0.0001$; *OPN5*: $F_{1,57} = 77.15$, FDR corrected $p < 0.0001$; *TRH*: $F_{1,57} = 160.51$, FDR corrected $p < 0.0001$, Figure 2, Table 1). We found no difference in expression levels of *VIP* between generations (data not shown).

Ovary

With the exception of *FSHR*, gonadotropin-inhibitory hormone receptor (*GnIHR*), prolaction receptor (*PRLR*) and steroidogenic acute regulatory protein (*StAR*), all candidate genes showed significant differences between generations and time points in ovary (Table 2, Figure S5), but only variation in mRNA expression of *IFRD1* (Figure 3A) and *VLDLR* (Figure 3B) is explained by timing of breeding (*IFRD1*: $F_{1,55} = 6.86$, FDR corrected $p = 0.03$, *VLDLR*: $F_{1,55} = 13.25$, FDR corrected $p < 0.001$, Table 2).

Liver

Early breeding females show increased mRNA expression for both insulin-like growth factor 1 (*IGF1*, Figure 4A) and *VTG2* (Figure 4B) in liver (*IGF1*: $F_{1,55} = 6.53$, FDR corrected $p = 0.03$, *VTG2*: $F_{1,58} = 6.62$, FDR corrected $p = 0.03$, Table 3) compared to late breeding females. Only in liver, we found differences in mRNA expression levels explained by laying date × time point interactions (Figure 5, Table 3). Females showed an increase in gene expression over time points for *APOB* ($F_{1,53} = 5.10$, FDR corrected $p = 0.03$), *APOV1* ($F_{2,53} = 11.58$, FDR corrected $p < 0.0001$), *BEST3* ($F_{2,53} = 6.53$, FDR corrected $p = 0.010$) and *CTSEAL* ($F_{2,53} = 7.21$, FDR corrected $p = 0.01$), with higher expression for

411 early laying females compared to late laying in time points 2 and 3. The genes *GR*, *HSPB1* and *MR*
412 only showed a generation effect (Table 3, Figure S6).

413

414 **Pairwise correlations between gene pairs**

415 Within and among the tissues examined, candidate genes, whether they reflect differences in timing or
416 not, tend to change in a strong and/or significantly similar way (Figure 6). For example, *CYP17A1*
417 expression in the ovary tends to change in a strong and similar way as *APOV1*, *CTSEAL* and *VTG2* in
418 liver. In addition, expression of *HSPB1* in ovary resembles that of *APOB* and *APOV1* in liver. The
419 mRNA expression of *GNIHR* in ovary shows a weak positive, however significant, correlation with
420 *VTG2* in liver. Interestingly, the genes examined in the hypothalamus show a high and significant
421 correlation among each other, but less so when correlated to genes in the ovary and liver. Between the
422 ovary and liver, more genes tend to change in a similar way, both positively and negatively.

423

424 # Discussion

425 Gene expression dynamics within the HPGL axis have not been well studied in seasonally breeding
426 females. Using a candidate gene approach, we set out to determine whether individual differences in
427 egg-laying dates (obtained from the first breeding season) are reflected in differences in candidate
428 gene expression levels, and if so, where these differences occur in the HPGL axis; upstream and/or
429 downstream and when these differences can be picked up towards the expected laying dates. We found
430 significant differences in mRNA expression of candidate genes between generations in all three tissues
431 examined. However, a correlation of candidate gene expression and egg-laying date (at the three
432 sampling time points) was found exclusively in ovary and liver, independent of generation. In
433 particular, individual differences in timing of breeding in females are significantly reflected in mRNA
434 expression for *IFRD1* and *VLDLR* in ovary and *IGF-1* in liver, and earlier breeding females show
435 increased expression of *APOB*, *APOV1*, *BEST3* and *CTSEAL* over time in liver. These findings,
436 together with other patterns found, suggest that fine-tuning of avian timing of breeding is regulated

downstream in the HPGL axis. This is in concurrence with the "alternative, female-specific hypothesis" (Caro et al 2009, Williams 2012), which awards a more prominent role for the ovary and/or liver in fine tuning timing of breeding (see "Downstream regulation of timing of breeding" below).

Experimental limitations

We compared gene expression levels at different time points approaching laying the first egg, but with different individuals per time point. The limitation is here, that an individual female could not be measured at each time point. There could be individual differences in responses to cues and (reproductive) physiology, which potentially decreased our power to detect patterns over time. In addition, a three-week interval between time points is quite long and with a wide range in laying dates, properly determining the last time point, where most females are supposed to have initiated vitellogenesis or egg-laying, posed a challenge. Further, due to practical reasons indicated in the materials and methods, we had to leave out the January photoperiods and temperatures for the second breeding season. However, because of increased expression levels for genes involved in for example vitellogenesis in both this study and the genome-wide study (Laine et al. *subm.*), and that several females had entered RYD or initiated laying (Appendix 2), we are positive that, given the narrow time window in which this occurs, the third time point was estimated correctly. Further, we avoided using the largest follicles, which prevented inflated expression levels for (certain) candidate genes and a possible misinterpretation of the results. We used two generations of selection lines in this study, which generated significant differences in gene expression levels in the three organs examined. Likely timing of the experiments and processing of the samples, for example, might be causing these differences. An alternative explanation is that year differences were causal to these differences, but we do not have enough years to test this.

Hypothalamus

Interestingly, temperature treatment affected genome-wide gene expression profiles early in the breeding season (time point 1) in the hypothalamus, but not in the ovary and liver in the same samples as used here of the F_2- generation females (Laine et al. *subm.*). In addition, we did *not* find an effect of temperature treatment on gene expression levels nor on the onset of egg-laying or follicular growth. The latter is contrary to previous studies in great tits housed in climate controlled aviaries, showing that the pattern of increase in ambient temperature has a direct effect on the onset of egg-laying (Schaper, Dawson, Sharp, et al. 2012; Visser, Holleman & Caro 2009), but agrees with other studies where gonadal size is not affected by ambient temperature (Schaper, Dawson, Sharp, et al. 2012; Visser et al. 2011). It seems that in these females in the beginning of the breeding season, the brain is able to perceive ambient temperatures, to 'switch on' the reproductive axis at an upstream level (perhaps in a similar way to photoperiod). However, even though temperature could possibly affect other tissues, it does not seem to directly affect gene expression in the ovary and liver to fine tune egg laying.

The F_2-generation females showed significantly lower expression levels of *DIO2*, *NFIL3*, *OPN5*, and *TRH* in all three time points compared to the F1 females, where *VIP* did not. These genes are involved in circadian rhythms (*DIO2*, *NFIL3*) (Cowell 2002; Yoshimura et al. 2003), photoperiodic perception (*OPN5*) (Nakane et al. 2014) and regulation of the hypothalamic-pituitary-thyroid axis (*TRH*) (McNabb 2007). A possible explanation for this generation difference could be that F_2-females were, on average, ~7.5 day later in onset of egg laying. However, both the F_1 and F_2 generation females followed the same photoperiod. Also, generation differences were found in ovary and liver, but again not for all genes. We are hesitant to attribute these generation differences to different biological functioning (see 'Limitations').

Ovary

The expression of *IFRD1*, a gene proposed to be involved in regulation of cell proliferation and

differentiation (Vadivelu et al. 2004; Vietor & Huber 2007), decreased in time point 3 compared to time point 1 for all females, as in Laine et al. (*subm*)., but significantly for the early laying females (Figure 3) . This is in contrast with a study in female Sprague-Dawley rats, where increased expression of *IFRD1* was found in granulosa cells and cumulus oocyte complexes after administration of human chorionic gonadotropin (to mimic the LH-surge and induce ovulation), indicating potential involvement of *IFRD1* in oocyte maturation (Li et al. 2016). However, this study was performed in a different time frame (hours) and on single cells compared to weeks and ovary homogenates, respectively in our study.

The mRNA expression of *VLDLR*, increased from time point 1 (early March) to time point 2 (late March) and decreased again in time point 3 (mid-April) in F_1-females. When taking into account that females in climate controlled aviaries lay ~3 weeks later (Visser, Holleman & Caro 2009) compared to wild females, this finding is consistent with expression in ovaries of European starlings (Perfito et al. 2015). However, we expected *VLDLR* expression to be lowest in non-breeding females (i.e. time point 1) (George, Barber & Schneider 1987) and highest in pre-laying females (i.e. time point 3) (Han, Haunerland & Williams 2009).

We find clear variation in expression, though not significantly explaining variation in laying dates, between early and late laying females over time for genes regulating, among others, processes involved in steroidogenesis (*CYP17A1*, *LHR*) (Johnson 2015) and follicular development through gonadotropin binding (*GnIHR*) (Maddineni et al. 2008). In the ovary both *CYP17A1* and *LHR* expression is higher for females laying early and peaks in time point 3. This increase over time, and nearing egg-laying, is consistent with findings in ovary homogenates in European starlings (Perfito et al. 2015) and in follicles of the dark-eyed junco (Needham et al. 2019). These studies support the idea that LHR seems to play a key role in the 'competence' of the ovary to respond to circulating gonadotropins (see "Downstream regulation of timing of breeding" below). Further, in chicken

(*Gallus gallus*), *CYP17A1* and *LHR* show increased expression when follicle selection takes place and is initiated by the signalling of several receptors via cyclic AMP (Johnson 2015).

Liver

Earlier breeding females showed increased mRNA expression levels over time in liver for genes involved in vitellogenesis and oocyte growth, which is consistent with differential expression levels found for these genes time point 1 and 3 (Laine et al. 2019). *APOV1* (alias apoVLDL-II) is a protein component of yolk-targeted very-low density lipoprotein (VLDLy), a lipoprotein synthesized by the liver under the influence of E2 and, together with VTG, the primary source of yolk protein and lipid for the developing embryo (Walzem 1996). APOB, a protein associated with VTG and VLDLy (Walzem 1996), and VTG2 (one of the three forms of VTG and the most abundant), show increased expression over time compared to *APOV1*, *BEST3* and *CTSEAL*. These expression patterns agree with concentrations of VTG and VLDL found in other seasonal breeders (Caro et al. 2009; Challenger et al. 2001). Like VTG and VLDL, synthesis of *CTSEAL* by the liver is estrogen-dependent (Zheng et al. 2018). Further, it is allegedly involved in sexual maturation of female chicken (Bourin, Gautron, Berges, Nys, et al. 2012) and may play a role in processing egg yolk macromolecules (Bourin, Gautron, Berges, Hennequet-Antier, et al. 2012), since it is found in egg yolk (Farinazzo et al. 2009). The function of *BEST3* in this study is unclear. *BEST3* is positioned closely to *CTSEAL* in the genome, and therefore its lower expression might be caused by an involvement of co-regulation with *CTSEAL* (Laine et al. *subm.*; Zheng et al. 2018).

We found expression of *IGF1* to reflect individual differences in egg laying, with early laying females showing higher *IGF1* expression compared to late laying females. There is little knowledge regarding the connection between IGF-1 and reproductive traits in birds. Few studies (mainly poultry) exist; ovaries have IGF-1 receptors and IGF-1 plays a regulatory role in ovarian functions, such as follicular growth and differentiation (Onagbesan et al. 1999) and stimulates ovarian progesterone production

(Williams 1994). Growth and reproduction are closely related and there is cross talk between the endocrine systems controlling these fundamental processes in vertebrates (Hull & Harvey 2014 and references therein). Studies in female chicken and rabbit suggests that IGF-1 is also produced by the ovary, together with and under the influence of growth hormone, where they act as paracrine/autocrine regulators during follicular development (Ahumada-Solorzano et al. 2016; Yoshimura et al. 1994, 1996). In addition, different variants of IGF-1 genes, as well as variation in IGF-1 levels in poultry resulted mainly in variation in productivity; different numbers of eggs produced or variation in egg quality (Hocking et al. 1994; Nagaraja et al. 2000; Wu et al. 2016).

Pairwise correlations between gene pairs

The limited number of candidate genes, which are not assessed in all the tissues examined hamper the construction of a gene network and a subsequent co-expression network analysis in order to associate genes (of unknown function in relation to timing of breeding) with biological processes. Even so, these preliminary results on correlated expression between gene pairs within and across tissues, highlight the importance to not only look within, but also across tissues in the HPGL axis. Further, co-expression of these genes might indicate the same transcriptional regulatory program (e.g. transcription factors, DNA methylation). In addition, these preliminary results emphasize the importance of the communication between ovary and liver as a potential mechanism in timing of breeding. For example, *CYP17A1* shows significantly correlated expression with genes expressed in liver (*CTSEAL*, *VTG2* and *APOV1*, Figure S5), that are involved in lipid metabolism and yolk formation (Walzem 1996, Zheng et al 2018). Of course, E2, for which CYP17A1 is a key enzyme in the steroidogenic pathway underlying its production, stimulates vitellogenesis (Mullinix et al. 1976). However, whether the 'decision' to lay is mechanistically linked to follicle selection and development, ovulation and ultimately egg-laying remains to be investigated.

Downstream regulation of timing of breeding

567 Currently, one can only speculate on where the 'switch' that initiates egg-laying resides within the
568 ovary and/or liver. A potential candidate is the 'competence' of the ovary to respond to gonadotropins
569 via their receptors (Johnson 2015, Caro et al. 2009; Ball 2007, Williams 2012, Schaper et al 2012,
570 Partecke et al 2005). Further, in starlings, it has also been shown that sex steroid secretion can be
571 regulated by local GnIH in the gonads (McGuire, Kangas & Bentley 2011, Kriegsfeld et al 2015). As
572 such, the gonadal GnIH system could be a potential mechanism in timing of breeding in females
573 (McGuire, Kangas & Bentley 2011; Needham et al 2018). Another potential mechanism is the
574 communication between the ovary and liver, where the E2-dependent shift in lipid metabolism or the
575 up-regulation of VTG/VLDL-receptors could be candidates. These potential mechanisms, however,
576 need to be regulated, and imply a more autonomous role for the ovary together with receiving signals
577 that bypass the classic neuro-endocrine pathway. As such, the ovary and brain might act more as
578 'partners' (Ball 2007). For example, a study in Japanese quail (*Coturnix japonica*) suggests that the
579 ovary regulates her own functioning through its circadian clock, because the largest follicle, through
580 production of circadian clock gene proteins, controls the LH surge that is essential for ovulation
581 (Nakao et al 2007).

582

583 *Outlook*

584 The exact downstream mechanisms that precede timing of breeding and how they are regulated
585 remains to be determined. Though gene expression is not the only mechanism regulating timing of
586 breeding, we have shown that variation in mRNA expression levels of several candidate genes in
587 ovary and liver, associated with reproductive functioning, explain variation in timing of breeding in
588 these females. Our study confirms that shifting the focus more towards *females* rather than males
589 (Caro 2012; Williams 2012) in future experimental studies investigating timing of breeding, is highly
590 important. Also, simultaneous examination of multiple, and preferably all, HPGL axis levels is of the
591 essence in understanding mechanisms underlying timing of breeding. This way, we gain knowledge on
592 the variation in the physiology underlying timing of avian breeding and what part of this variation is
593 genetically determined. Timing of breeding is currently under selection in wild populations due to

climate change (Both & Visser 2001; Visser et al. 1998). A better understanding of the variation in the physiological processes underlying seasonal timing will ultimately lead us to a better understanding of a species' adaptive potential to their warming world.

Acknowledgments

We thank Marylou Aaldering, Coretta Jongeling, Franca Kropman and Anouk de Plaa for taking care of the birds. We also thank Davide Dominoni and Barbara Helm for help in selecting candidate genes, Renske Jongen for assistance during tissue collection and Jeroen Laurens and Gilles Wijlhuizen for technical assistance prior and during the experiments. We are grateful to the Netherlands Institute of Neuroscience, Amsterdam, The Netherlands for making a laboratory available for cryo-sectioning the brain material.

Author contributions

MEV and IV designed the study. IV performed the experiments, with assistance of RdW, BvL, SPC and HV, analysed the data and wrote the manuscript. ACM and AP did the molecular work, to which WK contributed. KvO, VNL, WK, HV, TDW, SPC and SLM contributed to designing the study, train IV for experimental procedures, and data analysis and interpretation. All co-authors commented on the manuscript.

Competing interests

All authors declare to have no competing interests that might have influenced this manuscript.

Funding

This study was supported by an ERC Advanced Grant (339092 – E-response to MEV). SPC was supported by a grant from the French National Research Agency (ANR-15-CE02-0005-01).

References

Ahumada-Solorzano, S.M., Martinez-Moreno, C.G., Carranza, M., Avila-Mendoza, J., Luna-Acosta, J.L., Harvey, S., Luna, M. & Aramburo, C. 2016, 'Autocrine/paracrine proliferative effect of ovarian GH and IGF-I in chicken granulosa cell cultures.', *General and comparative endocrinology*, vol. 234, pp. 47–56.

Bergeon Burns, C.M., Rosvall, K.A., Hahn, T.P., Demas, G.E. & Ketterson, E.D. 2014, 'Examining sources of variation in HPG axis function among individuals and populations of the dark-eyed junco', *Hormones and Behavior*, vol. 65, no. 2, pp. 179–87.

Both, C. & Visser, M.E. 2001, 'Adjustment to climate change is constrained by arrival date in a long-distance migrant bird', *Nature*, vol. 411, p. 296.

Bourin, M., Gautron, J., Berges, M., Hennequet-Antier, C., Cabau, C., Nys, Y. & Rehault-Godbert, S. 2012, 'Transcriptomic profiling of proteases and antiproteases in the liver of sexually mature hens in relation to vitellogenesis.', *BMC genomics*, vol. 13, p. 457.

Bourin, M., Gautron, J., Berges, M., Nys, Y. & Rehault-Godbert, S. 2012, 'Sex- and tissue-specific expression of "similar to nothepsin" and cathepsin D in relation to egg yolk formation in Gallus gallus.', *Poultry science*, vol. 91, no. 9, pp. 2288–93.

Bradshaw, W.E. & Holzapfel, C.M. 2007, 'Evolution of Animal Photoperiodism', *Annual Review of Ecology, Evolution, and Systematics*, vol. 38, no. 1, pp. 1–25.

Cánovas, A., Reverter, A., DeAtley, K.L., Ashley, R.L., Colgrave, M.L., Fortes, M.R.S., Islas-Trejo, A., Lehnert, S., Porto-Neto, L., Rincón, G., Silver, G.A., Snelling, W.M., Medrano, J.F. & Thomas, M.G. 2014, 'Multi-Tissue Omics Analyses Reveal Molecular Regulatory Networks for Puberty in Composite Beef Cattle', *PLOS ONE*, vol. 9, no. 7, p. e102551.

Caro, S.P. 2012, 'Avian ecologists and physiologists have different sexual preferences.', *General and comparative endocrinology*, vol. 176, no. 1, pp. 1–8.

Caro, S.P., Charmantier, A., Lambrechts, M.M., Blondel, J., Balthazart, J. & Williams, T.D. 2009, 'Local adaptation of timing of reproduction: Females are in the driver's seat', *Functional Ecology*, vol. 23, no. 1, pp. 172–9.

Caro, S.P. & Visser, M.E. 2009, 'Temperature-induced elevation of basal metabolic rate does not affect testis growth in great tits.', *The Journal of experimental biology*, vol. 212, no. Pt 13, pp. 1995–9.

Challenger, W.O., Williams, T.D., Christians, J.K. & Vézina, F. 2001, 'Follicular development and plasma yolk precursor dynamics through the laying cycle in the European starling (Sturnus vulgaris)', *Physiological and Biochemical Zoology*, vol. 74, no. 3, pp. 356–65.

Cheviron, Z.A., Whitehead, A. & Brumfield, R.T. 2008, 'Transcriptomic variation and plasticity in rufous-collared sparrows (Zonotrichia capensis) along an altitudinal gradient', *Molecular Ecology*, vol. 17, no. 20, pp. 4556–69.

Cowell, I.G. 2002, 'E4BP4/NFIL3, a PAR-related bZIP factor with many roles.', *BioEssays : news and reviews in molecular, cellular and developmental biology*, vol. 24, no. 11, pp. 1023–9.

Dawson, A. 2008, 'Control of the annual cycle in birds: endocrine constraints and plasticity in response to ecological variability.', *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, vol. 363, no. 1497, pp. 1621–33.

Dawson, A. 2015, 'Annual gonadal cycles in birds: Modeling the effects of photoperiod on seasonal changes in GnRH-1 secretion', *Frontiers in Neuroendocrinology*, vol. 37, pp. 52–64.

Derks, M.F.L., Schachtschneider, K.M., Madsen, O., Schijlen, E., Verhoeven, K.J.F. & van Oers, K.

2016, ‘Gene and transposable element methylation in great tit (Parus major) brain and blood’, *BMC Genomics*, vol. 17, no. 1, p. 332.

Dijk, F., Kraal-Muller, E. & Kamphuis, W. 2004, ‘Ischemia-induced changes of AMPA-type glutamate receptor subunit expression pattern in the rat retina: a real-time quantitative PCR study’, *Investigative Ophthalmology & Visual Science*, vol. 45, no. 1, pp. 330–41.

Farinazzo, A., Restuccia, U., Bachi, A., Guerrier, L., Fortis, F., Boschetti, E., Fasoli, E., Citterio, A. & Righetti, P.G. 2009, ‘Chicken egg yolk cytoplasmic proteome, mined via combinatorial peptide ligand libraries.’, *Journal of chromatography. A*, vol. 1216, no. 8, pp. 1241–52.

Fidler, A.E., van Oers, K., Drent, P.J., Kuhn, S., Mueller, J.C. & Kempenaers, B. 2007, ‘Drd4 gene polymorphisms are associated with personality variation in a passerine bird’, *Proceedings of the Royal Society B: Biological Sciences*, vol. 274, no. 1619, pp. 1685–91.

George, R., Barber, D.L. & Schneider, W.J. 1987, ‘Characterization of the chicken oocyte receptor for low and very low density lipoproteins.’, *The Journal of biological chemistry*, vol. 262, no. 35, pp. 16838–47.

Gienapp, P.G., Calus, M.P.L., Laine, V.N. & Visser, M.E. 2019, ‘Genomic selection on breeding time in a wild bird population’, *Evolution Letters*, pp. 1–10.

Han, D., Haunerland, N.H. & Williams, T.D. 2009, ‘Variation in yolk precursor receptor mRNA expression is a key determinant of reproductive phenotype in the zebra finch (Taeniopygia guttata)’, *Journal of Experimental Biology*, vol. 212, no. 9, pp. 1277 LP – 1283.

Hocking, P.M., Bernard, R., Wilkie, R.S. & Goddard, C. 1994, ‘Plasma growth hormone and insulin- like growth factor- I (IGF- I) concentrations at the onset of lay in ad libitum and restricted broiler breeder fowl’, *British Poultry Science*, vol. 35, no. 2, pp. 299–308.

Hull, K.L. & Harvey, S. 2014, ‘Growth hormone and reproduction: a review of endocrine and autocrine/paracrine interactions.’, *International journal of endocrinology*, vol. 2014, p. 234014.

Johnson, A.L. 2015, ‘Ovarian follicle selection and granulosa cell differentiation.’, *Poultry Science*, vol. 94, no. 4, pp. 781–5.

Karten, H.J., Brzozowska-Prechtl, A., Lovell, P. V., Tang, D.D., Mello, C. V., Wang, H. & Mitra, P.P. 2013, ‘Digital atlas of the zebra finch (Taeniopygia guttata) brain: a high-resolution photo atlas’, *The Journal of comparative neurology*, vol. 521, no. 16, pp. 3702–15.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. & Drummond, A. 2012, ‘Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data’, *Bioinformatics*, vol. 28, no. 12, 2012/04/27., pp. 1647–9.

Kim, J.M., Santure, A.W., Barton, H.J., Quinn, J.L., Cole, E.F., Visser, M.E., Sheldon, B.C., Groenen, M.A.M., van Oers, K. & Slate, J. 2018, ‘A high density SNP chip for genotyping great tit (Parus major) populations and its application to studying the genetic architecture of exploration behaviour’, *Molecular Ecology Resources*, no. December 2017, pp. 1–15.

Laine, V.N., Gossmann, T.I., Schachtschneider, K.M., Garroway, C.J., Madsen, O., Verhoeven, K.J.F.F., de Jager, V., Megens, H., Warren, W.C., Minx, P., Crooijmans, R.P.M.A., Corcoran, P., Adriaensen, F., Belda, E., Bushuev, A., Cichon, M., Charmantier, A., Dingemanse, N., Doligez, B., Eeva, T., Erikstad, K., Fedorov, S., Hau, M., Hille, S., Hinde, C., Kempenaers, B., Kerimov, A., Krist, M., Mand, R., Matthysen, E., Nager, R., Norte, C., Orell, M., Richner, H., Slagsvold, T., Tilgar, V., Tinbergen, J., Torok, J., Tschirren, B., Yuta, T., Sheldon, B.C., Slate, J., Zeng, K., van Oers, K., Visser, M.E. & Groenen, M.A.M. 2016, ‘Evolutionary signals of selection on cognition from the great tit genome and methylome’, *Nature Communications*, vol. 7, p. 10474.

Laine, V.N., Verhagen, I., Mateman, A.C., Pijl, A., Gienapp, P., van Oers, K. & Visser, M.E. 2019,

709 'Gene expression patterns underlying timing of reproduction in a song bird', *Manuscript*
710 *submitted for publication*.

711 Lambrechts, M.M. & Visser, M.E. 1999, 'No', N.J. Adams & R.H. Slotow (eds),*Proceedings of the*
712 *22nd International Ornithol. Congress, Durban*, BirdLife South Africa, Johannesburg, pp. 231–
713 3.

714 Li, F., Liu, Y., Miao, X., Fu, G. & Curry Jr., T.E. 2016, 'Expression and regulation of the
715 differentiation regulators ERBB Receptor Feedback Inhibitor 1 (ERRFI1) and Interferon-related
716 Developmental Regulator 1 (IFRD1) during the periovulatory period in the rat ovary.',
717 *Molecular reproduction and development*, vol. 83, no. 8, pp. 714–23.

718 MacManes, M.D., Austin, S.H., Lang, A.S., Booth, A., Farrar, V. & Calisi, R.M. 2017, 'Widespread
719 patterns of sexually dimorphic gene expression in an avian hypothalamic–pituitary–gonadal
720 (HPG) axis', *Scientific Reports*, vol. 7, p. 45125.

721 Maddineni, S.R., Ocon-Grove, O.M., Krzysik-Walker, S.M., Hendricks, G.L. 3rd. & Ramachandran,
722 R. 2008, 'Gonadotropin-inhibitory hormone (GnIH) receptor gene is expressed in the chicken
723 ovary: potential role of GnIH in follicular maturation.', *Reproduction (Cambridge, England)*,
724 vol. 135, no. 2, pp. 267–74.

725 Mäkinen, H., Viitaniemi, H.M., Visser, M.E., Verhagen, I., van Oers, K. & Husby, A. 2019,
726 'Temporally replicated reduced representation bisulfite sequencing data on DNA methylation
727 patterns in great tit', *Accepted for publication in Scientific Data*.

728 Maruska, K.P. & Fernald, R.D. 2011, 'Plasticity of the Reproductive Axis Caused by Social Status
729 Change in an African Cichlid Fish: II. Testicular Gene Expression and Spermatogenesis',
730 *Endocrinology*, vol. 152, no. 1, pp. 291–302.

731 Maruska, K.P., Levavi-Sivan, B., Biran, J. & Fernald, R.D. 2011, 'Plasticity of the Reproductive Axis
732 Caused by Social Status Change in an African Cichlid Fish: I. Pituitary Gonadotropins',
733 *Endocrinology*, vol. 152, no. 1, pp. 281–90.

734 McGuire, N.L., Kangas, K. & Bentley, G.E. 2011, 'Effects of melatonin on peripheral reproductive
735 function: regulation of testicular GnIH and testosterone', *Endocrinology*, vol. 152, no. 9, pp.
736 3461–70.

737 McNabb, F.M.A. 2007, 'The Hypothalamic-Pituitary-Thyroid (HPT) Axis in Birds and Its Role in
738 Bird Development and Reproduction', *Critical Reviews in Toxicology*, vol. 37, no. 1–2, pp. 163–
739 93.

740 Mullinix, K.P., Wetekam, W., Deeley, R.G., Gordon, J.I., Meyers, M., Kent, K.A. & Goldberger, R.F.
741 1976, 'Induction of vitellogenin synthesis by estrogen in avian liver: relationship between level
742 of vitellogenin mRNA and vitellogenin synthesis', *Proceedings of the National Academy of*
743 *Sciences*, vol. 73, no. 5, pp. 1442 LP – 1446.

744 Nagaraja, S.C., Aggrey, S.E., Yao, J., Zadworny, D., Fairfull, R.W. & Kuhnlein, U. 2000, 'Trait
745 association of a genetic marker near the IGF-I gene in egg-laying chickens', *The Journal of*
746 *Heredity*, vol. 91, no. 2, pp. 150–6.

747 Nakane, Y., Shimmura, T., Abe, H. & Yoshimura, T. 2014, 'Intrinsic photosensitivity of a deep brain
748 photoreceptor', *Current Biology*, vol. 24, no. 13, pp. R596–7.

749 Nakane, Y. & Yoshimura, T. 2014, 'Universality and diversity in the signal transduction pathway that
750 regulates seasonal reproduction in vertebrates', *Frontiers in Neuroscience*, vol. 8, no. 8 MAY,
751 pp. 1–7.

752 Needham, K.B., Bergeon Burns, C.M., Graham, J.L., Bauer, C.M., Kittilson, J.D., Ketterson, E.D.,
753 Hahn, T.P. & Greives, T.J. 2019, 'Changes in processes downstream of the hypothalamus are
754 associated with seasonal follicle development in a songbird, the dark-eyed junco (Junco

hyemalis)', *General and Comparative Endocrinology*, vol. 270, pp. 103–12.

Onagbesan, O.M., Vleugels, B., Buys, N., Bruggeman, V., Safi, M. & Decuypere, E. 1999, 'Insulin-like growth factors in the regulation of avian ovarian functions.', *Domestic animal endocrinology*, vol. 17, no. 2–3, pp. 299–313.

Partecke, J., Van't Hof, T.J. & Gwinner, E. 2005, 'Underlying physiological control of reproduction in urban and forest-dwelling European blackbirds Turdus merula', *Journal of Avian Biology*, vol. 36, no. 4, pp. 295–305.

Perfito, N., Guardado, D., Williams, T.D. & Bentley, G.E. 2015, 'Social cues regulate reciprocal switching of hypothalamic Dio2/Dio3 and the transition into final follicle maturation in European starlings (Sturnus vulgaris)', *Endocrinology*, vol. 156, no. 2, pp. 694–706.

Schaper, S. V., Dawson, A., Sharp, P.J., Caro, S.P. & Visser, M.E. 2012, 'Individual variation in avian reproductive physiology does not reliably predict variation in laying date.', *General and comparative endocrinology*, vol. 179, no. 1, pp. 53–62.

Schaper, S. V., Dawson, A., Sharp, P.J., Gienapp, P., Caro, S.P. & Visser, M.E. 2012, 'Increasing Temperature, Not Mean Temperature, Is a Cue for Avian Timing of Reproduction.', *The American Naturalist*, vol. 179, no. 2, pp. E55–69.

Sharp, P.J. 1996, 'Strategies in avian breeding cycles', *Animal Reproduction Science*, vol. 42, no. 1–4, pp. 505–13.

Silverin, B., Massa, R. & Stokkhan, K.A. 1993, 'Photoperiodic adaptation to breeding at different latitudes in great tits', *General and Comparative Endocrinology*, vol. 90, no. 1, pp. 14–22.

Tsutsui, K., Ubuka, T., Bentley, G.E. & Kriegsfeld, L.J. 2012, 'Gonadotropin-inhibitory hormone (GnIH): discovery, progress and prospect.', *General and Comparative Endocrinology*, vol. 177, no. 3, pp. 305–14.

Underwood, H., Steele, C.T. & Zivkovic, B. 2001, 'Circadian organization and the role of the pineal in birds', *Microscopy Research and Technique*, vol. 53, no. 1, pp. 48–62.

Vadivelu, S.K., Kurzbauer, R., Dieplinger, B., Zweyer, M., Schafer, R., Wernig, A., Vietor, I. & Huber, L.A. 2004, 'Muscle regeneration and myogenic differentiation defects in mice lacking TIS7', *Molecular and cellular biology*, vol. 24, no. 8, pp. 3514–25.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. & Speleman, F. 2002, 'Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes', *Genome biology*, vol. 3, no. 7, p. research0034.1–0034.11.

Verhagen, I., Gienapp, P., Laine, V.N., van Grevenhof, E.M., Mateman, A.C., van Oers, K. & Visser, M.E. 2019, 'Genetic and phenotypic responses to genomic selection for timing of breeding in a wild songbird', *Functional Ecology*, vol. 0, no. 0.

Vietor, I. & Huber, L.A. 2007, 'Role of TIS7 family of transcriptional regulators in differentiation and regeneration.', *Differentiation; research in biological diversity*, vol. 75, no. 9, pp. 891–7.

Visser, M.E., Both, C. & Lambrechts, M.M. 2004, 'Global Climate Change Leads to Mistimed Avian Reproduction', *Advances in Ecological Research*, vol. 35, no. 04, pp. 89–110.

Visser, M.E., Holleman, L.J.M. & Caro, S.P. 2009, 'Temperature has a causal effect on avian timing of reproduction', *Proceedings of the Royal Society B: Biological Sciences*, vol. 276, no. 1665, pp. 2323 LP – 2331.

Visser, M.E., Holleman, L.J.M. & Gienapp, P. 2006, 'Shifts in caterpillar biomass phenology due to climate change and its impact on the breeding biology of an insectivorous bird', *Oecologia*, vol. 147, no. 1, pp. 164–72.

Visser, M.E., van Noordwijk, A.J., Tinbergen, J.M.M. & Lessells, C.M.M. 1998, 'Warmer springs lead to mistimes reproduction in Great Tits (Parus major)', *Proceedings of the Royal Society B*, vol. 265, no. May, pp. 1867–70.

Visser, M.E., Schaper, S. V., Holleman, L.J.M., Dawson, A., Sharp, P., Gienapp, P. & Caro, S.P. 2011, 'Genetic variation in cue sensitivity involved in avian timing of reproduction', *Functional Ecology*, vol. 25, no. 4, pp. 868–77.

Walzem, R.L. 1996, 'Lipoproteins and the laying hen: form follows function', *Poultry and Avian Biology Reviews*, vol. 7, no. 1, pp. 31–64.

Williams, T.D. 1994, 'Intraspecific variation in egg size and egg composition in birds: effects on offspring fitness.', *Biological reviews of the Cambridge Philosophical Society*, vol. 69, no. 1, pp. 35–59.

Williams, T.D. 2012, *Physiological adaptations for breeding in birds*.

Wingfield, J.C. 1993, 'Control of testicular cycles in the song sparrow, Melospiza melodia melodia: interaction of photoperiod and an endogenous program?', *General and Comparative Endocrinology*, vol. 92, no. 3, pp. 388–401.

Wu, X., Liu, X.T., Lan, J.J., Wan, X.P., Yan, M.J., Lian, S.Y. & Li, A. 2016, 'Investigation of insulin-like growth factor-1 gene with egg-laying traits in the Muscovy duck (Cairina moschata)', *Canadian Journal of Animal Science*, vol. 96, no. 2, pp. 203–14.

Yoshimura, T., Yasuo, S., Watanabe, M., Iigo, M., Yamamura, T., Hirunagi, K. & Ebihara, S. 2003, 'Light-induced hormone conversion of T4 to T3 regulates photoperiodic response of gonads in birds.', *Nature*, vol. 426, no. 6963, pp. 178–81.

Yoshimura, Y., Ando, M., Nagamatsu, S., Iwashita, M., Adachi, T., Sueoka, K., Miyazaki, T., Kuji, N. & Tanaka, M. 1996, 'Effects of insulin-like growth factor-I on follicle growth, oocyte maturation, and ovarian steroidogenesis and plasminogen activator activity in the rabbit.', *Biology of reproduction*, vol. 55, no. 1, pp. 152–60.

Yoshimura, Y., Iwashita, M., Karube, M., Oda, T., Akiba, M., Shiokawa, S., Ando, M., Yoshinaga, A. & Nakamura, Y. 1994, 'Growth hormone stimulates follicular development by stimulating ovarian production of insulin-like growth factor-I.', *Endocrinology*, vol. 135, no. 3, pp. 887–94.

Zheng, H., Li, H., Tan, W., Xu, C., Jia, L., Wang, D., Li, Z., Sun, G., Kang, X., Yan, F. & Liu, X. 2018, 'Oestrogen regulates the expression of cathepsin E-A-like gene through ERBeta in liver of chicken (Gallus gallus).', *Journal of genetics*, vol. 97, no. 1, pp. 145–55.

Figure legends and tables

Figure 1. **PCA was performed on normalized and subsequently transformed gene expression data from individual females for hypothalamus (A, B), ovary (C, D) and liver (E, F)**. Every data point represents an individual female. Potential clusters associated with generation are shown (panels A, C, E); green = F_1, orange = F_2, and with time point (B, D, F); green = time point 1, purple= time point 2, orange = time point 3. PCA identified two very distinct clusters separated over the first principal component (PC1) in ovary (C), which are overlapping in both hypothalamus (A) and liver (E). Time point revealed a gradient in clustering along PC2 and PC1 for ovary (D) and liver (F) respectively, but not for hypothalamus (B).

Figure 2. **Normalized and subsequently log10 transformed mRNA levels of candidate genes in *hypothalamus* in three time points.** Females from the F_1-generation (red) have significantly higher expression levels in hypothalamus for *DIO2*, *NFIL3*, *OPN5* and *TRH* compared to the F_2-generation (blue) in the three time points, but independent of time point (Table 1). Regression lines are dashed, as not to imply significant interactions. Note that the scale of the y-axis differs per panel.

Figure 3. **Normalized and subsequently log10 transformed mRNA levels of candidate genes in *ovary* in three time points.** F_1-generation (red) females have lower expression levels for *IFRD1* (A), but higher expression levels for *VLDLR* (B). Early breeding females have lower expression of *IFRD1* in time points 2 and 3 and late laying females show and increased *VLDLR* expression in time point 2 (Table 2). Regression lines are dashed, as not to imply significant interactions. Note that the scale of the y-axis differs per panel.

Figure 4. **Normalized and subsequently log10 transformed mRNA levels of candidate genes in *liver* for three time points.** Females differ in expression of *IGF1* (A) between generations (F_1 = red, F_2 = blue), but not *VTG2* (B, no distinction is made between generations, data shown in black). Early laying females show increased expression compared to late laying females in both *IGF1* and *VTG2*, independent of time point (Table 3). Regression lines are dashed, as not to imply significant interactions. Note that the scale of the y-axis differs per panel.

Figure 5. **Normalized and subsequently log10 transformed mRNA levels of candidate genes in *liver* in three time points.** Variation in laying date, in interaction with time point, explains mRNA expression for *APOB*, *APOV1*, *BEST3* and *CTSEAL* (Table 3). In addition, APOB expression differs between generations (F_1 = red, F_2 = blue), but not *APOV1*, *BEST3* and *CTSEAL* where, therefore, no distinction between generations is made (data shown in black). Regression lines for APOB expression are dashed, as not to imply a significant three-way interaction. Note that the scale of the y-axis differs per panel.

Figure 6. **Correlation matrix of the pairwise correlations between all possible gene pairs. Positive correlations are displayed in blue and negative correlations in red**. Colour intensity and the size of the circle are proportional to the correlation coefficients (see the colour legend on the right). Only the significant ($p < 0.05$) correlations are shown. Genes *OPN5* until *VIP* are assessed in hypothalamus, *VLDLR* until *MMP15* expressed in ovary and *VTG2* until *APOV1* in liver. Please note that *HSPB1* has been assessed in both ovary and liver.

875 Table 1. The degree of variation per gene explained by laying date, time point, their interaction and
876 generation in hypothalamus, where bold *p*-values indicate significance.

877

	Laying date × time point		**Laying date**		**Time point**		**Generation**	
Gene	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*
DIO2	$F_{(2,52)} = 1.20$	0.677	$F_{(1,56)} = 2.99$	0.297	$F_{(2,54)} = 0.81$	0.677	$F_{(1,57)} = 82.52$	**<0.0001**
NFIL3	$F_{(2,52)} = 1.42$	0.677	$F_{(1,56)} = 0.02$	0.898	$F_{(2,54)} = 1.07$	0.677	$F_{(1,57)} = 58.03$	**<0.0001**
OPN5	$F_{(2,52)} = 0.70$	0.677	$F_{(1,56)} = 0.31$	0.705	$F_{(2,54)} = 0.90$	0.677	$F_{(1,57)} = 77.15$	**<0.0001**
TRH	$F_{(2,52)} = 0.42$	0.705	$F_{(1,56)} = 1.12$	0.677	$F_{(2,54)} = 0.90$	0.677	$F_{(1,57)} = 160.51$	**<0.0001**
VIP	$F_{(2,52)} = 0.40$	0.705	$F_{(2,55)} = 0.46$	0.705	$F_{(1,54)} = 0.44$	0.677	$F_{(1,57)} = 3.06$	0.297

878

879

880 Table 2. The degree of variation per gene explained by laying date, time point, their interaction and
881 generation in ovary, where bold *p*-values indicate significance.

882
883

	Laying date × timepoint		**Laying date**		**Time point**		**Generation**	
Gene	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*
AR	$F_{(2,53)} = 1.18$	0.488	$F_{(1,55)} = 0.67$	0.592	$F_{(2,55)} = 8.37$	**<0.0001**	$F_{(1,56)} = 435.59$	**<0.0001**
CYP17A1	$F_{(2,53)} = 1.91$	0.300	$F_{(1,56)} = 3.19$	0.159	$F_{(2,56)} = 9.91$	**<0.0001**	$F_{(1,55)} = 1.11$	0.480
C1D	$F_{(2,53)} = 0.96$	0.565	$F(_{1,55)} = 0.03$	0.922	$F_{(2,56)} = 11.39$	**<0.0001**	$F_{(1,56)} = 27.47$	**<0.0001**
ERα	$F_{(2,53)} = 0.41$	0.807	$F_{(1,55)} = 1.05$	0.488	$F_{(2,56)} = 18.11$	**<0.0001**	$F_{(1,56)} = 48.47$	**<0.0001**
FABP4	$F_{(2,53)} = 0.67$	0.701	$F_{(1,55)} = 0.35$	0.745	$F_{(2,56)} = 1.74$	0.327	$F_{(1,56)} = 7.70$	**0.023**
FSH-R	$F_{(2,53)} = 1.35$	0.444	$F_{(1,57)} = 0.02$	0.922	$F_{(2,55)} = 0.18$	0.917	$F_{(1,57)} = 1.78$	0.327
GnIH-R	$F_{(2,53)} = 2.81$	0.146	$F_{(1,55)} = 0.06$	0.899	$F_{(2,56)} = 0.26$	0.877	$F_{(1,56)} = 2.21$	0.276
HSPB1	$F_{(2,53)} = 3.18$	0.118	$F_{(1,55)} = 4.56$	0.094	$F_{(2,55)} = 22.98$	**<0.0001**	$F_{(1,55)} = 13.65$	**<0.0001**
HSPB7	$F_{(2,53)} = 0.30$	0.865	$F_{(1,55)} = 0.01$	0.946	$F_{(2,56)} = 3.07$	0.122	$F_{(1,56)} = 17.42$	**<0.0001**
IFRD1	$F_{(2,53)} = 3.18$	0.118	$F_{(1,55)} = 6.86$	**0.032**	$F_{(2,55)} = 25.61$	**<0.0001**	$F_{(1,56)} = 9.73$	**0.010**
MMP15	$F_{(2,53)} = 2.82$	0.146	$F_{(1,55)} = 1.95$	0.309	$F_{(2,56)} = 27.51$	**<0.0001**	$F_{(1,56)} = 346.91$	**<0.0001**
LH-R	$F_{(2,53)} = 1.12$	0.493	$F_{(1,55)} = 0.32$	0.750	$F_{(2,56)} = 5.53$	**0.023**	$F_{(1,56)} = 38.26$	**<0.0001**
PRL-R	$F_{(2,53)} = 1.15$	0.490	$F_{(1,55)} = 0.00$	0.953	$F_{(2,53)} = 0.16$	0.920	$F_{(1,55)} = 0.23$	0.779
StAR	$F_{(2,53)} = 0.75$	0.661	$F_{(1,57)} = 0.14$	0.850	$F_{(2,55)} = 1.61$	0.354	$F_{(1,57)} = 5.63$	0.057
VLDL-R	$F_{(2,53)} = 0.30$	0.865	$F_{(1,55)} = 13.25$	**<0.001**	$F_{(2,55)} = 5.13$	**0.027**	$F_{(1,55)} = 63.70$	**<0.0001**
ZP4	$F_{(2,35)} = 3.60$	0.089	$F_{(1,58)} = 0.01$	0.933	$F_{(2,57)} = 2.74$	0.150	$F_{(1,53)} = 12.27$	**0.004**

884 Table 3. The degree of variation per gene explained by laying date, time point, their interaction and
885 generation in liver, where bold *p*-values indicate significance.

	Laying date × timepoint		**Laying date**		**Time point**		**Generation**	
Gene	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*
APOB	$F_{(2,53)} = 5.10$	**0.029**	$F_{(1,58)} = 1.58$	0.285	$F_{(2,57)} = 3.36$	0.079	$F_{(1,53)} = 17.98$	**<0.0001**
APOV1	$F_{(2,53)} = 11.58$	**<0.0001**	$F_{(1,58)} = 4.97$	0.063	$F_{(2,57)} = 39.09$	**<0.0001**	$F_{(1,53)} = 4.75$	0.068
BEST3	$F_{(2,53)} = 6.53$	**0.010**	$F_{(1,58)} = 1.75$	0.264	$F_{(2,57)} = 84.47$	**<0.0001**	$F_{(1,53)} = 3.57$	0.115
CTSEAL	$F_{(2,53)} = 7.21$	**0.007**	$F_{(1,58)} = 2.33$	0.203	$F_{(2,57)} = 103.80$	**<0.0001**	$F(_{1,53)} = 0.01$	0.954
GR	$F_{(2,53)} = 2.39$	0.167	$F_{(1,55)} = 2.30$	0.203	$F_{(2,56)} = 1.90$	0.229	$F_{(1,55)} = 35.29$	**<0.0001**
HSPB1	$F_{(2,53)} = 0.51$	0.658	$F_{(1,56)} = 0.41$	0.589	$F_{(2,56)} = 0.86$	0.548	$F_{(1,55)} = 18.57$	**<0.0001**
IGF1	$F_{(2,53)} = 0.72$	0.573	$F_{(1,55)} = 6.53$	**0.032**	$F_{(2,55)} = 0.75$	0.570	$F_{(1,55)} = 6.77$	**0.032**
MR	$F_{(2,53)} = 0.34$	0.757	$F_{(1,57)} = 3.32$	0.126	$F_{(2,55)} = 0.83$	0.548	$F_{(1,57)} = 9.01$	**0.013**
VTG2	$F_{(2,53)} = 3.90$	0.063	$F_{(1,58)} = 6.63$	**0.032**	$F_{(2,57)} = 56.00$	**<0.0001**	$F_{(1,55)} = 0.00$	0.993

886

Appendix 1 – explaining variation in reproductive phenotypes

Differences in laying dates and largest follicle widths between selection lines and treatments were tested by performing ANOVA, with laying dates or log10 transformed follicle widths as dependent variable and 'selection line' (i.e. genomic phenotype), 'treatment' (i.e. warm and cold temperature environment, see "Experimental setup" in Materials and methods) and 'generation' as explanatory variables, with 'time point' added as extra variable in testing differences in follicle widths.

Laying dates (first breeding season) ranged from January date 99 (9 April) to 159 (8 June). Treatment ($F_{1,95}$ = 0.91, p = 0.34), selection line ($F_{1,95}$ = 0.08, p = 0.78), or their interaction ($F_{1,94}$ = 0.02, p = 0.90) did not explain variation in laying dates. We did, however, find an effect of generation ($F_{2,96}$ = 3.38, p = 0.04); a difference in mean laying date of ~7.5 days between the F_1- and F_2-generation ($t_{61.99}$ = -2.50, p = 0.02) (Figure A1.1).

Also, we found no effect of selection line ($F_{1,55}$ = 0.27, p = 0.36) or treatment ($F_{1,56}$ = 0.27, p = 0.61), or their interaction ($F_{2,50}$ = 0.03, p = 0.86) on the width of the largest follicle (second breeding season). Follicles were larger for the F_1-generation ($F_{1,58}$ = 7.24, p = 0.01) and increased over time ($F_{2,59}$ = 32.68, p < 0.0001), with the largest follicles measured in time point 3 (p < 0.0001 for both comparisons with time points 2 and 3, Figure A1.2).

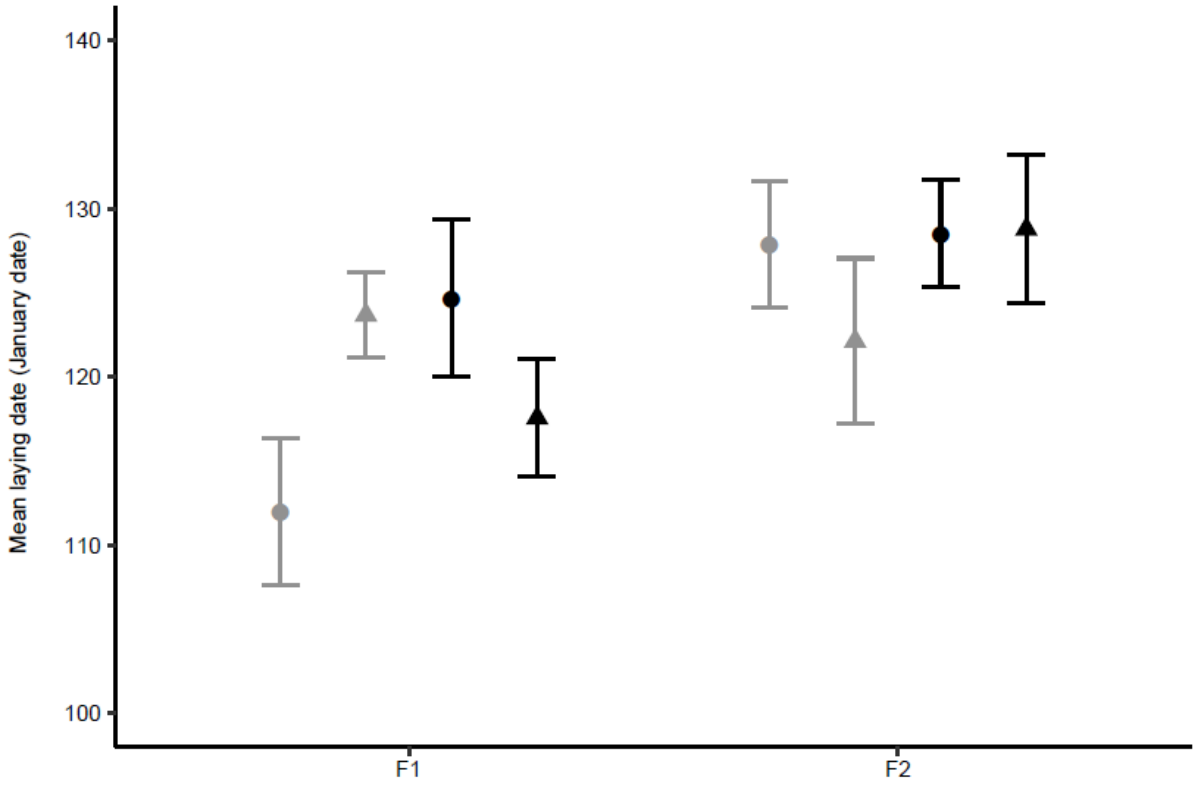
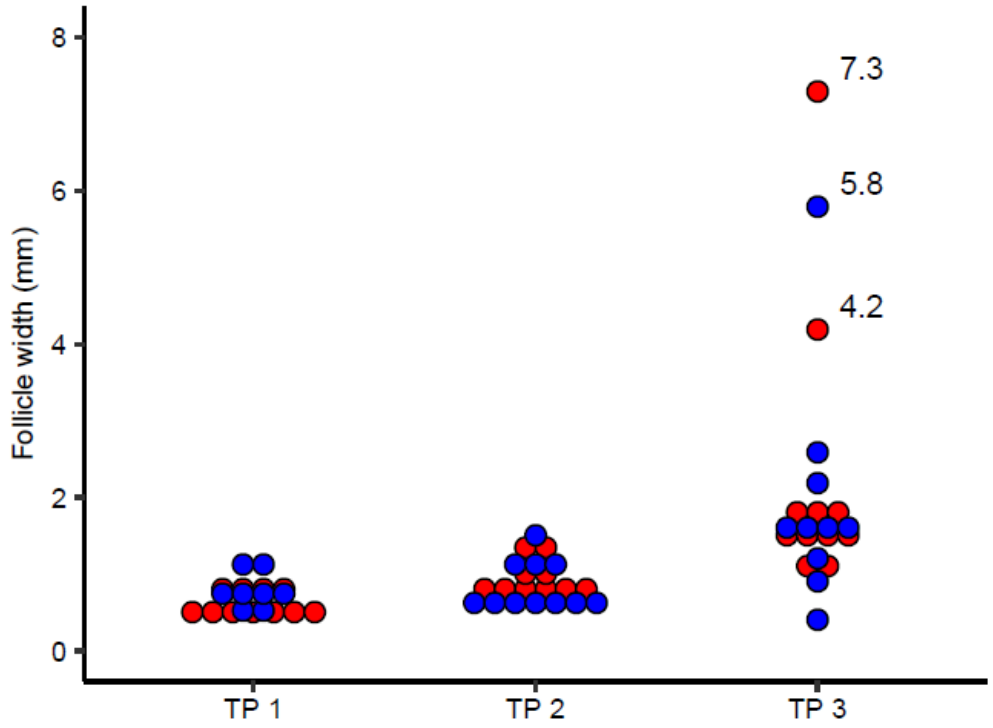


Figure A1.1. Mean laying dates (mean ± SEM) in January dates (y-axis, 1 = 1 January, 100 = 10 April etc.) for the F_1- and F_2-generation (x-axis) shown for females based on their selection line × treatment groups. No significant differences were found in mean laying date between early (dark grey) and late (black) selection line females, or the warm (triangles) and cold (circles) treatment, nor their interaction. We adjusted the horizontal position of the data shown to prevent overlap and so facilitate clarity of the graph.



910

911 Figure A1.2. Follicle widths of the largest follicles for the three different time points (TP 1 = time
912 point 1, TP 2 = time point 2, TP 3 = time point 3) from the females of the F_1 (red) and F_2 (blue)
913 generation.

914

Appendix 2 – reference gene validation for RT-qPCR data normalization

For proper gene expression analysis, the data needs normalization against the expression level of a set of stable reference genes. This approach controls for factors such as the amount of cDNA load in a sample, variations in the efficiency of the RT reaction, and RNA quality (Vandesompele et al., 2002). Optimal reference genes, exhibit stable expression levels that are not influenced by the (experimental) condition. Ideally, a set of multiple reference genes is compilated, for which then the normalization factor (NF) is calculated (Vandesompele et al., 2002).

Absolute amounts of reference gene cDNA were calculated by converting the C_t values ($C \times E^{-Ct}$, with $C = 10^{10}$ and $E = 2$) (Dijk, Kraal-Muller, & Kamphuis, 2004). Then, the measure of reference gene expression stability (M) was calculated in the application geNorm. M is defined as the average pairwise variation ($V_{n/n+1}$) between the normalization factors (NF_n and NF_{n+1}) of a particular reference gene with all other reference genes (Vandesompele et al., 2002).

There is a cut-off point of $V_{n/n+1} = 0.15$, below which it is not necessary to include an additional reference gene for normalization (Vandesompele et al., 2002). However, we did not take this cut-off point too strictly. When all $V_{n/n+1}$-values are slightly >0.15, but the NF_n and NF_{n+1} show a high correlation, we decided to stop adding reference genes. Using at least three reference genes with highly correlated expression levels is already a significant improvement opposed to the common practise of using a single gene.

Hypothalamus

The genes *PRKCA*, *RPL19* and *SDHA* were selected as potential reference genes for hypothalamus samples. All three showed an M < 1.5 according to geNorm, with M_{RPL19} and $M_{PRKCA} = 0.632$ and $M_{SDHA} = 0.691$. Further analysis resulted in $V_{PRKCA\text{-}RPL19\text{-}SDHA} = 0.217$, which is above the recommend V = 0.15 (see above). Close inspection of the data lead to three individual samples having strongly deviating amounts, both in the reference gene dataset and the candidate gene dataset indicating decreased cDNA quality in these samples. These individuals were therefore removed from the dataset and the remaining data rerun in geNorm. Stability slightly increased (M_{RPL19} and $M_{SDHA} = 0.583$, $M_{PRKCA} = 0.603$) and V dropped closer to the cut-off point ($V_{PRKCA\text{-}RPL19\text{-}SDHA} = 0.179$). We found very low variation ($R^2 = 0.980$, Figure A2.1) between NF_2 and NF_3, meaning that addition of the third reference gene does not add much to the overall normalization. Nevertheless, we use these three reference gene for normalization of mRNA expression data in the hypothalamus.

Ovary

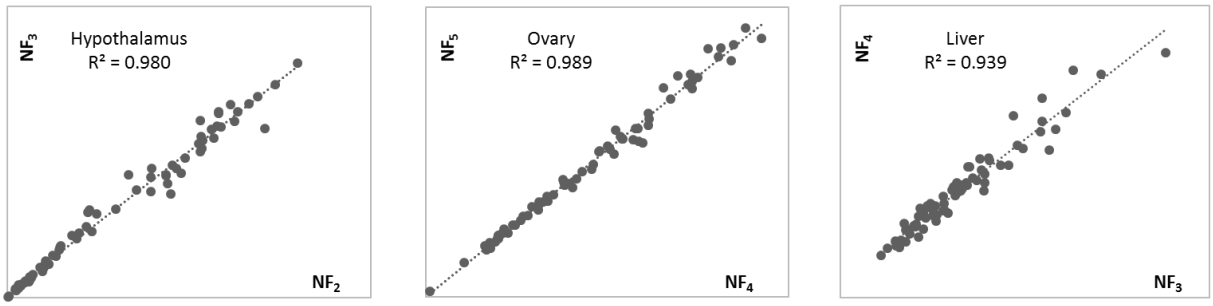
We started with *HPRT*, *PRKCA* and *YWHAZ* as potential reference genes for ovary samples. All three reference genes showed an M < 1.5, with M_{HPRT} and $M_{YWHAZ} = 0.451$ and $M_{PRKCA} = 1.0262$. Further analysis of V in geNorm resulted in $V_{HPRT\text{-}PRKCA\text{-}YWHAZ} = 0.431$, which is far above the recommended V = 0.15 (see above). Therefore, we ran two extra potential reference gene, *RPL19* and *PRL13*, in order to decrease V. Analysis in geNorm resulted in M_{PRKCA} and $M_{RPL19} = 0.583$, $M_{RPL13} = 0.741$, $M_{HPRT} = 1.010$ and $M_{YWHAZ} = 1.038$. Using these five reference genes abled us to reduce V ($V_{HPRT\text{-}PRKCA\text{-}RPL13\text{-}RPL19\text{-}YWHAZ} = 0.189$). Here, V is still >0.15, but we find very low variation ($R^2 = 0.989$, Figure A2.1) between NF_4 and NF_5 and decided to not add a sixth reference gene for normalization of mRNA expression data in the ovary.

Liver

We started with *PRKCA*, *RPL19* and *SDHA* as potential reference genes for liver samples. All three reference genes showed an M<1.5, with M_{RPL19} and $M_{SDHA} = 0.606$ and $M_{PRKCA} = 0.830$. Further analysis of V in geNorm resulted in $V_{RPL19\text{-}SDHA\text{-}PRKCA} = 0.298$, which is far above the recommended V = 0.15 (see above). Therefore, we ran an extra potential reference gene, *B2M*, in order to decrease V. In addition, one individual was removed from the dataset due to strongly deviating amounts. Analysis of the four reference genes in geNorm resulted in all genes showing M<1.5 (M_{RPL19} and M_{SDHA} =

962 0.606, M_{PRKCA} = 0.650 and M_{B2M} = 0.677) and a decreased V ($V_{B2M\text{-}PRKCA\text{-}RPL19\text{-}SDHA}$ = 0.149). Here,
963 $V_{B2M\text{-}PRKCA\text{-}RPL19\text{-}SDHA\text{-}PRKC}$ < 0.15 and together with the high R^2 (Figure A2.1) found between NF_3 and
964 NF_4, adding a fifth reference gene is not necessary for accurate normalization.

965



966

967 Figure A2.1. Scatterplots for hypothalamus, ovary and liver of normalization factors before (x-axis)
968 and after (y-axis) the addition of a reference gene. The high R^2 indicates that the inclusion of a second,
969 fourth and third reference gene is not necessary for hypothalamus, ovary and liver respectively. But,
970 because V was still slightly above 0.15 for hypothalamus and ovary and V < 0.15 after adding the
971 fourth reference gene in liver, we did add the extra reference gene.

972

Appendix 3 – *Validation for linking breeding season 1 to breeding season 2*

The largest follicle width measured during the second breeding season was 7.3mm and measured on the day that this females should have laid her 3rd egg in the first breeding season. Therefore, this width is taken as a measure for a fully developed follicle (F1-follicle). Note: in this appendix, F1, F2 etc. is used in relation to follicle order/size and not generation, as elsewhere. We back calculated the approximate follicle sizes (Table A3.1), as we did not measure individual size differences of the F5, F4, F3 and F2 follicles, by using the traditional hierarchical model of follicle development (Astheimer & Grau, 1990). This model predicts that the first follicle to enter rapid yolk development (RYD) is the first to ovulate and first to be laid. Hepatic production of vitellogenin (VTG) and very low density lipoprotein (VLDL), both yolk-targeted lipoproteins, are essential for vitellogenesis (i.e. yolk formation through nutrient deposition in the oocyte) and oocyte growth (Bacon, Musser, & Brown, 1974; Walzem, 1996). We found increasing follicle widths over time (Figure S4), with three females carrying follicles similar to F3-F1 approximate sizes (Table A3.1) and five females likely to have entered RYD and others close. We found VTG2 mRNA expression reflecting individual differences in egg-laying ($F_{1,58} = 6.625$, $p = 0.032$) and increasing over time ($F_{2,57} = 56$, $p < 0.0001$). In addition, we found a significant relationship between laying dates and follicle widths, especially in time point 3 (Figures S5).We are therefore confident that the mRNA expression levels from the second breeding season are representative of the phenotypes (i.e. laying dates) recorded and assume these breeding seasons to be similar.

Table A3.1. Calculation of approximate size for the F5 - F1 follicles based on the traditional hierarchical model of follicle development (Astheimer and Grau, 1990). First the radius was calculated by dividing the width (or diameter). Subsequently, the radius was used to calculate the volume by the formula of a sphere ($\frac{4}{3}\pi r^3$)

Follicle	Percentage of F1 follicle	Diameter (mm)	Radius (mm)	Volume (mm3)
F1	100	7.30	3.65	203.688
F2	64	6.30	3.15	130.361
F3	30	4.88	2.44	61.107
F4	10	3.38	1.69	20.369
F5	3	2.36	1.13	6.111

References used in appendices

Astheimer, L. E. E. B., & Grau, C. R. (1990). A comparison of yolk growth rates in seabird eggs. *Ibis*, *132*(3), 380–394. https://doi.org/10.1111/j.1474-919X.1990.tb01057.x

Bacon, W. L., Musser, M. A., & Brown, K. I. (1974). Plasma free fatty acid and neutral lipid concentrations in immature, laying and broody turkey hens. *Poultry Science*, *53*(3), 1154–1160. https://doi.org/10.3382/ps.0531154

Dijk, F., Kraal-Muller, E., & Kamphuis, W. (2004). Ischemia-induced changes of AMPA-type glutamate receptor subunit expression pattern in the rat retina: a real-time quantitative PCR study. *Investigative Ophthalmology & Visual Science*, *45*(1), 330–341. https://doi.org/10.1167/iovs.03-0285

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, *3*(7), research0034.1–0034.11. https://doi.org/10.1186/gb-2002-3-7-research0034

Walzem, R. L. (1996). Lipoproteins and the laying hen: form follows function. *Poultry and Avian Biology Reviews*, *7*(1), 31–64.